



CERIS

CENTRE FOR ECONOMIC RESEARCH
ON INCLUSIVITY AND SUSTAINABILITY

On the Frontlines:

**An exploratory analysis of unequal exposure to air pollution and COVID-19
in the United States**

Working Paper Series 2021/07

Cite as: Vila, D and McDermott TKJ (2021). On the Frontlines: An exploratory analysis of unequal exposure to air pollution and COVID-19 in the United States. Centre for Economic Research on Inclusivity and Sustainability (CERIS) Working Paper Series, 2021/07.

On the Frontlines

An exploratory analysis of unequal exposure to air pollution and COVID-19 in the United States

Daniel Vila¹ Thomas K.J. McDermott¹

1: Centre for Economic Research on Inclusivity and Sustainability (CERIS)
J.E. Cairnes School of Business and Economics, National University of Ireland Galway

September 29, 2021

Recent literature has suggested a link between poor air quality and worse COVID-19 outcomes. In the United States, this link is particularly noteworthy because of residential sorting along ethnic lines within the US population; minorities are disproportionately exposed to health hazards, including air pollution. The impacts of the COVID-19 pandemic have also been disproportionately concentrated amongst minorities. We explore the association between air quality and COVID-19 outcomes, using county level data for the United States from the first wave of the pandemic in 2020, and test whether exposure to more polluted air can account for some of the observed disparities in COVID-19 outcomes among minorities.

Keywords: COVID-19, air quality, race, inequality

Introduction

During the first wave of the COVID-19 outbreak in the United States (US), Black and Hispanic people had a higher COVID-19 death rate. Both academic literature (Millett et al. (2020)) and journalistic reporting (Oppel et al. (2020)) claims that this can be explained at least partially by a higher prevalence of underlying health conditions (such as diabetes, obesity, and asthma) in these populations. However, these underlying conditions are themselves often a result of differences in exposure to environmental hazards. Non-white populations, especially Black populations, are typically exposed to higher levels of air pollution in the US (American Lung Association (2020)). According to the American Lung Association, certain communities are also disproportionately exposed to toxic and hazardous waste. These communities most often reside in urban settings, have low socioeconomic status, and include a large proportion of ethnic minorities (American

Lung Association (2001)). This is relevant to an analysis of COVID-19 because of the airborne and respiratory nature of the virus. Indeed previous studies have demonstrated a link between local air pollution and COVID-19 outcomes (Wu et al. (2020); Cole et al. (2020); Travaglio et al. (2020); Conticini et al. (2020)).

Other recent academic literature highlights that COVID-19 has disproportionately impacted racial and ethnic minorities and socioeconomically disadvantaged communities in the US (Pierce et al. (2021); Wadhwa et al. (2020); Price-Haywood et al. (2020); Berkowitz et al. (2020)). For example, using census tract level data from the city of Chicago, Pierce et al. (2021) conclude that communities with larger populations of Non-Hispanic Black and Hispanic individuals experienced not just higher mortality, but also markedly higher loss of potential life due to COVID-19. Historical discrimination in housing policies, and residential sorting along ethnic lines, has resulted in particular groups experiencing long-standing neighborhood disadvantage that includes a number of aspects relevant to COVID-19 outcomes, such as crowded housing, food insecurity, employment instability, higher rates of uninsured residents, and increased reliance on crowded public transit (Pierce et al. (2021); Berkowitz et al. (2020); Egede & Walker (2020)).

We add a more explicit environmental angle to this emerging discussion on structural inequalities and COVID-19, by exploring the relationship between COVID-19 outcomes, air quality, and race directly. Using county level data for the US, we specifically examine whether or not higher levels of air pollution are associated with worse COVID-19 outcomes in the US, and if this can account for worse COVID-19 outcomes among minorities.

Data

We construct a dataset at the county level for the continental US, combining information on COVID-19 outcomes, socio-demographic data, air pollution and workplace attendance for some 2,757 counties, for the period 01 January 2020 to 10 June 2020, i.e. the first wave of the pandemic.

Data on COVID-19 outcomes including deaths and confirmed cases, as well as socio-demographic data, all observed at the county level, are from the Johns Hopkins University Center for Systems Science and Engineering dataset (JHU CSSE) (Dong et al. (2020)). The mean number of cases per county over this period was 600, with a max of 80,204 in Cook County, Illinois. The mean number of deaths per county was 31, with a max of 3,780 also in Cook County, Illinois.¹

For air pollution, we obtain data on NO₂ concentrations from NASA’s Goddard Earth Sciences Data and Information Services Center (GES DISC). NO₂ is primarily produced by burning fuel. As such, local sources of NO₂ pollution include power plants, cars, trucks, buses, and other combustion engines. According to the US EPA, breathing air with a high concentration of NO₂ can irritate airways in the human respiratory

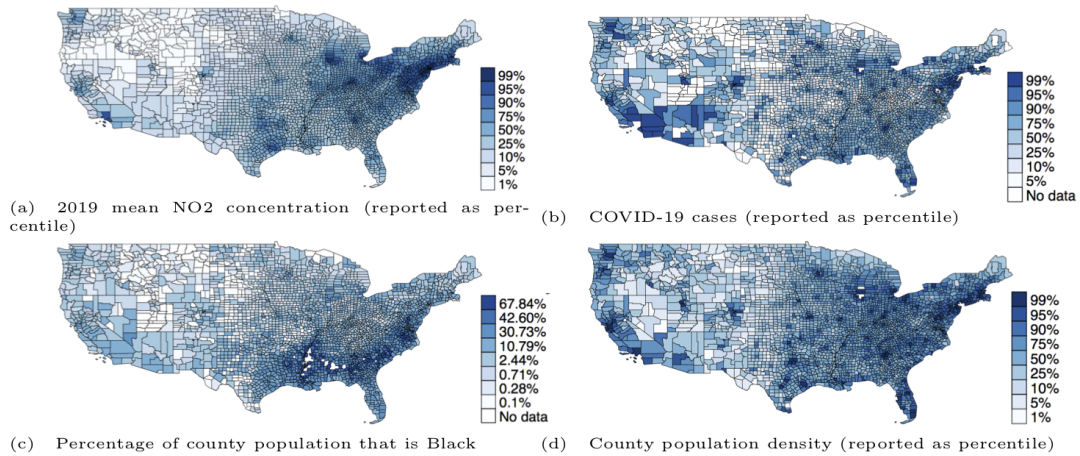
¹Full summary statistics for the main variables used in our analysis are included in Appendix Table A1.

system. Such exposures over short periods can aggravate respiratory diseases, particularly asthma, leading to respiratory symptoms (such as coughing, wheezing or difficulty breathing), hospital admissions and visits to emergency rooms. Longer exposure to elevated concentrations of NO_2 may contribute to the development of asthma and potentially increase susceptibility to respiratory infections (Moshhammer et al. (2020)). We take the average NO_2 concentration over the period January-June for 2019, as a measure of baseline (pre-pandemic) exposure to air pollution. We then merge this data with county boundary shapefiles using GIS software to obtain county level estimates for NO_2 concentrations. For ease of interpretation we normalise the NO_2 concentrations to have mean 0 and standard deviation of 1.

Finally, data on workplace attendance is based on Google’s COVID-19 Community Mobility Reports (Google LLC (2020)). These reports contain daily observations of workplace attendance at the county level, which we aggregate to obtain a measure of the average change over the period, relative to the same period in 2019. Not surprisingly, all counties experienced a decline in workplace mobility, with an average 16% decline relative to the pre-COVID-19 baseline.

Figure 1 displays maps at the county level for NO_2 concentration, COVID-19 cases, percentage of county population that is Black, and population density, in our sample. Evident from these maps is the geographic concentration of NO_2 concentrations and cases, both of which are associated primarily with high density urban areas.

Figure 1



Methods

Our aim is to investigate if counties with a higher share of Black or Hispanic population experienced worse COVID-19 outcomes, and the extent to which these outcomes may be related to exposure to poor air quality. We start by estimating models of the number of COVID-19 related deaths as the outcome variable. Given the over-dispersion

of the outcome variable (county level deaths has a mean of 31 and a standard deviation of 166 in our sample period) and the relatively large number of zeros (nearly 1200 counties had zero Covid deaths), zero inflated negative binomial (ZINB) regressions were used to estimate these models. The ZINB model is a combination of a binary model of the zero distribution, and a count model of the non-zero data, allowing for the possibility that there are distinct processes in the data leading to zeros and non-zeros in the outcome variable (Silva et al. (2015)). In our case, this corresponds to counties that experienced zero Covid deaths during our sample period. Many of these counties also experience very few or even no Covid cases over the period.

In addition to mortality outcomes, we are also interested in examining exposure to the virus, and how this has varied across counties in our data. To do so, we estimate models with cases per 1,000 of population as the outcome, using ordinary least squares regression. Finally, we also examine differences in rates of attendance at work across counties, again using ordinary least squares regressions with workplace mobility as the outcome of interest.

Results

The results of our regression analysis are presented in Table 1.² In the first column, we see that controlling for the number of confirmed cases, the number of deaths from COVID-19 was higher in counties with higher shares of Black or Hispanic population, as well as in counties with higher rates of poverty. In the second column, where we include NO₂ concentration, we see that the number of deaths from COVID-19 is also higher in counties with worse air quality. However, the inclusion of the NO₂ variable does not appear to reduce the size or precision of the estimated coefficients on the three socio-demographic variables. Contrary to our initial hypothesis, this suggests that air quality is not the primary mechanism behind worse Covid outcomes for minority groups.

In each of these regressions we also control for population density and the share of county population over the age of 65. In additional specifications not reported, we included a range of further controls, including the number of intensive care unit beds per capita, rate of mortality from respiratory disease (2018), and the percentage of county population without health insurance, or who smoke. In each case the results on the main variables of interest remain qualitatively unchanged.

A higher count of COVID-19 deaths could reflect greater exposure to the virus, or higher mortality conditional on contracting the virus, or some combination of these. However, we found no statistically significant differences in COVID-19 mortality, conditional on exposure to the virus, for counties with higher shares of Black or Hispanic populations, or higher rates of poverty.³ This suggests that whatever explains the significant differences in the count of COVID-19 deaths by county level socio-demographic

²Table 1 presents a summary of key findings from our exploratory regression analysis. Results from the first stage “inflation model” for the mortality outcomes are included in the Appendix results Table A2.

³These results not reported, but available on request.

Table 1

	(1)	(2)	(3)	(4)	(5)
	ZINB	ZINB	OLS	OLS	OLS
	Deaths	Deaths	Cases/pop.	Workplace attd.	Workplace attd.
Black pop. as percent of total	0.348*** (0.0839)	0.393*** (0.107)	0.0725** (0.0294)	0.0494*** (0.0171)	0.0825*** (0.0247)
Hispanic pop. as percent of total	0.883*** (0.144)	1.223*** (0.220)	0.0429 (0.0304)	-0.0110 (0.0135)	-0.0111 (0.0132)
Percentage of pop. below fed. poverty line	0.901*** (0.240)	1.430*** (0.356)	-0.0397 (0.0654)	0.191*** (0.0462)	0.197*** (0.0455)
2019 NO2 Concentration		11.21*** (2.364)	2.759*** (0.455)		
Cases/pop.				-0.0654** (0.0323)	0.000755 (0.0480)
Cases/pop. * Black pop. as percent of total					-0.00516** (0.00225)
Constant			4.636** (1.841)	-12.92*** (0.911)	-13.28*** (0.928)
Observations	2,756	2,741	2,741	2,724	2,724
R-squared			0.450	0.503	0.509

Notes:

attd.: attendance

Standard errors in parentheses.

All regressions include (log) population density as an additional control.

Columns (1) and (2) include controls for the percent of county population over age 65 and the number of COVID-19 cases.

Cases/pop. is the number of cases per 1,000 of county population.

Regressions 3, 4, and 5 are also analytically weighted by the total pop. and clustered by state.

*** p<0.01, ** p<0.05, * p<0.1

characteristics, is driven by differences in exposure to the virus, as opposed to differences in outcomes conditional on exposure.

In Column (3) of Table 1, where the outcome is confirmed cases per 1,000 of population, the results show that counties with a higher share of Black population experienced a higher incidence of COVID-19 cases (relative to population), controlling for population density. We also find that the incidence of COVID-19 is higher in counties with worse air pollution.

Finally, in Columns (4) and (5), we test for differences in attendance at work. As expected, in Column (4) we find a negative association with COVID-19 cases, suggesting that where the virus is more prevalent, people engage in avoidance behaviour in the form of less attendance at work. We also find that in counties with a higher share of Black population, attendance at work was significantly higher during the first wave. However, in Column (5) the negative coefficient on the interaction between cases and share of Black population indicates that when infection rates are high, avoidance behaviour in counties with high share of Black population is stronger than average.

Discussion

Existing literature has demonstrated a link between air quality and COVID-19 mortality, including in the context of the US (Wu et al. (2020)). Minority, and particularly Black, populations in the US are typically exposed to worse air quality. The effects of COVID-19 have also been disproportionately concentrated amongst minorities.

In our analysis, we explore the association between air quality and COVID-19 outcomes, using county level data for the United States from the first wave of the pandemic in 2020, and test whether exposure to more polluted air can account for some of the observed disparities in COVID-19 outcomes among minorities.

Our findings show that counties with worse air quality have more Covid cases, and higher mortality. We also find that counties with a higher share of Black population experienced both higher cases and higher mortality from COVID-19 during the first wave. However, these two effects – minorities and air quality – appear to operate independently; that is, at least at the county level, bad air quality appears not to be the primary mechanism behind worse COVID-19 outcomes for minority groups.

Further analysis suggests that counties with a higher share of Black population experienced a higher incidence of the disease, but not necessarily worse outcomes conditional on exposure. This led us to explore differences in exposure. In general people try to avoid the disease, and this is evident in our data – attendance at work is lower in counties with a higher incidence of disease. But we also found that attendance at work is higher for counties with higher minority population share. This could be one of the potential mechanisms behind higher exposure to the disease amongst ethnic minorities.

If people generally prefer to avoid risks, higher exposure to the disease presumably represents some form of constrained choice – for example related to observed differences in the capacity to work from home during the pandemic (see e.g. Bloom et al. (2021)) – or incomplete information about the risk. However, our finding that avoidance behaviour (reduced attendance at work in response to a high incidence of the disease) is if anything stronger in counties with high minority share, suggests that higher exposure among minorities is not a result of lower awareness or concern about the disease among these populations.

An important limitation to our analysis is the spatial resolution of using county level data, which may miss more local neighborhood level effects. In particular, identifying the full causal relationships between these inter-related and overlapping factors would require higher resolution data on COVID-19 outcomes, ideally at a neighborhood level. This we leave to future research.

References

- American Lung Association (2001), ‘Urban air pollution and health inequities: a workshop report.’, *Environmental Health Perspectives* **109**(suppl 3), 357–374.
- American Lung Association (2020), ‘Disparities in the impact of air pollution’.
URL: <https://www.lung.org/clean-air/outdoors/who-is-at-risk/disparities>
- Berkowitz, R. L., Gao, X., Michaels, E. K. & Mujahid, M. S. (2020), ‘Structurally vulnerable neighbourhood environments and racial/ethnic covid-19 inequities’, *Cities & Health* pp. 1–4.
- Bloom, N., Davis, S. J. & Zhestkova, Y. (2021), Covid-19 shifted patent applications toward technologies that support working from home, *in* ‘AEA Papers and Proceedings’, Vol. 111, pp. 263–66.
- Cole, M., Ozgen, C. & Strobl, E. (2020), ‘Air pollution exposure and covid-19’.
- 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* p. 114465.
- Dong, E., Du, H. & Gardner, L. (2020), ‘An interactive web-based dashboard to track covid-19 in real time’, *The Lancet infectious diseases* **20**(5), 533–534.
- Egede, L. E. & Walker, R. J. (2020), ‘Structural racism, social risk factors, and covid-19—a dangerous convergence for black americans’, *New England Journal of Medicine* **383**(12), e77.
- Google LLC (2020), ‘Google covid-19 community mobility reports’.
URL: <https://www.google.com/covid19/mobility/>
- Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., Honermann, B., Lankiewicz, E., Mena, L., Crowley, J. S. et al. (2020), ‘Assessing differential impacts of covid-19 on black communities’, *Annals of Epidemiology* .
- Moshhammer, H., Poteser, M., Kundi, M., Lemmerer, K., Weitensfelder, L., Wallner, P. & Hutter, H.-P. (2020), ‘Nitrogen-dioxide remains a valid air quality indicator’, *International Journal of Environmental Research and Public Health* **17**(10), 3733.
- Oppel, R. A., Gebeloff, R., Lai, K. R., Wright, W. & Smith, M. (2020), ‘The fullest look yet at the racial inequity of coronavirus’, *The New York Times* **5**.
- Pierce, J. B., Harrington, K., McCabe, M. E., Petito, L. C., Kershaw, K. N., Pool, L. R., Allen, N. B. & Khan, S. S. (2021), ‘Racial/ethnic minority and neighborhood disadvantage leads to disproportionate mortality burden and years of potential life lost due to covid-19 in chicago, illinois’, *Health & place* **68**, 102540.

- Price-Haywood, E. G., Burton, J., Fort, D. & Seoane, L. (2020), ‘Hospitalization and mortality among black patients and white patients with covid-19’, *New England Journal of Medicine* **382**(26), 2534–2543.
- Silva, J. M. S., Tenreiro, S. & Windmeijer, F. (2015), ‘Testing competing models for non-negative data with many zeros’, *Journal of Econometric Methods* **4**(1), 29–46.
- Travaglio, M., Yu, Y., Popovic, R., Leal, N. S. & Martins, L. M. (2020), ‘Links between air pollution and covid-19 in england’, *medRxiv* .
- Wadhwa, R. K., Wadhwa, P., Gaba, P., Figueroa, J. F., Maddox, K. E. J., Yeh, R. W. & Shen, C. (2020), ‘Variation in covid-19 hospitalizations and deaths across new york city boroughs’, *Jama* **323**(21), 2192–2195.
- 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’, *MedRxiv* .

Appendix

Table A1

	Mean	Std. Dev.	Min.	Max.
Deaths	31.40261	165.8946	0	3780
Cases	599.9804	2909.401	0	80,204
Mean Workplace Attendance	-16.0958	5.4417	-58.33333	-3.614035
Black pop. as percent of total	9.579032	14.58247	0	87.41228
Hispanic pop. as percent of total	9.290927	13.71558	0	99.06877
Percentage of pop. below fed. poverty line	15.69659	6.216326	2.3	55.1
2019 NO2 Concentration	1.21e-10	1	-1.747016	7.288707
Percentage of pop. over age of 65	14.65527	7.685659	0	55.59633
Natural log of population density	3.176304	1.50806	-3.148377	10.23119
Respiratory illness mortality rate per 100,000 (2018)	112.3384	58.60863	0	574.4
Population	114,264	344,241.1	1605	10,100,000
Cases/pop.	.0035391	.0063653	0	.1456179
ICU beds per capita	.0002558	.0002703	0	.0041118

Table A2

	(1) (ZINB) Deaths	(2) (ZINB) Deaths
Marginal Effects		
Black pop. as percent of total	0.348*** (0.0839)	0.393*** (0.107)
Hispanic pop. as percent of total	0.883*** (0.144)	1.223*** (0.220)
Percentage of pop. below fed. poverty line	0.901*** (0.240)	1.430*** (0.356)
2019 NO2 Concentration		11.21*** (2.364)
Inflation Model (Cases)	-0.0594*** (0.007)	-0.0584*** (0.007)
Observations	2,756	2,741
R-squared		

Notes:

Standard errors in parentheses.

Additional controls: (Log) of population density and percent of pop. over 65.

*** p<0.01, ** p<0.05, * p<0.1