I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13480

Does the COVID-19 Pandemic Improve Global Air Quality? New Cross-National Evidence on Its Unintended Consequences

Hai-Anh H. Dang
Trong-Anh Trinh

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13480

Does the COVID-19 Pandemic Improve Global Air Quality? New Cross-National Evidence on Its Unintended Consequences

Hai-Anh H. Dang
World Bank, IZA, Indiana University and Vietnam Academy of Social Sciences

Trong-Anh Trinh
World Bank

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Does the COVID-19 Pandemic Improve Global Air Quality? New Cross-National Evidence on Its Unintended Consequences*

Despite a growing literature on the impacts of the COVID-19 pandemic, scant evidence currently exists on its impacts on air quality. We offer the first study that provides cross-national evidence on the causal impacts of COVID-19 on air pollution. We assemble a rich database consisting of daily, sub-national level data of air quality for 178 countries before and after the COVID-19 lockdowns, and investigate their impacts on air quality using a Regression Discontinuity Design approach. We find the lockdowns to result in significant decreases in global air pollution. These results are consistent across measures of air quality and data sources and robust to various model specifications. Some limited evidence emerges that countries with a higher share of trade and manufacturing in the economy or with an initially lower level of air pollution witness more reduced air pollution after the lockdowns; but the opposite result holds for countries near the equator. We also find that mobility restrictions following the lockdowns are a possible explanation for improved air quality.

JEL Classification: D00, H00, O13, Q50

Keywords: COVID-19, air pollution, mobility restriction, regression discontinuity design

Corresponding author:
Hai-Anh H. Dang
Data Production & Methods Unit
Development Data Group
World Bank
1818 H St. N.W.
Washington, D.C. 20433
USA

E-mail: hdang@worldbank.org

* We would like to thank Dean Jolliffe and Paolo Verme for helpful discussions on earlier versions. Dang would also like to thank the UK Department of International Development for additional funding assistance through a Knowledge for Change (KCP) grant for the World Development Report 2021 "Data for Better Lives". Trinh acknowledges support from the World Bank Young Fellows Program in Forced Displacement.

1. Introduction

It has by now become clear that the COVID-19 pandemic is not only a global health emergency but has also led to a major global economic downturn. An emerging body of economic literature has examined the impacts of COVID-19 on a wide range of outcomes including unemployment (Fairlie et al., 2020), household consumption (Baker et al., 2020), and individual income (loss) and behaviour changes for the whole population or for different income groups (Akesson et al., 2020; Dang et al., 2020). Most studies generally confirm the adverse effects of the pandemic on these various outcomes in richer and poorer countries alike.

Yet, scant evidence currently exists on the impacts of the COVID-19 crisis on air quality, and among the few existing studies there appears no conclusive evidence. Employing a difference-in-difference model that compare Chinese cities with and without the pandemic-induced lockdown policies, He et al. (2020) find that city lockdowns led to considerable improvement in air quality as measured by Air Quality Index (AQI) and $PM_{2.5}$. This result is consistent with Brodeur et al.'s (2020) findings for the United States that 'safer-at-home' policies decreased PM2.5 emissions. Research in other disciplines such as environment studies also suggest a considerable decline in pollutant parameters during and after the lockdown.[1] However, using a similar difference-in-difference approach to examine the linkage between COVID-19 and air pollution in Hubei, the province at the center of the outbreak in China, Almond et al. (2020) show that COVID-19 had ambiguous impacts on China's pollution, such as even some relative deterioration in air quality near the pandemic's epicenter. Furthermore, to our knowledge, the emerging literature on COVID-19 focuses on country-specific case studies rather than investigates the impacts of the pandemic on the global scale.[2]

[1] For example, Ma et al. (2020) show a decrease in concentration of nitrogen dioxide (NO_2) by 14 percent in Wuhan, China. Similarly, Tobías et al. (2020) find that NO_2 concentration was reduced by half during the lockdown in Spain, another hot spot of COVID-19.

[2] Other studies examine instead a related outcome, the impacts on health outcomes caused by the pandemic-induced changes in air quality. For example, Cicala et al. (2020) use a sample of more than 3,100 counties in the US and show that reductions in emissions from travel and electricity usage reduced deaths by over 360 deaths per

We fill in this gap in the literature and offer the first assessment of the pandemic impacts on air pollution in a multi-country setting. Given the negative impacts of air pollution on health in the recent studies cited above and its general linkage to heart and lung damage and other diseases (Brunekreef and Holgate, 2002; Liu et al., 2019), understanding how air quality is affected during the COVID-19 pandemic will provide important empirical evidence for health and environmental policies.

Specifically, we make several new contributions in this study. First, we offer global estimates for the causal impacts of COVID-19 on air quality, using a Regression Discontinuity Design (RDD) approach in a short window of time before and after each country implemented its lockdown policies. Since the lockdown—as most society-wide regulations or policies—cannot be randomized across countries, the RDD offers us the most rigorous evaluation model that is available. Second, we provide estimates for several different measures of air quality. While most existing studies restrict analysis to one or two indicators of air quality, we employ two indicators NO_2 and $PM_{2.5}$ for our main analysis and several other indicators for robustness check including O_3, PM_{10}, and SO_2. These various indicators help strengthen the estimation results.

Finally, we combine a variety of real-time data sources for richer analysis. We obtain daily data on air pollution at the sub-national level from two sources: satellite data (from the European Union's Copernicus programme) and station-based data (from the World Air Quality Index). We then combine these air quality data with the Oxford COVID-19 Government Response Tracker (OxCGRT), which provides a unique measure of government responsiveness to COVID-19. We also supplement our analysis with data from several other reliable sources including the National Oceanic and Atmospheric Administration, Google

month. Isphording and Prestel (2020) find air pollution to affect the severity of already realized infections rather than the probability of infection itself in Germany. On the other hand, Cole et al. (2020) find 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 in the Netherlands. Notably, these studies focus on one specific country only.

Community Mobility Reports, World Bank World Development Indicators, WHO Global Ambient Air Quality Database, and Economist Intelligence Unit.

The rich database that we assemble allows us to address a key issue in cross-country analysis, which is to construct lockdown dates for different countries. Indeed, identifying comparable cut-off dates across different countries is challenging. The term 'lockdown' can refer to anything from mandatory quarantines to bans on events and gatherings, closures of certain types of businesses or non-mandatory stay-at-home recommendations. Some governments immediately respond to the outbreak by implementing a complete (regional or national) lockdown (e.g., China, Italy), while some implement a gradual lockdown in a staggering manner for different locations (e.g., the United States). We also present a number of robustness tests regarding our constructed lockdown dates. Once we establish the causal relationship of COVID-19-induced lockdowns on air pollution, we explore the role of movement and travel restrictions as potential mechanisms.

We find strong evidence for reduced air pollution after the lockdowns, with the reduction becoming stronger as the lockdowns go into effect for a longer period. In particular, the global decreases in NO_2 and $PM_{2.5}$ hover around 9 percent and 4 percent, respectively, 90 days after the lockdowns. Our estimation results are qualitatively similar for different indicators of air quality and government policy indexes, and remain robust to alternative model specifications regarding bandwidths, functional forms, and the inclusion of different covariates. We also find some limited evidence that countries with a higher share of trade and manufacturing in the economy have more reduced air pollution after the lockdowns, as do countries with an initially lower level of air pollution. But the opposite result holds for countries near the equator. Our findings suggest that mobility restrictions following the lockdowns can be a channel that explains the improvement of air quality.

The remainder of the paper is organized as follows. We describe the database we constructed for analysis in Section 2 before discussing the empirical models in Section 3. We present the estimation results in Section 4 and conclude in Section 5.

2. Data

To examine the relationship between COVID-19 and air quality, we use two measures of air pollution, namely fine particulate matter $PM_{2.5}$ (mass concentration of particles with diameters ≤2.5 um) and nitrogen dioxide NO_2. While other pollutants are available in our dataset, we select the $PM_{2.5}$ and NO_2 given their direct link to human health. $PM_{2.5}$ is a common cause for adverse health outcomes such as chronic obstructive pulmonary disease (COPD) and lower respiratory infection (LRI) causing death of nearly three million people globally (Gakidou et al., 2017). At the same time, NO_2 is the leading source of childhood asthma in urban areas globally (Achakulwisut et al., 2019). In this study, we collect data on these measures from October 1st, 2019 to June 1st, 2020. We also use other pollutants, such as NO_2, SO_2 and O_3, for robustness checks.

The NO_2 data are derived from images of pollution-monitoring satellites released by the National Aeronautics and Space Administration (NASA) and European Space Agency (ESA). In particular, we use data from the Sentinel-5P/TROPOMI (S5P) instrument of the European Union's Copernicus programme. The Copernicus S5P provides daily global coverage of atmospheric parameters at high resolution (i.e., a pixel size of about 5.5 km x 3.5 km after August 6th, 2019).[3] We then use Google Earth Engine to process and average air quality data at the sub-national level using administrative areas from Database of Global Administrative Areas (GADM). In particular, we measure air pollution at the first-order administrative division

[3] The data have recently been used to study changes in air quality caused by COVID-19 in some health and environmental studies (see, e.g., Chen et al. (2020) and Zambrano-Monserrate et al. (2020)).

(ADM1).[4] While the Copernicus S5P records a wide range of pollutants including NO_2 and others (O_3, SO_2, CO, CH_4, and aerosols), we focus on NO_2 because this is a noxious gas emitted by motor vehicles, power plants, and industrial facilities (see, e.g., Dutheil et al. (2020)). Among other pollutants, NO_2 is also a particularly well-suited data to analysis of emission because it has a short lifetime; this implies that molecules of NO_2 stay fairly close to their sources and thus offer an appropriate measure of changes in emissions.

A potential concern of using satellite air quality, however, is cloud cover. This can bias results by obscuring the sensor's view of the lower atmosphere. Concentrations of NO_2 in the atmosphere are also highly variable in space and time due to factors such as varying traffic flows on weekdays versus weekends and changes in weather conditions. Therefore, we follow suggestions from the Copernicus program and perform a cloud masking which excludes results from pixels with > 10 percent cloud fraction.[5] We also average data over weekly periods as a robustness test. Finally, we include data on daily rainfall and temperature to control for weather conditions, which are derived from the National Center for Environmental Prediction (NCEP) at the National Oceanic and Atmospheric Administration (NOAA). The global dataset provides four 6-hour daily records of temperature and precipitation at the resolution of approximately 25 km. We extract the weather data at the sub-national level using a similar process as with the air pollution data.

As an alternative measure of air quality, we use daily station-based air quality index (AQI) from the World Air Quality Index (WAQI) project. The AQI provides accurate and reliable information on different air pollutant species from more than 12,000 ground-based air quality monitoring stations (primarily located at/near the US embassies and consulates)

[4] In some countries, the ADM1 refers to province level while for others, it refers to state/region level. The administrative data are available at https://gadm.org/about.html.
[5] For more details, see: https://atmosphere.copernicus.eu/flawed-estimates-effects-lockdown-measures-air-quality-derived-satellite-observations?q=flawed-estimates-effects-lockdown-measures-air-quality-satellite-observations

situated in 1,000 major cities in more than 100 countries from 2014 to present. However, there are certain limitations with station-based data. One is that station-based data are often reported more slowly, and not in a 'real-time' fashion as satellite data. Another limitation is the locations of air quality monitoring stations are likely not random, so they may not provide representative data on an area's air quality. Consequently, the satellite data are our preferred data for analysis.

We subsequently match the air pollution data with the government stringency data from the Oxford COVID-19 Government Response Tracker (OxCGRT). The OxCGRT is a novel country-level dataset published by the Blavatnik School of Government at the University of Oxford, which contains information on various lockdown measures, such as school and workplace closings, travel restrictions, bans on public gatherings, and stay-at-home requirements (Hale et al., 2020). It measures government stringency responses on a scale of 0 to 100%. We provide a description of the index components in Table A1 (Appendix 1).

To explore a potential channel through which COVID-19 affects air quality, we collect data on mobility from Google Community Mobility Reports. The Google Community Mobility Reports provide daily data on Google Maps users who have opted-in to the 'location history' in their Google accounts settings across 132 countries. The reports calculate changes in movement compared to a baseline, which is the median value for the corresponding day of the week from January to present. The purpose of travel has been assigned to one of the following categories: retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential. In our analysis, we expect that the lockdowns will lead to reduced mobility of all categories, except for the residential category. We also examine data from several additional sources for robustness checks. The data sources are listed in Table A3 (Appendix 1).

3. Empirical Model

We first employ a panel data model with country fixed effects and time fixed effects to examine whether air quality improves in response to government COVID-19 lockdown policies

$$A_{it} = \beta S_{it} + \gamma X_{it} + \alpha_i + \tau_t + \epsilon_{it} \quad (1)$$

The coefficient of interest in Equation (1) is β, which measures how the air quality (A_{it}) in country *i* and date *t* changes in response to the stringency of government COVID-19 policies (S_{it}). Because A_{it} varies by country and date, this fixed-effects model allows for the inclusion of country fixed effects (α_i) and time fixed effects (τ_t) to absorb the effects of unobservable time-invariant country or time characteristics. X_{it} is a vector of time-varying control variables such as daily temperature and rainfall (or humidity). We estimate Equation (1) using global data at the sub-national level, and we also replicate our analysis at the country level as a robustness test.

Yet, while Equation (1) provides a useful evaluation of the relationship between air quality and the strictness of government COVID-19 policies, it will yield an inconsistent estimate of β if there are omitted variables that simultaneously correlate with both air quality and government policies. For example, countries with strong institutions likely implement stringent policies during the pandemic, and at the same time, may have had better programs in place that ensures better air quality. Furthermore, since the model analyzes a sample of countries after the lockdown date (i.e., with the stringency index being positive), it does not take into account the fact that different countries can differ in terms of pre-COVID-19 characteristics such as governance quality and public preferences for protecting the environment. Another potential threat to Equation (1) is reverse causality. If air pollution is positively associated with the number of COVID-19 cases (Cicala et al., 2020; Cole et al., 2020; Isphording and Prestel, 2020), this can lead to governments implementing more stringent policies on air quality.

In order to identify the causal effects of COVID-19 on air quality, we take advantage of the pandemic-induced lockdowns as an exogenous policy shock and apply a sharp Regression Discontinuity Design (RDD) approach. In this approach, the observations immediately before the lockdowns provide the counterfactual outcomes for those observations immediately after the lockdowns because the lockdown (treatment) status is randomized in a small neighborhood of the lockdowns (Hahn et al., 2001).

More formally, the treatment effect can be estimated as the change in air quality (A) in the neighborhood of lockdown dates

$$\tau_{RD} = \lim_{\varepsilon \downarrow 0} E[\mathrm{A}|d = 0 + \varepsilon] - \lim_{\varepsilon \uparrow 0} E[\mathrm{A}|d = 0 + \varepsilon] \tag{2}$$

where d is the number of days before and after the official dates of lockdowns. We thus estimate the following reduced form

$$A_{it} = \delta L_{it} + f(d_{it}) + \theta X_{it} + \mu_i + \pi_t + \varepsilon_{it} \tag{3}$$

where L_{it} (treatment variable) is a dummy variable that equals 1 after the lockdown and 0 otherwise, and δ is the parameter of interest. $f(d_{it})$ denotes a function of the running variable d_{it} (number of days from the lockdown dates). Similar to Equation (1), μ_i and π_t respectively denote the country fixed effects and the time fixed effects, and ε_{it} denotes the error term. For comparison and robustness checks, we employ a parametric approach and use different functional forms of the running variable d_{it} to estimate Equation (3). These include (i) the linear model, (ii) the linear model with the interaction term of the running variable and the treatment variable (L_{it}*d_{it}), (iii) the quadratic model, and (iv) the quadratic model with the interaction term of the running variable and the treatment variable (L_{it}*d_{it}^2). We present results for a broad range of bandwidths including 30, 60, and 90 days before and after the official lockdown dates. As shown by Figure 1 in the next section, the impacts of lockdowns become stronger over time, so these different bandwidths help capture the impacts of lockdowns over different time windows. We cluster the standard errors at the sub-national level in all models. We also offer a number of robustness checks in Section 4.2.

A key challenge with estimating Equation (3) is to identify lockdown dates that are comparable across different countries. As discussed earlier, different countries likely implement lockdowns with different degrees of strictness (i.e., business activities and travels can continue to varying extents after the lockdown dates for different countries). Indeed, in quite a few countries, while all schools are shut down, universities operate on a different schedule. Furthermore, there can be multiple lockdown dates even within the same country where different regions/ states impose different lockdown dates (with different levels of intensity). To address this issue, the OxCGRT data provides a unique composite measure which combines indicators on different aspects of lockdown policies regarding school, workplace and public transportation and public events into a general index (Table A1, Appendix 1). By using a range of different indicators, this stringency index accounts for any indicator that may be over- or mis-interpreted, thus allows for a better and more systematic comparison across countries (Hale et al., 2020).

For each country, we define the official lockdown date as the first day on which the stringency index becomes positive. Using our constructed measure, Figure A1 (Appendix 1) shows that most countries introduced lockdown policies somewhere between the last week of January and the first week of February 2020. Notably, the start dates of lockdowns do not correspond to the intensity of the stringency index as countries that implemented lockdown policies later tend to be more stringent in their response.

Although the OxCGRT data provides a systematic comparison across different countries, it is still possible that not all business activities and travel cease exactly by the time of our constructed lockdown dates. In that case, a better approach is to employ the fuzzy RDD model rather than the sharp RDD model where the treatment variable L_{it} can assume the value of 0 for $S_{it} > 0$ for some countries. However, we do not have such additional information for L_{it} in our case and have to uniformly define L_{it} as 1 after the lockdown date for each country.

But we offer a multi-layered approach to ensure that estimation results are robust. First, the estimates using Equation (1) above provide the first set of evidence over the (correlational) relationship between air quality and the government stringency index. Second, examining the outcomes over three different time bandwidths helps average out any lingering impacts after the lockdowns and provides comparisons for estimation results. If estimates remain (qualitatively) similar, this indicates that they are robust to this concern. Third, we also consider different versions of the bandwidths (such as using weekly air quality data instead of daily air quality data) and the stringency index (such as probing more deeply into its different components or using different thresholds). Finally, we also offer a battery of other additional robustness tests in Section 4.2.

An advantage of the RDD design is that the identification assumptions offer testable predictions. To validate our design, we present two types of tests. First, we investigate the distribution of observations (ADM1 level for satellite data and city level for station-based data) around the cut-off date. Table A2 (Appendix 1) provides results of the manipulation test suggested by McCrary (2008) and Cattaneo et al. (2018) based on the nonparametric local polynomial density estimator. The t-tests and their corresponding p-values confirm there is no evidence of systematic manipulation of the running variable. Second, we test for discontinuity in the other covariates around the dates of lockdowns. The results, shown in Figure A2 (Appendix 1), rule out this concern. Table A3 (Appendix 1) provides the summary statistics of the main variables used in this study.

4. Results

4.1. Main findings

We present in Table 1 the estimation results for Equation (1) using two data samples at the sub-national level (columns 1 and 2) and at the country level (columns 3 and 4). Our preferred estimates are shown in columns (2) and (4), which control for daily temperature and precipitation (humidity for station-based data)[6]. But we also show the estimates without these control variables in columns (1) and (3) for comparison and robustness checks. The estimation results are strongly statistically significant in our preferred models (columns 2 and 4) and point to reduced air pollution where government policies are more stringent. Overall, our findings suggest that global air quality improved in response to COVID-19-induced lockdown policies.

In particular, column (2) indicates that a one-point increase in the stringency index is associated with a 0.040 (mol/km^2) decrease in NO_2 (Panel A). When using station-based data, the corresponding figure is a 0.130 (μg/m^3) decrease in $PM_{2.5}$ (Panel B). Estimates are rather similar when we analyze the data at the country level (column 4). However, as discussed earlier, these estimates are likely biased since employing Equation (1) does not allow us to properly account for the unobservables that may be correlated with both the stringency index and air quality.

We subsequently present our main analysis which examines the lockdown impacts on air quality using the RDD model. For illustrative purpose, Figure 1 provides prima facie evidence of the impact of lockdowns on air quality. The figure shows the results from a data-driven RDD regression of air pollution, measured by NO_2 (Panel A) and $PM_{2.5}$ (Panel B), on the days before and after the lockdown dates.[7] We observe a negative jump at the threshold of cut-off date, which suggests a reduction of air pollution after the lockdowns. The downward sloping trend for air pollution in Figure 1 also suggests that the reduction in air pollution becomes stronger

[6] We do not use precipitation from the station-based data due to its low frequency (a large number of stations do not record rainfall level).
[7] We use the optimal bandwidth proposed by Imbens and Kalyanaraman (2012).

as the lockdowns go into effect for a longer period. This is understandable, since a short period of time may not be sufficient to detect the changes in air quality.

We report the estimation results for Equation (3) in Table 2, which shows estimates using two data samples: the satellite data (panel A) and the station-based data (panel B). We consider three different bandwidths, 30 days, 60 days, and 90 days before and after the lockdowns. As suggested by Figure 1, a wider time window from the lockdown date can capture a stronger impact of the lockdowns on air quality. Our preferred models are, again, those that control for weather conditions (columns 2, 4, and 6). In each panel of Table 2, we estimate four models using four different functional forms of the running variable as discussed earlier.

The estimation results using the satellite data, our main data for analysis, shows that air quality improves after the lockdowns, and the results are strongly statistically significant at the 5% level or less (panel A). The estimates are rather qualitative similar whether we include control variables or not. Specifically, the estimated coefficient on the lockdown variable is negative and statistically significant at the 1 percent level using the linear model (panel A, column 2). This indicates that the lockdown leads to a 2.097 (mol/km^2) decrease in the global concentration of NO_2 after 90 days. This translates into a 9.1 percent decrease compared to an average value of NO_2 of 22.914 mol/km^2 before the lockdowns. We also find that using different functional forms (models 2 to 4) results in similar estimates. Finally, the negative impacts of lockdowns on NO_2 are rather consistent across bandwidths, but have a smaller magnitude with narrower bandwidths (as also seen with Figure 1). The decreases in concentration of NO_2 are roughly 5 percent for 60 days (panel A, column 4) and 2 percent for 30 days (panel A, column 6) after the lockdowns, respectively.

We turn next to the alternative station-based data and find a strong impact of the lockdowns on $PM_{2.5}$ using the bandwidth of 90 days (panel B, column 2). The global decrease in $PM_{2.5}$ for 90 days after the lockdowns hovers around 3 to 4 percent depending on the functional form that we employ. But estimates become statistically insignificant for the shorter bandwidths of 60 days and 30 days. We then use different measures of air pollution available from the station-based data and reach a similar conclusion. Specifically, the results presented in Table A4 (Appendix 1) confirm the beneficial effects of lockdowns on air quality, as measured by NO_2 and PM_{10}, 90 days before and after the lockdowns. While there is no evidence of the lockdown effect on SO_2, the indicator O_3 is found to be positively associated with the lockdowns at the windows of 90 days and 60 days. A possible explanation for the increase in concentration of O_3 is warmer weather during this period (Tobías et al., 2020).

4.2. Robustness tests

In this section, we conduct a battery of robustness tests on the estimation results. These include employing a nonparametric RDD method, adding different covariates to the regressions, using wider time bandwidths and different thresholds and versions of the stringency index, controlling for potentially differential time trends across countries, and converting the air quality variables into logarithmic form.

First, since employing specific functional forms can affect the parametric RDD estimation results, we adopt a nonparametric RDD method for robustness checks. An important feature of the nonparametric method is that the bandwidth is not selected arbitrarily; instead, it is calculated on a data-driven basis. In Table A5 (Appendix 1), we report the results of non-parametric specifications using two optimal bandwidths: the one common mean squared error (MSE) bandwidth and the two selectors MSE-optimal bandwidth (MSE-2). We find consistent impacts of the lockdowns on NO_2 using the satellite data, while there is little evidence of the impacts on $PM_{2.5}$ using the alternative station-based data.

Second, our estimation results are rather similar whether we control for weather conditions in our RDD regressions. For further checks, we include additional covariates to control for the pre-pandemic country characteristics, namely country's log of GDP per capita (in constant 2010 USD), population density, log of energy consumption per capita, the number of motor vehicles per 1,000 inhabitants, and the share of electricity generated by coal power. These country characteristics come from the World Development Indicators (WDI) database in the latest year when data is available. The estimation results in Table A6 (Appendix 1) are consistent with our main findings.

Third, a potential issue with daily air pollution data is that these data can substantially vary from one day to another because of variations in emission and changes in weather conditions. Therefore, we replicate our parametric RDD approach using a weekly indicator. We employ different bandwidths of 5, 10 and 15 weeks before and after the lockdown date. The results are presented in Table A7 (Appendix 1), which are generally consistent with the main findings in Table 2.

Fourth, the lockdown dates are identified based on the stringency index becoming positive. As a robustness check, we also plot the estimated impacts of lockdowns from Model 4 in Table 2 for other thresholds of the stringency index that range from 0 to 50% (on a scale of 0-100%). The estimation results, shown in Figure A3 (Appendix 1), indicate that using thresholds up to roughly 10% yields very similar results. The lockdown impacts, however, become much weaker for thresholds of larger values but remain strongly statistically significant. Fifth, we use alternative measures of stringency index taken from the OxCGRT dataset. There are two versions of the stringency index: (i) a "regular" version which returns null values if there are insufficient data to calculate the index, and (ii) a "display" version which extrapolates to smooth over the last seven days of the index based on the most recent complete data. We use the latter indicator for our main analysis, but we also find consistent results using the "regular" version (Appendix 1, Table A8).

Sixth, the stringency index in the OxCGRT dataset is calculated using a simple additive unweighted approach. It is thus possible that some dimensions with higher weights will be underestimated in the index. To address this issue, we create a new index based on the Principal Component Analysis (PCA) method for all the dimensions of stringency index. Table A9 in Appendix 1 shows similar estimation results for our own index.[8] Seventh, we further explore other indexes that are available from the OxCGRT dataset. They include: (i) Government response index, (ii) Containment and health index, and (iii) Economic support index.[9] Compared to our main measure, the government response index and the containment and health index include two additional dimensions: testing policy and contact tracing. Still, we find a consistent impact of the lockdowns on air pollution when using these indexes, except for the economic support index when using small bandwidths of 60 days and 30 days (Appendix 1, Table A10). However, the economic support index only includes income support programs and debt relief programs, so it does not fully capture the overall responsiveness of the government.

Eighth, we also check whether our results are driven by differential time trends across countries. We include in the regressions the interaction terms of country dummies with linear time trends. The results, presented in Table A11 (Appendix 1), are generally consistent with our main findings. They become somewhat weaker for NO2 but stronger for PM2.5 for the 30 day bandwidth. Finally, our findings also remain consistent when we use the logarithmic form of the air quality variable (Appendix 1, Table A12).

4.3. Heterogeneity analysis

[8] A notable change is that the lockdowns have positive impacts on NO_2 concentration for the 30 day bandwidth, but this is only marginally statistically significant. Another change is that the lockdowns now have negative impacts on $PM_{2.5}$, and these impacts are strongly statistically significant.

[9] Another index is Legacy stringency index; however, it is not recommended by the OxCGRT team (Hale et al., 2020).

Having shown that changes in air quality are driven by COVID-19, it is useful to understand whether the impacts of lockdowns differ by certain country characteristics. In particular, the impacts of lockdowns can vary according to a country's geographic location. For example, cities near the deserts are often affected by sand and dust storms, which can strongly impact air quality. We thus interact a dummy variable indicating whether a country is near the equator with the treatment variable. The results presented in panel A of Table 3 show that countries near the equator have a higher concentration of NO_2 after the lockdowns.

A country's institution may also affect the impacts of lockdowns. A large body of economic literature has shown the important role of institutions and culture in shaping economic development (e.g. Gorodnichenko and Roland, 2017; Acemoglu et al., 2019). Consequently, we use the democracy index from the 2019 report of the Economist Intelligence Unit. We expect that countries with strong institutions likely implement stringent policies during the time of COVID-19, and therefore have a better performance in terms of air quality. The results in panel B of Table 3, however, provide little support for this argument. In contrast, partial democratic countries and countries with hybrid regime appear to have less reduced air pollution after the lockdowns than authoritarian countries.

Another useful heterogeneity analysis is whether countries with a high level of openness have more reduced air pollution after the lockdowns. Whether trade is good or bad for the environment has been a topic of debate in the literature. While evidence exists on the beneficial effects of trade on the environment (e.g. Antweiler et al., 2001; Frankel and Rose, 2005), other studies show that trade openness could in fact lead to higher emissions (Managi et al., 2009, Li et al., 2015). To answer this question, we interact a country's share of manufacturing and share of trade in its GDP (from the 2019 World Development Indicators (WDI) database) with the treatment variable. The estimation results, presented in panels C and D of Table 3, show that countries with a larger share of trade or manufacturing have more reduced air pollution after the lockdowns.

Finally, we examine whether countries with existing lower levels of air pollution may reduce air pollution more. We use the WHO Global Ambient Air Quality Database that summarizes concentration of $PM_{2.5}$ at the country level in 2018. We then split our sample into five quintiles and interact each with our treatment variable. The results in panel E indicate that countries with an initially lower level of air pollution (i.e., the 1st quintile) have more reduced air pollution compared to those with initially higher levels of air pollution. For further illustration, we interact our treatment variable with the country dummies and plot the estimated interaction terms against countries' initial level of air quality in Figure A4 (Appendix 1). We highlight a country's population size by drawing a bubble graph, where the size of a country's circle is proportionate to its population size. Figure A4 shows countries bunching to the left of the graph and below the zero (no change) line, confirming that countries with better air quality before the pandemic tend to have higher reduction of NO_2. Several countries with a large population size that improved air quality stand out, including China, Iraq, Norway, Russia, South Korea, and the United States.

4.4. Stringency index and mobility restriction

Once we established the causal impacts of COVID-19 on air pollution, we shift our attention to the role of mobility restrictions as a potential mechanism. Due to COVID-19, human mobility and relevant production and consumption activities have since decreased significantly. Given that one main source of air pollution comes from traffic mobility (Viard and Fu, 2015), it is reasonable to argue that more stringent policies will result in less mobility, thereby improving air quality.

We directly test this hypothesis by using data from the Google Community Mobility Reports. Since mobility data were not available before the lockdown date, we are unable to apply the more rigorous RDD approach. Consequently, we estimate the panel data model with the country and time fixed effects in Equation (1). The estimation results obtained by the panel model are in fact qualitatively similar to those obtained by the RDD approach. As such, applying the panel data can provide some qualitative evidence on the mechanism of impacts.

We present the estimation results in Table 4, which show that geographic mobility has declined significantly where government policies are more stringent. In particular, a higher stringency index is associated with less mobility in both 'essential services' (e.g., grocery and pharma, workplace) and 'non-essential services' (retail and recreation, parks), but more mobility in the 'residential' category.

5. Conclusions

We contribute to the emerging literature on COVID-19 by offering the first study that provides cross-national evidence on the causal impacts of COVID-19 on air pollution. We assemble a rich database from a number of different reliable sources, which we analyze with panel data and RDD econometric models.

Our findings provide a better understanding of the unexpected positive impacts of the pandemic on air quality. We find heterogeneous impacts for different country characteristics such as shares of trade and manufacturing in the economy, initial levels of air pollution or proximity to the equator. We identify reduced mobility as a potential channel that can help reduce air pollution. Our findings suggest that while mobility restrictions appear not to be a long-term solution to address air pollution, reducing nonessential individual movements can help improve air quality on a global scale. A promising direction for future research can be more in-depth country studies on the impacts of the pandemic on air quality.

References

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy, 127*(1), 47-100.

Achakulwisut, P., Brauer, M., Hystad, P., & Anenberg, S. C. (2019). Global, national, and urban burdens of paediatric asthma incidence attributable to ambient NO2 pollution: estimates from global datasets. *The Lancet Planetary Health, 3*(4), e166-e178.

Akesson, J., Ashworth-Hayes, S., Hahn, R., Metcalfe, R. D., & Rasooly, I. (2020). Fatalism, Beliefs, and Behaviors During the COVID-19 Pandemic (No. w27245). National Bureau of Economic Research.

Almond, D., Du, X., & Zhang, S. (2020). Did COVID-19 Improve Air Quality Near Hubei? (No. w27086). National Bureau of Economic Research.

Antweiler, W., Copeland, B. R., & Taylor, M. S. (2001). Is free trade good for the environment? *American Economic Review, 91*(4), 877-908.

Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., & Yannelis, C. (2020). How does household spending respond to an epidemic? Consumption during the 2020 covid-19 pandemic (No. w26949). National Bureau of Economic Research.

Brodeur, A., Cook, N., & Wright, T. (2020). On the Effects of COVID-19 Safer-At-Home Policies on Social Distancing, Car Crashes and Pollution. (No. 13255). IZA Discussion Paper.

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *The Lancet*, 360(9341), 1233-1242.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics, 101*(3), 442-451.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal, 18*(1), 234-261.

Chen, K., Wang, M., Huang, C., Kinney, P. L., & Anastas, P. T. (2020). Air pollution reduction and mortality benefit during the COVID-19 outbreak in China. *The Lancet Planetary Health*, *4*(6), e210-e212.

Cicala, S., Holland, S. P., Mansur, E. T., Muller, N. Z., & Yates, A. J. (2020). Expected Health Effects of Reduced Air Pollution from COVID-19 Social Distancing (No. w27135). National Bureau of Economic Research.

Cole, M., Ozgen, C., & Strobl, E. (2020). Air Pollution Exposure and COVID-19 (No. 13367). IZA Discussion Paper.

Dang, Hai-Anh, Toan Huynh, and Manh-Hung Nguyen. (2020). Does the COVID-19 Pandemic Disproportionately Affect the Poor? Evidence from a Six-Country Survey (No. 13352). IZA Discussion Paper.

Dutheil, F., Baker, J. S., & Navel, V. (2020). COVID-19 as a factor influencing air pollution? *Environmental Pollution, 263*, 114466.

Fairlie, R. W., Couch, K., & Xu, H. (2020). The impacts of COVID-19 on minority unemployment: First evidence from April 2020 CPS microdata (No. w27246). National Bureau of Economic Research.

Frankel, J. A., & Rose, A. K. (2005). Is trade good or bad for the environment? Sorting out the causality. *Review of Economics and Statistics, 87*(1), 85-91.

Gakidou, E., Afshin, A., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., ... & Abu-Raddad, L. J. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1345-1422.

Gorodnichenko, Y., & Roland, G. (2017). Culture, institutions, and the wealth of nations. *Review of Economics and Statistics*, *99*(3), 402-416.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201-209.

Hale, T., Petherick, A., Phillips, T., & Webster, S. (2020). Variation in government responses to COVID-19. Blavatnik school of government working paper, 31.

He, G., Pan, Y. & Tanaka, T. (2020). The short-term impacts of COVID-19 lockdown on urban air pollution in China. *Nature Sustainability*. https://doi.org/10.1038/s41893-020-0581-y

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, *79*(3), 933-959.

Isphording, I. E., & Pestel, N. (2020). Pandemic Meets Pollution: Poor Air Quality Increases Deaths by COVID-19. (No. 13418). IZA Discussion Paper.

Li, Z., Xu, N., & Yuan, J. (2015). New evidence on trade-environment linkage via air visibility. *Economics Letters*, *128*, 72-74.

Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A. M., Guo, Y., Tong, S., ... & Valdes Ortega, N. (2019). Ambient particulate air pollution and daily mortality in 652 cities. *New England Journal of Medicine*, 381(8), 705-715.

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J. & Luo, B. (2020). Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Science of the Total Environment*, 138226.

Managi, S., Hibiki, A., & Tsurumi, T. (2009). Does trade openness improve environmental quality? *Journal of Environmental Economics and Management*, *58*(3), 346-363.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698-714.

Tobías, A., Carnerero, C., Reche, C., Massagué, J., Via, M., Minguillón, M. C., Alastuey, A. & Querol, X. (2020). Changes in air quality during the lockdown in Barcelona (Spain) one month into the SARS-CoV-2 epidemic. *Science of the Total Environment*, 138540.

Viard, V. B., & Fu, S. (2015). The effect of Beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics*, *125*, 98-115.

Zambrano-Monserrate, M. A., Ruano, M. A., & Sanchez-Alcalde, L. (2020). Indirect effects of COVID-19 on the environment. *Science of the Total Environment*, 138813.

Figure 1: COVID-19 lockdowns and air pollution

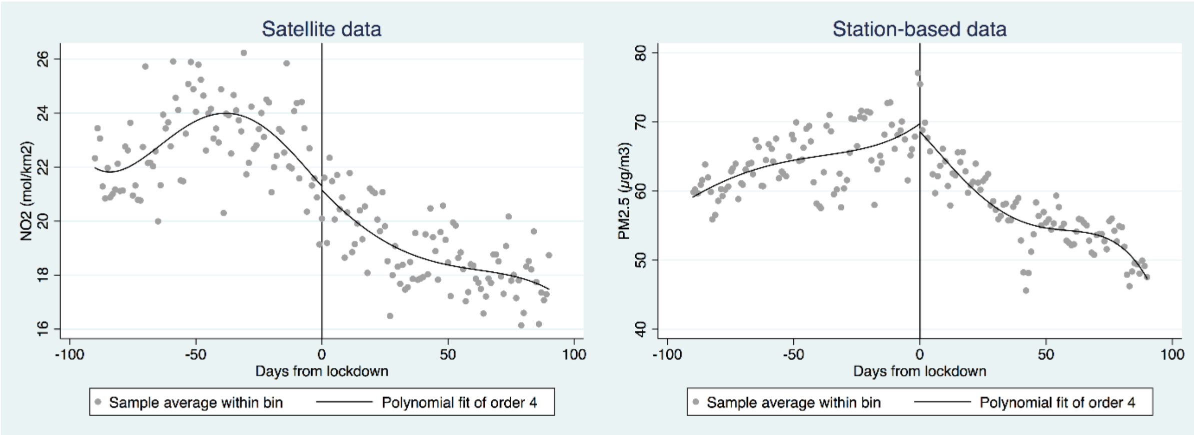


Table 1: Government response to COVID-19 and air pollution

	ADM1/City level		Country level	
	(1)	(2)	(3)	(4)
Panel A: Air quality is measured by NO_2 (satellite data)				
Stringency index	-0.032***	-0.040***	-0.040***	-0.033***
	(0.003)	(0.003)	(0.012)	(0.010)
Controls	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	250,838	248,120	14,850	14,712
R-squared	0.381	0.381	0.657	0.658
Panel B: Air quality is measured by $PM_{2.5}$ (station-based data)				
Stringency index	-0.165***	-0.130***	-0.176***	-0.150***
	(0.019)	(0.017)	(0.050)	(0.041)
Controls	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	81,478	75,048	12,784	11,986
R-squared	0.449	0.456	0.591	0.614

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of panel model. Robust standard errors in parentheses. Standard errors are clustered at ADM1/city level in columns (1) and (2), and country level in columns (3) and (4). Regressions in columns (2) and (4) include country dummies and week dummies. Control variables are daily temperature and rainfall (humidity for station-based data).

Table 2: COVID-19 lockdowns and air pollution

Panel A: Satellite air pollution

Air quality:	+/-90 days		+/-60 days		+/-30 days	
NO_2	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: Linear model						
Lockdown=1	-2.037***	-2.097***	-1.209***	-1.219***	-0.568**	-0.535**
	(0.231)	(0.229)	(0.220)	(0.221)	(0.223)	(0.225)
Model 2: Linear interaction model						
Lockdown=1	-1.973***	-2.057***	-1.182***	-1.185***	-0.540**	-0.505**
	(0.229)	(0.229)	(0.219)	(0.220)	(0.221)	(0.223)
Model 3: Quadratic model						
Lockdown=1	-1.987***	-2.078***	-1.197***	-1.206***	-0.551**	-0.517**
	(0.229)	(0.229)	(0.220)	(0.221)	(0.222)	(0.224)
Model 4: Quadratic interaction model						
Lockdown=1	-1.986***	-2.075***	-1.184***	-1.192***	-0.613***	-0.583***
	(0.229)	(0.229)	(0.220)	(0.221)	(0.224)	(0.226)
Means before lockdowns	22.914	22.914	23.316	23.316	22.719	22.719
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	381,872	378,101	255,628	252,991	128,041	126,799

Panel B: Station-based air pollution

Air quality:	+/-90 days		+/-60 days		+/-30 days	
$PM_{2.5}$	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: Linear model						
Lockdown=1	-4.539***	-2.665***	-1.952*	-0.288	0.903	-0.669
	(0.925)	(0.912)	(1.007)	(0.914)	(1.008)	(1.047)
Model 2: Linear interaction model						
Lockdown=1	-3.905***	-2.182**	-1.386	0.127	1.195	-0.493
	(0.917)	(0.900)	(0.996)	(0.905)	(1.002)	(1.049)
Model 3: Quadratic model						
Lockdown=1	-4.057***	-2.271**	-1.520	0.063	1.104	-0.570
	(0.919)	(0.901)	(0.999)	(0.905)	(1.001)	(1.049)
Model 4: Quadratic interaction model						
Lockdown=1	-3.886***	-2.159**	-1.331	0.081	1.295	-0.367
	(0.918)	(0.900)	(0.996)	(0.903)	(1.007)	(1.048)
Means before lockdowns	64.599	64.599	66.015	66.015	67.544	67.544
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	93,941	82,193	63,779	52,502	33,151	24,910

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Results of parametric RDD. Robust standard errors in parentheses. Standard errors are clustered at city level. Air pollution in panel A and panel B is respectively measured by NO_2 from satellite data and $PM_{2.5}$ from station-based data. Model 1 uses running variable in linear form, Model 2 includes interaction of running variable and treatment variable, Model 3 includes quadratic term of running variable, Model 4 includes interactions of running variable (linear and quadratic terms) with treatment variable. All regressions include country dummies and week dummies. Control variables are daily temperature and humidity.

Table 3: Heterogeneity tests

Air quality: NO_2	+/-90 days	+/-60 days	+/-30 days
	(1)	(2)	(3)
Panel A: Location			
Lockdown*Countries near equator	4.069***	3.531***	1.945***
	(0.279)	(0.274)	(0.246)
Observations	378,101	252,991	126,799
Panel B: Democracy			
Reference: Authoritarian			
Lockdown*Hybrid regime	1.097*	1.399**	0.852
	(0.656)	(0.692)	(0.574)
Lockdown*Partial democracy	1.055*	1.141*	-0.210
	(0.605)	(0.662)	(0.555)
Lockdown*Full democracy	-0.504	0.458	-0.677
	(0.797)	(0.764)	(0.663)
Observations	341,845	229,054	114,787
Panel C: Share of trade			
Lockdown*Trade	-0.038***	-0.034***	-0.041***
	(0.008)	(0.009)	(0.013)
Observations	293,697	196,456	98,313
Panel D: Share of manufacturing			
Lockdown*Manufacturing	-0.441***	-0.445***	-0.363***
	(0.048)	(0.057)	(0.058)
Observations	252,656	169,888	85,258
Panel E: Air pollution index			
Reference: 1^{st} quintile			
Lockdown*2^{nd} quintile	0.995**	0.837**	0.644**
	(0.406)	(0.384)	(0.326)
Lockdown*3^{rd} quintile	1.600***	1.463***	1.675***
	(0.344)	(0.344)	(0.336)
Lockdown*4^{th} quintile	-1.433**	-0.866	0.243
	(0.649)	(0.610)	(0.518)
Lockdown*5^{th} quintile	-0.038	-0.686	-0.208
	(0.586)	(0.661)	(0.583)
Observations	373,444	249,859	125,275
Means before lockdowns	22.914	23.316	22.719
Controls	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Results of parametric RDD that include interactions of running variable (linear and quadratic terms) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1 level. Air pollution is measured by NO_2 from satellite data. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall.

Table 4: Stringency index and mobility restriction

Mobility changes	Retail and recreation	Grocery and pharmacy	Park	Transit	Workplaces	Residential
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sub-national level						
Stringency index	-0.819***	-0.384***	-0.586***	-0.766***	-0.619***	0.291***
	(0.009)	(0.013)	(0.013)	(0.012)	(0.007)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	364,599	351,147	211,813	245,560	458,450	254,625
R-squared	0.726	0.461	0.424	0.605	0.656	0.724
Panel B: Country level						
Stringency index	-0.764***	-0.478***	-0.536***	-0.787***	-0.594***	0.284***
	(0.021)	(0.019)	(0.028)	(0.018)	(0.019)	(0.008)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,284	13,284	13,284	13,284	13,284	13,238
R-squared	0.802	0.608	0.670	0.846	0.710	0.788

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of panel model. Robust standard errors in parentheses. Standard errors are clustered at sub-region level in Panel A and country level in Panel B. Results of panel analysis. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall.

Appendix 1: Additional Figures and Tables

Figure A1: Number of countries that introduced lockdowns and policy stringency index, OxCGRT database

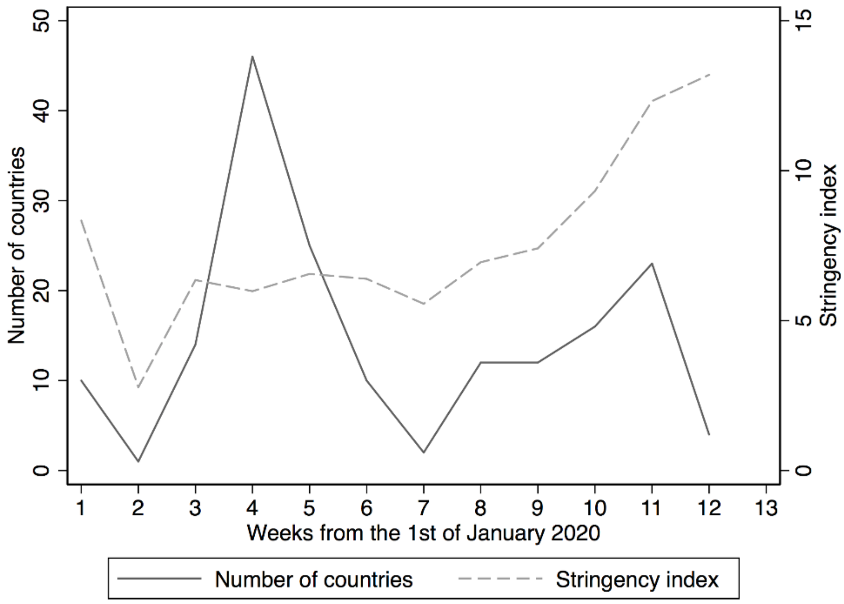


Figure A2: COVID-19 lockdowns and temperature/precipitation

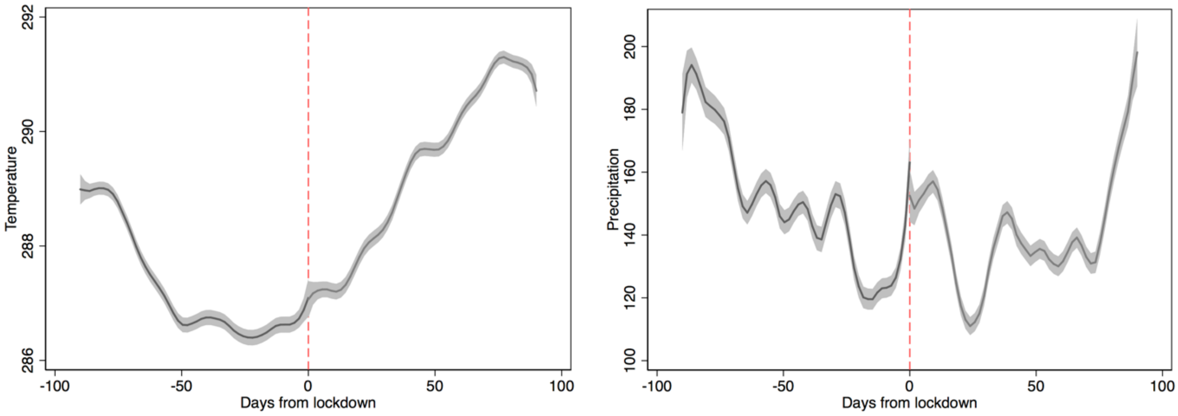
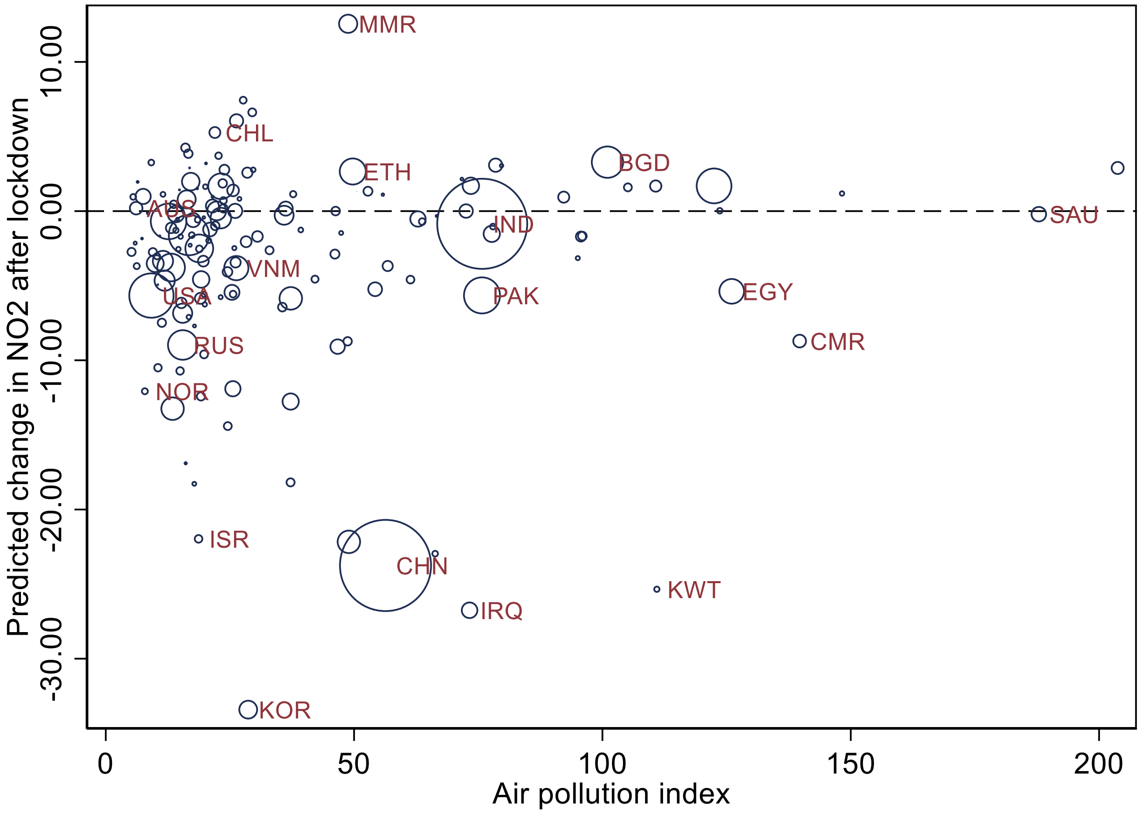


Figure A3: Reduction of air pollution using alternative cut-offs of stringency index

Notes: Air pollution is measured by NO_2 from satellite data. Each point in the figure shows point estimate and 95 percent confidence interval of treatment variable (lockdown) using different percentiles of stringency index to construct lockdown date. The parametric RDD model includes interactions of running variable (linear and quadratic terms) with treatment variable. The running variable is number of days from the lockdown date. We use bandwidth of 90 days before and after the lockdowns. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall.

Figure A4: Air pollution reduction by country



Notes: Air index is measured by concentration of $PM_{2.5}$ at country level from the 2018 WHO Global Ambient Air Quality Database, where a higher index indicates a higher level of pollution. The figure shows the point estimates of the interaction of the treatment variable (lockdown) and the country dummies using parametric RDD model. The running variable is the number of days from the lockdown date. We use bandwidth of 90 days before and after the lockdowns. Countries are depicted with their population sizes taken from 2019 WDI database. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall.

Table A1: Stringency index components

Number	Components	Description
1	School closing	Record closings of schools and universities
2	Workplace closing	Record closings of workplaces
3	Cancel public events	Record cancelling public events
4	Restrictions on gatherings	Record the cut-off size for bans on private gatherings
5	Close public transport	Record closing of public transport
6	Stay at home requirements	Record orders to "shelter-in- place" and otherwise confine to home.
7	Restrictions on internal movement	Record restrictions on internal movement
8	International travel controls	Record restrictions on international travel
9	Public info campaigns	Record presence of public info campaigns

Notes: Each component is measured by an ordinal scale. The stringency index is measured by the OxCGRT team as simple averages of the individual component indicators. Each component is measured by an ordinal scale (e.g. 0 – no measures, 1 – recommended closing, 2 – require partial closing, 3 – require closing all levels). It is then rescaled by maximum value to create a score between 0 and 100. These scores are then averaged to get the stringency index.

Table A2: Manipulation test

	Bandwidth		Observations		t-test	
	left	right	left	right	t-test	p-value
NO2 (satellite)	7.367	6.694	16,378	16,396	1.170	0.242
PM2.5 (station-based)	6.741	6.964	4,101	4,308	-0.225	0.822

Notes: Results of manipulation test developed by McCrary (2008). We implement a regression discontinuity (RD) manipulation test using local polynomial density estimation following Cattaneo et al. (2018). T-tests and their corresponding p-values indicate there is no statistical evidence of systematic manipulation of the running variable.

Table A3: Data sources and summary statistics

Variable	Descriptions	Mean	Standard deviation	Min	Max
Oxford COVID-19 Government Response Tracker (OxCGRT)					
Source: Blavatnik School of Government at the University of Oxford (https://covidtracker.bsg.ox.ac.uk/)					
Stringency index	Government responses to COVID-19 (Score between 0 and 100)	44.751	35.254	0	100
Government response index		41.404	31.499	0	96.15
Containment and health index		44.148	33.202	0	100
Economic support index		26.331	32.501	0	100
Satellite air quality (daily)					
Source: European Union's Copernicus programme (https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-5p)					
NO_2	Nitrogen dioxide	20.458	26.334	-43.400	886
Satellite weather data (daily)					
Source: National Oceanic and Atmospheric Administration (NOAA) (https://www.ncep.noaa.gov)					
Rainfall	Average rainfall (m)	0.0002	0.0003	0.000	0.015
Temperature	Average temperature (K)	289.715	10.399	232.625	313.183
Station-based data (daily)					
Source: World Air Quality Index (WAQI) project (https://waqi.info/)					
$PM_{2.5}$	Particles with a diameter of 2.5 micrometres or less.	56.291	43.799	1	999
PM_{10}	Particles with a diameter of 10 micrometres or less.	27.338	25.403	1	999
NO_2	Nitrogen dioxide	10.118	8.442	0	500
SO_2	Sulfur dioxide	4.126	7.895	0	500
O3	Ozone	19.459	12.670	0	500
Humidity	Average humidity (%)	69.084	19.276	0	122
Temperature	Average temperature (°C)	14.393	9.200	-67.7	93.3
Mobility rates					
Source: Google Community Mobility Reports (https://www.google.com/covid19/mobility/)					
Retail & Recreation	Changes in people's mobility (%) in different categories	-22.801	28.661	-100	313

Grocery & pharmacy		-6.118	21.645	-100	345
Park		-2.925	51.956	-100	616
Transit		-27.151	30.046	-100	497
Workplaces		-23.812	21.033	-94	258
Residential		10.669	9.177	-25	56
Other control variables (Table A6)					
Source: World Bank World Development Indicators (https://databank.worldbank.org/source/world-development-indicators)					
Energy consumption	Energy consumption per capita (kWh)	24,620	25,452	706.246	215,883
Vehicles	Number of motor vehicles per 1,000 inhabitants	200.713	217.914	1.000	797
GDP	GDP per capita (in constant 2010 USD)	13,260	17,763	208.075	111062
Population density	People per sq. km of land area	164.668	586.711	0.137	20480
CO2 emissions	CO2 emissions (kg per 2010 US$ of GDP)	0.516	0.374	0.056	2.004
Electricity	Electricity production from coal sources (% of total)	19.917	24.166	0.000	96.360
Other control variables (Table 3)					
Democracy index	2019 Economist Intelligence Unit Report *(https://www.eiu.com/topic/democracy-index)*	54.714	20.579	13.200	98.700
Air index	2018 WHO Global Ambient Air Quality Database *(https://www.who.int/airpollution/data)*	36.234	31.953	4.071	203.744

Table A4: COVID-19 lockdowns and air pollution – Other parameters of pollution

	(1)	(2)	(3)
Bandwidths	+/-90 days	+/-60 days	+/-30 days
Panel A: Air quality is measured by PM_{10}			
Lockdown=1	-1.529***	-0.423	-0.694
	(0.552)	(0.546)	(0.582)
Means before lockdowns	30.676	30.904	31.242
Observations	80,024	51,207	24,208
Panel B: Air quality is measured by NO_2			
Lockdown=1	-1.523***	-0.709***	0.180
	(0.233)	(0.214)	(0.200)
Means before lockdowns	12.747	12.899	12.815
Observations	79,912	51,094	24,090
Panel C: Air quality is measured by O_3			
Lockdown=1	2.214***	1.132***	-0.608**
	(0.308)	(0.214)	(0.251)
Means before lockdowns	14.982	14.493	14.464
Observations	74,209	47,295	22,372
Panel D: Air quality is measured by SO_2			
Lockdown=1	-0.457	-0.325	0.466
	(0.331)	(0.417)	(0.649)
Means before lockdowns	4.535	4.697	4.866
Observations	67,689	43,341	20,628
Controls	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic terms) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at city level. Air pollution parameters are derived from station-based data. All regressions include country dummies and week dummies. Control variables are daily temperature and humidity.

Table A5: COVID-19 lockdowns and air pollution – Optimal bandwidth

	Satellite NO_2		Station-based $PM_{2.5}$	
Optimal bandwidth:	MSE-1	MSE-2	MSE-1	MSE-2
Lockdown=1 (Conventional)	-0.718***	-1.046***	0.264	-0.029
	(0.210)	(0.214)	(0.912)	(0.945)
Means before lockdowns	22.914	22.914	64.599	64.599
Controls	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	184,782	216,911	36,162	36,875

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of RDD use the optimal bandwidths based on Calonico et al. (2019). Standard errors are in parentheses and are obtained by clustering at ADM1 level (satellite data) and city level (station-based data). Mean of air quality before lockdowns is calculated 90 days before the official dates of lockdowns. Control variables are daily temperature and rainfall (humidity for station-based data).

Table A6: COVID-19 lockdowns and air pollution – RDD with additional covariates

	Air pollution: NO_2			Air pollution: $PM_{2.5}$		
Bandwidth	+/-90 days	+/-60 days	+/-30 days	+/-90 days	+/-60 days	+/-30 days
Lockdown=1	-3.345***	-1.944***	-0.964***	-3.322***	-1.195	-0.240
	(0.365)	(0.299)	(0.340)	(1.075)	(1.160)	(1.146)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Means before lockdowns	22.914	23.316	22.719	64.599	66.015	67.544
Observations	272,653	182,163	90,934	76,486	48,857	23,215
R-squared	0.201	0.205	0.204	0.428	0.448	0.472

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1 level. Control variables are daily temperature and rainfall (humidity for station-based data), log of GDP per capita (constant 2010 USD), population density, log of energy consumption per capita, motor vehicles per 1,000 inhabitants, and share of electricity generated by coal power. All regressions include week dummies.

Table A7: COVID-19 lockdowns and air pollution - Weekly data

Panel A: Satellite air pollution

Air quality:	+/-15 weeks		+/-10 weeks		+/-5 weeks	
NO_2	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: Linear model						
Lockdown=1	-2.291***	-2.380***	-1.235***	0.129	-0.231	-0.216
	(0.239)	(0.239)	(0.229)	(0.467)	(0.254)	(0.257)
Model 2: Linear interaction model						
Lockdown=1	-2.087***	-2.201***	-1.163***	-1.146***	-0.150	-0.114
	(0.235)	(0.235)	(0.227)	(0.227)	(0.247)	(0.250)
Model 3: Quadratic model						
Lockdown=1	-2.110***	-2.232***	-1.196***	-1.193***	-0.172	-0.142
	(0.236)	(0.236)	(0.227)	(0.228)	(0.249)	(0.252)
Model 4: Quadratic interaction model						
Lockdown=1	-2.098***	-2.217***	-1.191***	-1.185***	-0.257	-0.240
	(0.235)	(0.236)	(0.228)	(0.228)	(0.256)	(0.259)
Means before lockdowns	22.756	22.756	23.225	23.225	22.955	22.955
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	435,571	431,344	296,819	293,832	146,864	145,443

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD. Robust standard errors in parentheses. Standard errors are clustered at ADM1 level. Air pollution is measured by NO_2 from satellite data. Model 1 uses running variable in linear form, Model 2 includes interaction of running variable and treatment variable, Model 3 includes quadratic term of running variable, Model 4 includes interactions of running variable (linear and quadratic terms) with treatment variable. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall.

Panel B: Station-based air pollution

Air quality:	+/-15 weeks		+/-10 weeks		+/-5 weeks	
$PM_{2.5}$	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: Linear model						
Lockdown=1	-4.950***	-3.191***	-2.481**	-0.756	-0.115	-1.099
	(0.914)	(0.941)	(1.054)	(0.989)	(1.131)	(1.144)
Model 2: Linear interaction model						
Lockdown=1	-4.160***	-2.597***	-1.619	-0.203	0.732	-0.560
	(0.902)	(0.919)	(1.044)	(0.979)	(1.100)	(1.129)
Model 3: Quadratic model						
Lockdown=1	-4.466***	-2.797***	-1.852*	-0.358	0.595	-0.689
	(0.905)	(0.923)	(1.046)	(0.980)	(1.098)	(1.127)
Model 4: Quadratic interaction model						
Lockdown=1	-4.173***	-2.505***	-1.461	-0.050	0.436	-0.660
	(0.907)	(0.930)	(1.044)	(0.978)	(1.141)	(1.167)
Means before lockdowns	63.477	63.477	65.666	65.666	66.682	66.682
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	107,611	95,855	73,452	61,863	38,031	28,633

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD. Robust standard errors in parentheses. Standard errors are clustered at city level. Air pollution is measured by $PM_{2.5}$ from station-based data. Model 1 uses running variable in linear form, Model 2 includes interaction of running variable and treatment variable, Model 3 includes quadratic term of running variable, Model 4 includes interactions of running variable (linear and quadratic terms) with treatment variable. All regressions include country dummies and week dummies. Control variables are daily temperature and humidity.

Table A8: COVID-19 lockdowns and air pollution – 'Regular' stringency index

	Air pollution: NO_2			Air pollution: $PM_{2.5}$		
Bandwidth	+/-90 days	+/-60 days	+/-30 days	+/-90 days	+/-60 days	+/-30 days
Lockdown=1	-2.097***	-1.219***	-0.535**	-2.665***	-0.288	-0.669
	(0.229)	(0.221)	(0.225)	(0.912)	(0.914)	(1.047)
Means before lockdowns	22.914	23.316	22.719	64.599	66.015	67.544
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	378,101	252,991	126,799	82,193	52,502	24,910
R-squared	0.386	0.397	0.411	0.525	0.562	0.603

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1/city level. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall (humidity for station-based data).

Table A9: Stringency index and air pollution – Principal Component Analysis

	Air pollution: NO_2			Air pollution: $PM_{2.5}$		
Bandwidth	+/-90 days	+/-60 days	+/-30 days	+/-90 days	+/-60 days	+/-30 days
Lockdown=1	-0.878***	-0.296	0.393*	-1.758*	1.118	-2.563**
	(0.201)	(0.233)	(0.218)	(0.910)	(1.040)	(1.157)
Means before lockdowns	22.914	23.316	22.719	64.599	66.015	67.544
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	351,376	256,242	128,826	82,214	52,795	25,344
R-squared	0.378	0.373	0.354	0.503	0.546	0.570

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1/city level. Stringency index is constructed using Principal Component Analysis. For all dimensions of stringency index, see Table A1 (Appendix). All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall (humidity for station-based data).

Table A10: Stringency index and air pollution – Alternative stringency indexes

	Air pollution: NO_2		
Bandwidth	+/-90 days	+/-60 days	+/-30 days
Panel A: Government response index			
Lockdown=1	-2.372***	-1.376***	-0.715***
	(0.231)	(0.215)	(0.248)
Observations	378,254	251,973	125,976
R-squared	0.388	0.399	0.411
Panel B: Containment and health index			
Lockdown=1	-2.463***	-1.454***	-0.872***
	(0.230)	(0.215)	(0.245)
Observations	377,986	251,983	125,978
R-squared	0.389	0.399	0.411
Panel C: Economic support index			
Lockdown=1	0.129	0.492***	0.638***
	(0.172)	(0.172)	(0.218)
Observations	323,154	242,563	125,164
R-squared	0.384	0.368	0.377
Means before lockdowns	22.914	23.316	22.719
Controls	Yes	Yes	Yes
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1 level. All indexed are taken from "display" version of OxCGRT which will extrapolate to smooth over the last seven days of the index based on the most recent complete data. All regressions include country dummies and week dummies. Control variables are daily temperature and rainfall (humidity for station-based data).

Table A11: COVID-19 lockdowns and air pollution – Country linear time trend

	Air pollution: NO_2			Air pollution: $PM_{2.5}$		
Bandwidth	+/-90 days	+/-60 days	+/-30 days	+/-90 days	+/-60 days	+/-30 days
Model 1: Linear model						
Lockdown=1	-2.116***	-1.113***	-0.441*	-2.979***	-0.347	-2.702**
	(0.232)	(0.221)	(0.238)	(0.923)	(0.916)	(1.099)
Model 2: Linear interaction model						
Lockdown=1	-2.081***	-1.067***	-0.384	-2.484***	0.104	-2.508**
	(0.231)	(0.220)	(0.235)	(0.909)	(0.903)	(1.089)
Model 3: Linear interaction model						
Lockdown=1	-2.106***	-1.092***	-0.407*	-2.575***	0.043	-2.599**
	(0.232)	(0.220)	(0.236)	(0.910)	(0.904)	(1.092)
Model 4: Quadratic interaction model						
Lockdown=1	-2.102***	-1.073***	-0.550**	-2.460***	0.067	-2.655**
	(0.231)	(0.221)	(0.240)	(0.910)	(0.903)	(1.122)
Means before lockdowns	22.914	23.316	22.719	64.599	66.015	67.544
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Country linear time trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	378,101	252,991	126,799	82,193	52,502	24,910

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1/city level. Model 1 uses running variable in linear form, Model 2 includes interaction of running variable and treatment variable, Model 3 includes quadratic term of running variable, Model 4 includes interactions of running variable (linear and quadratic terms) with treatment variable. All regressions include country dummies, week dummies, and interaction of country dummies with linear time trend. Control variables are daily temperature and rainfall (humidity for station-based data).

Table A12: COVID-19 lockdowns and air pollution – Air pollution in log form

	Air pollution: NO_2			Air pollution: $PM_{2.5}$		
Bandwidth	+/-90 days	+/-60 days	+/-30 days	+/-90 days	+/-60 days	+/-30 days
Model 1: Linear model						
Lockdown=1	-0.066***	-0.036***	-0.009	-0.050***	-0.020	-0.013
	(0.006)	(0.006)	(0.007)	(0.014)	(0.013)	(0.017)
Model 2: Linear interaction model						
Lockdown=1	-0.065***	-0.034***	-0.007	-0.045***	-0.015	-0.011
	(0.006)	(0.006)	(0.007)	(0.013)	(0.013)	(0.017)
Model 3: Linear interaction model						
Lockdown=1	-0.066***	-0.035***	-0.008	-0.046***	-0.017	-0.012
	(0.006)	(0.006)	(0.007)	(0.013)	(0.013)	(0.017)
Model 4: Quadratic interaction model						
Lockdown=1	-0.065***	-0.033***	-0.008	-0.045***	-0.016	-0.008
	(0.006)	(0.006)	(0.007)	(0.013)	(0.013)	(0.017)
Means before lockdowns	22.914	23.316	22.719	64.599	66.015	67.544
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	374,218	250,161	125,309	82,193	52,502	24,910

Notes: *** p<0.01, ** p<0.05, * p<0.1. Results of parametric RDD that include interactions of running variable (linear and quadratic) with treatment variable. Robust standard errors in parentheses. Standard errors are clustered at ADM1/city level. Model 1 uses running variable in linear form, Model 2 includes interaction of running variable and treatment variable, Model 3 includes quadratic term of running variable, Model 4 includes interactions of running variable (linear and quadratic terms) with treatment variable. All regressions include country dummies and week dummies. Air pollutants are in log form. Control variables are daily temperature and rainfall (humidity for station-based data).