



Grantham Research Institute on  
Climate Change and  
the Environment

**Estimating the effect of air pollution on road  
safety using atmospheric temperature  
inversions**

**Lutz Sager**

**October 2016**

**Grantham Research Institute on Climate Change and  
the Environment**

**Working Paper No. 251**

**The Grantham Research Institute on Climate Change and the Environment** was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has nine research programmes:

1. Adaptation and development
2. Carbon trading and finance
3. Ecosystems, resources and the natural environment
4. Energy, technology and trade
5. Future generations and social justice
6. Growth and the economy
7. International environmental negotiations
8. Modelling and decision making
9. Private sector adaptation, risk and insurance

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: <http://www.lse.ac.uk/grantham>.

This working paper is intended to stimulate discussion within the research community and among users of research, and its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper represent those of the author(s) and do not necessarily represent those of the host institutions or funders.

# Estimating the Effect of Air Pollution on Road Safety using atmospheric Temperature Inversions

Lutz Sager\*

London School of Economics

3 October 2016

**Abstract:** This paper estimates the causal effect of increased ambient air pollution on the frequency of road traffic accidents in the United Kingdom between 2009 and 2014. An instrumental variable approach is applied, exploiting atmospheric temperature inversions as a source of plausibly exogenous variation in daily air pollution levels. The paper estimates the local average treatment effect for a geographic grid cell ( $1^\circ \times 1^\circ$ ) as an increase of 0.3 accidents per day for each additional  $1\mu\text{g}/\text{m}^3$  in the daily concentration of  $\text{NO}_2$ . The effect is equivalent to an increase of 2% relative to the average number of daily accidents. The results appear robust to multiple sources of potential confounding, measurement error and co-emission bias.

**Keywords:** Air pollution, cognitive performance, health, nitrogen dioxide, temperature inversions, traffic accidents.

---

\*Grantham Research Institute on Climate Change and the Environment, Department of Geography and Environment, LSE, Houghton Street, London, WC2A 2AE, United Kingdom; telephone: +44-79-3454-3798; e-mail: [l.sager@lse.ac.uk](mailto:l.sager@lse.ac.uk). I gratefully acknowledge financial support by the Grantham Foundation for the Protection of the Environment and the UK's Economic and Social Research Council (ESRC). I thank Simon Dietz and Dominik Hangartner for valuable comments. All remaining errors are my own.

# 1. Introduction

Air pollution has long been suspected to be responsible for serious adverse health effects including respiratory illness and cardiovascular disease (Dockery and Pope, 1994; Seaton et al., 1995). Quasi-experimental methods based on observational data have helped to solidify the evidence regarding the magnitude and causal nature of the effects from exposure to air pollutants such as nitrogen dioxide ( $NO_2$ ), carbon monoxide ( $CO$ ), and small particulate matter ( $PM$ ) (for recent surveys see Currie et al., 2011; Graff Zivin and Neidell, 2013).

Recently, air pollution has also been shown to negatively affect other outcomes relevant to economic productivity and wellbeing. Graff Zivin and Neidell (2012) find that heightened levels of ozone ( $O_3$ ) negatively affect the labour productivity of agricultural workers. Lay et al. (2014) find that elevated levels of small particulate matter below  $2.5\mu m$  ( $PM_{2.5}$ ) and carbon monoxide ( $CO$ ) impair cognitive performance and reduce standardised test scores of high school students.

This paper contributes to this emerging literature concerned with the adverse effects of air pollution on productivity broadly and cognitive performance in particular. Specifically, we ask whether or not air pollution affects driving performance in road traffic. To the best of our knowledge, this paper represents the first attempt to relate air quality to road safety. In particular, we aim to identify the causal effect of an increase in air pollution, measured as the daily average concentration of  $NO_2$ , on the number of road traffic accidents within spatial grid cells (regions) of dimension  $1^\circ \times 1^\circ$  in the United Kingdom between 2009-2014. Our identification of the causal effect relies on atmospheric temperature inversions as an instrument inducing plausibly exogenous variation in pollution levels.

Our main hypothesis is that increases in air pollution result in more traffic accidents due to impaired performance of drivers, likely by affecting cognitive ability such as attention span or reaction time. This research question is highly relevant, as road traffic accidents represent a significant source of premature deaths and injuries, as well as large costs in damages, both in the United Kingdom and beyond.

This paper proceeds as follows. Section 2 discusses the existing literature. Section 3 introduces the research design. Section 4 describes the data. Section 5 presents the main results. Section 6 discusses robustness and limitations. Section 7 concludes.

## 2. Previous Literature

Air pollution refers to a variety of pollutants including carbon monoxide ( $CO$ ), nitrogen dioxide ( $NO_2$ ), sulphur dioxide ( $SO_2$ ), small particulate matter ( $PM_{2.5}$  or  $PM_{10}$ ), and ozone ( $O_3$ ). A large literature exists focusing on the adverse effects on health from exposure to air pollution (Currie et al., 2011; Graff Zivin and Neidell, 2013). Air pollution is likely responsible for serious adverse health effects such as increased risk of cardiovascular disease and respiratory illness (Seaton et al., 1995; Dockery and Pope, 1994). These health shocks appear to disproportionately affect vulnerable populations such as infants (Currie and Neidell, 2005), members of poorer households (Jans et al., 2014), and those living in developing nations (Arceo et al., 2016). The improved understanding and quantification of adverse health effects from pollution has shaped a consensus and helped inform policies to reduce air pollution.

However, more recent work has also found adverse effects outcomes beyond cardiovascular and respiratory health. Notably, it has been shown that air pollution significantly lowers the productivity of agricultural labourers (Graff Zivin and Neidell, 2012) and, in the case of  $PM_{2.5}$ , even in-door workers inside a pear-packing factory (Chang et al., 2014) and call-center (Chang et al., 2016). Lay et al. (2014) show that pollution levels inside of classrooms negatively affects test performance among high-school students. This is in line with recent findings in the medical literature indicative of adverse effects of air pollution exposure on cognitive function. For example, Chen and Schwartz (2009) find a negative association between exposure to air pollution ( $O_3$  and  $PM_{10}$ , yearly variation) on performance in neurobehavioural tests designed to measure reaction time, attention, perceptual function, and short-term memory. A large portion of medical research has thus far focused on the association between *long-term* exposure to air pollution and cognitive as well as psychological function in both children and, more recently, adults (for a recent survey see Tzivian et al., 2015). A number of these studies explicitly focus on the role of traffic as the source of pollution exposure (e.g. Freire et al., 2010; Power et al., 2011).

These findings suggest that the adverse effects from air pollution may extend well beyond cardiovascular and respiratory health. Indeed, if air pollution impedes productivity in a wide array of activities, we may currently underestimate the cost of air pollution to human well-being. This paper adds to this emerging literature by asking if air pollution

may affect productivity in another dimension - namely the ability to be a safe participant in road traffic. Arguably, safely steering a car and avoiding accidents is positively related to ones mental alertness, reaction time, and a host of other dimensions of mental acuity. Based on previous findings that pollution affects both physical and cognitive productivity, we hypothesise that exposure to air pollution may impair driving performance and consequently increase the number of traffic accidents. In doing so, we focus on contemporaneous effects from relatively short-term exposure (i.e. daily variations) rather than developmental effects from long-term exposure.

The possible effect of air pollution on driving performance is highly relevant for a number of reasons. Firstly, participating in road traffic takes up a significant amount of time for many working age adults. Consequently, traffic accidents are responsible for significant costs in the form loss of life, bodily harm, and material damages. Secondly, road vehicles constitute a major source for air pollutants and it is arguably while participating in road traffic that one tends to be exposed to one of the highest concentrations of air pollution. Knittel et al. (2016) even use variation in traffic as instrument for pollution levels in order to better estimate the causal effect of pollution on health. However, this present paper is to the best of our knowledge the first to analyse the possibility that pollution levels may affect traffic and in particular the number of traffic accidents. Finally, evidence that contemporaneous exposure to air pollution may impair driving performance may suggest further previously undiscovered costs from air pollution during activities with high risk or high cognitive strain, such as operating heavy machinery.

### 3. Research Design

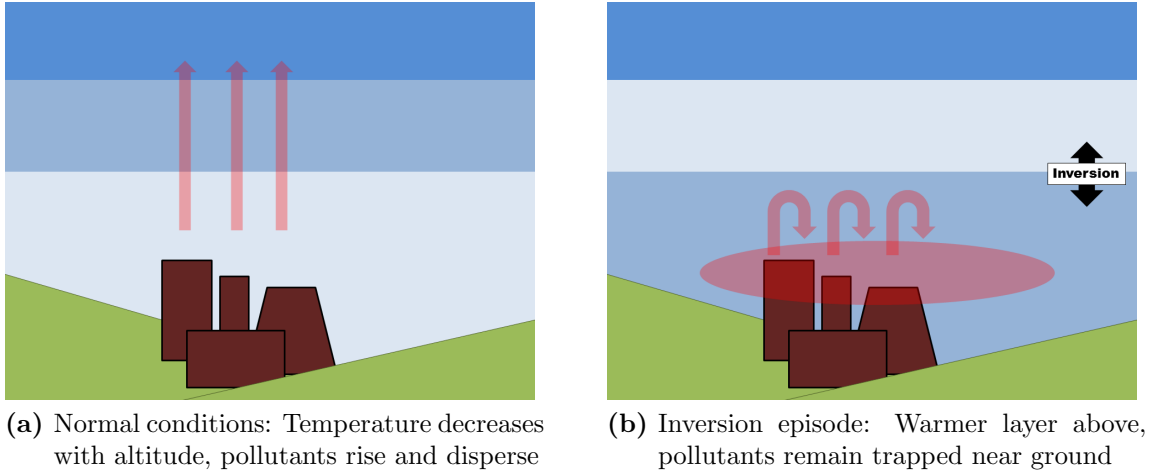
We are interested in estimating the causal effect of an increase in the level of air pollution on the number of traffic accidents. This section introduces the empirical challenges when estimating such a causal effect and presents our identification strategy alongside identifying assumptions. In doing so, we follow the potential outcome terminology of Angrist et al. (1996) by defining  $Y$  as the realised outcome (number of road traffic accidents) and  $D$  as the treatment indicator (concentration of air pollutant). We further define  $Y_1$  and  $Y_0$  as potential outcomes under treatment (heightened level of pollution) and control respectively.

We are then interested in the causal effect of an increase in pollution on the outcome, i.e.  $(Y_1 - Y_0)$ . However, we face what is known as the 'fundamental problem of causal inference', namely that we cannot observe both potential outcomes at the same time.

**The empirical challenge:** Identifying the causal effect of pollution on accidents from observed outcomes is challenging because air pollution is largely a product of human activity and we thus cannot reasonably assume that it is randomly assigned across locations. In the context of this paper, it is likely that both the number of accidents per day and the level of pollution in a certain region are related to unobserved confounding factors. These factors may be temporary, such as for example weather conditions affecting both accident frequencies and pollution levels, or structural, such as population density of regions, prevalence of road types, or speed regulations.

In addition, it is possible that there is a certain degree of reverse causation, in the sense that accidents affect traffic flows and thus air quality. In all these cases, endogeneity of pollution levels will result in biased estimates when applying simple regression analysis with OLS estimators (or simply comparing the difference between realised outcomes of treated and untreated subjects in a binary setting).

**Identification strategy:** In order to overcome the problem posed by potential endogeneity of treatment, we adopt an instrumental variable (IV) approach to estimate the causal effect of air pollution on accident frequency. We use as an instrument the occurrence of atmospheric temperature inversions. We denote  $Z$  as an instrumental variable, which takes the value 1 on inversion days and 0 otherwise.



**Figure 1:** Temperature inversions as instrument for air pollution

While on most days, temperature decreases with altitude, inversion periods are characterised by increasing temperature. During these inversion periods, the warmer air at higher altitude prevents pollutants from rising and dispersing, but rather traps them close to the ground (see Figure 1). We will argue that such inversion periods present an exogenous source of variation in air pollution levels (after controlling for potentially related weather conditions).

Our approach is most closely related to that applied by Jans et al. (2014), who use temperature inversions as an instrumental variable to estimate the effect of  $PM_{10}$  concentrations on children’s respiratory health. Similarly, Arceo et al. (2016) use the number of temperature inversions per week to estimate effects of  $PM_{10}$  and  $CO$  on child mortality in Mexico City.

**Identifying assumptions:** In order for our IV approach to yield an unbiased estimate of the suspected causal effect, certain assumptions need to be met. The identifying assumptions in this framework are:

**Assumption A1 Independence:**  $(Y_0, Y_1, D_0, D_1) \perp\!\!\!\perp Z$

**Assumption A2 First Stage:**  $0 < P(Z = 1) < 1$  and  $P(D_1 = 1) \neq P(D_0 = 1)$

**Assumption A3 Monotonicity:**  $D_1 \geq D_0$  (no defiers)

In essence, First Stage requires that the instrument  $Z$  is informative with regard to the treatment  $D$ . In our application, we require inversion episodes to significantly affect the concentration of air pollutants (*instrument relevance*). Independence requires



inversion episodes to be randomly assigned. Crucially, it implies that inversions  $Z$  only affect accident frequency  $Y$  through their effect on pollution  $D$  (*exclusion restriction*). Monotonicity rules out 'defiers', in our case regions where inversions systematically have the opposite effect on pollution than in general. We will discuss the plausibility of these assumptions alongside the results presented in Section 4.

**Estimators:** It has been shown by Imbens and Angrist (1994) that under assumptions (A1) – (A3) it is possible to estimate the local average treatment effect (LATE), which identifies average treatment for 'compliers', using the Wald estimator:

$$\alpha_{LATE} = E[Y_1 - Y_0 | D_1 > D_0] = \frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]} = \frac{cov(Y, Z)}{cov(D, Z)} \quad (1)$$

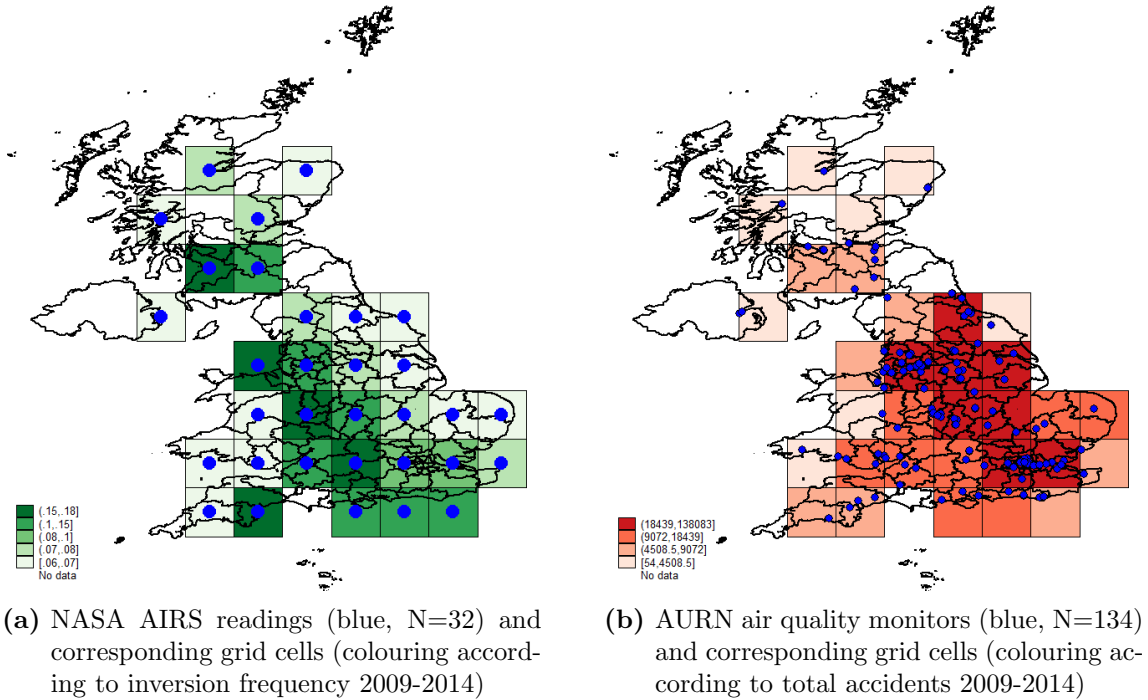
We report below the standard two-stage least squares (2SLS) estimator, which reduces to the Wald estimator in the absence of covariates. In the presence of covariates, it has been shown by Abadie (2003) that 2SLS estimates do not have a straightforward causal interpretation. However, they again represent LATE under the additional assumption that treatment effects are constant conditional on covariates  $X$  (and assumptions (1) – (3) conditional on  $X$ ). Again, results in Section 4 will be accompanied by a brief discussion of the likely causal nature of estimated effects.

**Empirical specification:** The empirical specification that we implement for 2SLS estimation is represented by the following two equations:

$$D_{it} = \tau + \rho Z_{it} + \delta' \mathbf{x}_{it} + \eta_t + \theta_i + v_{it} \quad (\text{First Stage}) \quad (2)$$

$$Y_{it} = \mu + \alpha \hat{D}_{it} + \beta' \mathbf{x}_{it} + \mu_t + \gamma_i + \epsilon_{it} \quad (\text{Second Stage}) \quad (3)$$

Here,  $Y_{it}$  is the number of accidents in region  $i$  on day  $t$ ,  $D_{it}$  is the average daily concentration of an air pollutant,  $Z_{it}$  is a binary indicator of night-time inversion,  $\mathbf{x}_{it}$  is a vector of controls for weather conditions,  $\eta_t$  and  $\mu_t$  are time fixed-effects,  $\theta_i$  and  $\gamma_i$  are region fixed-effects,  $\tau$  and  $\mu$  are constants, and  $v_{it}$  and  $\epsilon_{it}$  are error terms.



**Figure 2:** Geographic coverage of observational units

## 4. Data

We use data from four sources to construct a panel data set on air quality, traffic accidents, and weather conditions in the United Kingdom between 2009 and 2014.

### 4.1. Data Sources

**Atmospheric climate data:** This paper uses data extracts from NASA’s Atmospheric Infrared Sounder (AIRS) launched onboard the NASA satellite AQUA in 2002. With the primary intention of improving weather forecasts, the AIRS produces bi-daily a 3-dimensional (the third dimension are vertical pressure layers) map of atmospheric temperature and water vapor. Readings are provided for an ascending phase (ca. 2pm local time) and descending phase (2am). We use the readily available L3 data of AIRS Version 6 (AIRS Science Team/Joao Teixeira, 2015), which provides readings for 24 layers at a spatial resolution of  $1^\circ \times 1^\circ$  (ca.  $110km \times 70km$ ). We use data on the two pressure levels closest to the earth’s surface ( $1000hPa$  at sea level and  $925hPa$  at ca.  $600m$  above sea level). Specifically, we use readings from the descending phase (ca. 2am local time) to construct a binary indicator of inversion periods when the temperature at the  $925hPa$ -level exceeds that of the  $1000hPa$ -level. We follow Jans et al. (2014) in focusing on nighttime

inversions to counter concerns that daytime inversions may be perceivable for road traffic participants (inversions during daylight can be occasionally be seen with the naked eye) and because these inversions occur before the bulk of recorded traffic accidents on a given day. Furthermore, inversions are more frequent at night. We complement information on night-time inversion periods by readings on relative humidity at the surface over liquid phase (in per cent) and cloud fraction for all pressure levels combined (between 0 – 1).

Our main unit of observation are the NASA L3 grid cells and the geographic coverage is most of the United Kingdom (spatial box  $49N, 10W, 61N, 2E$ ), observed for the 2,191 days of the years 2009–2014. Figure 2a provides a visual representation of the geographic coverage of grid cells used in the final analysis.

**Air pollution data:** Air pollution data come from the United Kingdom Automatic Urban and Rural Network (AURN), which includes automatic air quality monitoring stations measuring a variety of air pollutants. The data is provided by the Department for Environment, Food and Rural Affairs (2016). The network comprises a total of 198 monitoring sites with data available at different degrees since 22 February 1973. For the 2,191 days in the period between 2009-2014, we obtain daily readings for 150 sites, of which 134 report nitrogen dioxide ( $NO_2$ ) levels (see Figure 2b) and 75 report concentrations of small particulate matter below  $10\mu m$  ( $PM_{10}$ ).

**Road accident statistics:** Observations for road traffic performance are obtained from the Road Safety Data maintained by the United Kingdom Department of Transport (Department for Transport, 2015). This publicly available data provides yearly lists of road traffic accidents involving personal injury which have been reported to the police and recorded using STATS19 forms. For the six years between 2009 and 2014, we have obtained details for a total of 899,995 individual accidents with exact timing and geo-coordinates (longitude to 6 digits after the comma, latitude to 5 digits). Each accident record is complemented by details about road type, vehicle make, casualty count, as well as age and sex of drivers' involved. For the purpose of this paper, we collapse these data to daily counts of reported accidents within each geographic grid cell (see Figure 2b).

**Additional weather data:** Finally, we complement atmospheric climate data with ground-level observations on weather conditions from the United Kingdom Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations (Met Office, 2012). We use the daily amount of rainfall over a 24 hour period in *mm* (usually 09:00-09:00) and hourly readings of wind speed in *knots* averaged over the day (00:00-00:00).

## 4.2. Descriptive Statistics

The final sample used in the analysis consists of 41,651 observations, each representing one grid cell on a given day, for which at least one non-zero accident count,  $NO_2$  concentration reading, and atmospheric temperature readings were available.

Summary statistics for key variables are presented in Table 1. The average for the number of accidents per grid cell per day is 14.05 and the average  $NO_2$  concentration is  $23.69\mu g/m^3$ . Inversion periods are observed 9.6% of the time. Data availability introduces a likely sample bias, as we expect less populated and more rural grid cells to more likely have no reported accidents on a given day or suffer from missing values in pollution levels. In sum, our sample accounts for 585,076 (i.e. 65% of the total 899,995) accidents reported between 2009 and 2014.

**Table 1:** Descriptive statistics

	(1) Observations	(2) Mean	(3) Std. Deviation	(4) Minumum	(5) Maximum
# of accidents	41,651	14.05	18.67	1	157
NO2	41,651	23.69	15.34	0	147.7
PM10	28,714	19.24	10.61	0	118
Inversion (binary)	41,651	0.0959	0.294	0	1
Temperature	41,651	280.5	4.848	259.2	294.8
Humidity	39,925	85.38	7.876	40.25	100
Cloud coverage	41,651	0.505	0.263	0	1
Wind speed	37,800	8.037	3.678	0	31.42
Rainfall	41,651	1.629	3.230	0	57.19

Notes: The unit of observation is one grid cell ( $1^\circ \times 1^\circ$ ) on one day. Data sources are the UK Department of Transport for the number of accidents; UK DEFRA / AURN for  $NO_2$  and  $PM_{10}$  data ( $\mu g/m^3$ , averaged for all available readings per grid cell per day); NASA AIRS for Temperature (Kelvin,  $1000hPa$  pressure layer), Cloud coverage (fraction), Humidity (percentage, relative), and Inversions (binary indicator equal to 1 when Temperature in  $925hPa$  is greater than Temperature in  $1000hPa$  layer); and UK MIDAS for Rainfall (*mm*, daily total) and Wind speed (Knots, daily average of hourly readings).

## 5. Results

This section presents results of the empirical model detailed in Section 2. The causal nature and economic significance of estimates is assessed, and threats to internal validity are discussed. Section 5 provides further robustness tests and discusses limitations.

### 5.1. Baseline Results

As outlined above, this paper aims to estimate the causal effect of increased concentrations of air pollutants on the frequency of road traffic accidents in the United Kingdom between 2009-2014.

**Table 2:** First Stage - Effect of inversions on pollution

	(1) OLS	(2) OLS	(3) OLS
Mean NO <sub>2</sub>	23.69	23.52	23.52
Inversion (binary)	5.011*** (0.272)	1.525*** (0.426)	1.126*** (0.246)
Effect	21.15%	6.48%	4.79%
Weather Covariates		YES	YES
Region FE		YES	
Region-Year FE			YES
Weekday FE			YES
Month FE			YES
Observations	41,651	36,274	36,274
R-squared	0.009	0.333	0.498
F-test	339	34.99	64.76

Notes: This table provides estimates of the effect of inversion episodes on the average daily  $NO_2$  concentration in a grid cell. Robust standard errors are reported in parentheses (\*\* $p < 0.01$ , \* $p < 0.05$ , \* $p < 0.1$ ). A detailed version of this table including coefficients for weather covariates can be found in the Appendix.

**First stage estimates:** Table 2 presents results from the first stage regression (equation 2) assessing the effect of inversion periods on daily  $NO_2$  pollution levels. Column 1 shows that days on which night-time inversions occur have on average a concentration of  $NO_2$  which is  $5.011\mu g/m^3$  higher than on non-inversion days. This amounts to an increase of 21.15% relative to the sample average. However, one concern for the analysis presented is that inversion episodes may be correlated with other weather conditions

that affect pollution and accidents. For example, inversion frequency, accident count and pollution levels may all systematically differ between regions.

Column 2 shows that when controlling for weather conditions (ground-level temperature, humidity, cloud coverage, wind speed, and rainfall) as well as region fixed effects, the coefficient is lower at  $1.525\mu\text{g}/\text{m}^3$ . It is thus apparent that some of the correlation between inversion episodes and pollution levels is related to other weather conditions and time-invariant characteristics of grid cells. Column 3 reports results for the first stage regression, when letting regional fixed effects vary by year (to account for structural breaks) and further controlling for month fixed effects (as both inversion frequency and pollution levels vary by season) and day of the week. An inversion episode is then estimated to relate to an increase of  $1.126\mu\text{g}/\text{m}^3$  in daily average  $\text{NO}_2$  concentrations (4.79% relative to the average).

In all specifications, coefficients of night-time inversions are statistically significant at the 99%-level and F-tests exceed 30 in each case. While the relationship with other weather conditions is non-trivial, we conclude that the First Stage assumption (A2) holds and night-time inversions represent a relevant instrument for daily  $\text{NO}_2$  concentrations.

**Table 3:** Main Results - Effect of pollution on accidents

	(1)	(2)	(3)	(4)
	OLS	LATE	2SLS	2SLS
Mean # Accidents	14.05	14.05	15.10	15.10
$\text{NO}_2$	0.498***	0.276***	0.279***	0.273***
	(0.00893)	(0.0589)	(0.0773)	(0.103)
Effect	3.55%	1.96%	1.85%	1.81%
Weather Covariates			YES	YES
Region FE			YES	
Region-Year FE				YES
Weekday FE				YES
Month FE				YES
Observations	41,651	41,651	36,274	36,274

Notes: This table provides estimates of the effect of an increase in  $\text{NO}_2$  concentration on the number of accidents per grid cell per day. Estimates in Column 1 are from a simple OLS estimator. Estimates in Columns 2-4 are from 2SLS estimators using the occurrence of temperature inversions on a given day as instrumental variable for pollution. Robust standard errors are reported in parentheses (\*\*\*)  $p < 0.01$ , (\*\*)  $p < 0.05$ , (\*)  $p < 0.1$ ). A detailed version of this table including coefficients for weather covariates can be found in Appendix D.

**Causal estimates of the effect of pollution on accidents:** Table 3 reports the main results of this paper - estimates of the causal effect of  $NO_2$  pollution on the number of daily traffic accidents in a grid cell. Column 1 reports the result from a simple OLS regression, which is included for comparison. As discussed in Section 2, this estimate likely suffers from multiple sources of endogeneity bias, in particular due to reverse causation and confounding from factors such as road network density.

To obtain a more reliable estimate of the causal effect in question, estimates from an instrumental variable approach are reported in Columns 2-4. Column 2 reports the 2SLS estimate from an IV specification without covariates (separate second stage results can be found in Appendix D alongside the detailed versions of Tables 2 and 3). The interpretation of the estimate is as follows: An increase of  $1\mu g/m^3$  in the daily average concentration of  $NO_2$  in a grid cell ( $1^\circ \times 1^\circ$ ), on average causes an increase of 0.276 in the number of reported road accidents in that cell on that day (an increase of 1.96% relative to the mean). We interpret the significant difference to the OLS estimate as a first indication of bias in the OLS model. Under assumptions (1) – (3), the 2SLS estimate has a direct causal interpretation as the local average treatment effect (LATE).

**Internal validity:** We have already demonstrated, both in theory (Figure 1) and empirically (Table 2), that the First Stage (A2) assumption likely holds. Monotonicity (A3) is not a testable assumption. However, we would like to argue that, given the nature of the instrument as a weather phenomenon, it is likely to hold that there are 'no defiers' in the sense of regions in which night-time inversions always result in lower pollution levels.

The independence assumption (A1) is crucial to the causal interpretation and unbiasedness of the reported estimates. While it is not testable, we discuss below our case for independence in the given application. We argue that night-time inversions present a source of plausibly exogenous variation in daily air pollution levels. In particular, we require it to be true that inversions do not affect road safety, except through their effect on pollution levels (exclusion restriction).

This might be violated if inversion episodes themselves affect the performance of traffic participants. We are not aware of any evidence showing adverse effects of inversion episodes on well-being or cognitive performance. It is also possible that inversions coincide with other weather patterns that affect road safety or that inversions are more frequent in

regions or at times with a higher number of accidents. While independence (A1) cannot be tested for formally, results from additional model specifications presented in Table 3 further strengthen our confidence in the assumption. They are discussed below.

## 5.2. Alternative Specifications

**Weather confounding:** One possible threat to independence might be that inversion periods are related to other weather phenomena that affect the accident frequency. In addition, it is possible that regions systematically differ in terms of weather patterns (including inversion frequency) as well as structural characteristics affecting accident rates (e.g. road network density between coastal and mountainous regions). To control as much as possible for these threats to independence, Column 3 of Table 3 reports results from a model specification which includes additional weather controls and grid cell fixed effects.

Indeed, comparison between Columns 1 and 2 in Table 2 indicates that the magnitude of the correlation between inversions and pollution is reduced when controlling for weather covariates. However, the estimated effect of  $NO_2$  on accidents (0.279) in Column 3 Table 3 is not significantly different from that without covariates.

**Seasonality and time trends:** Column 4 of Table 3 presents results from a specification with added time fixed effects. These are intended to control for confounding over time, especially by including month fixed effects to control for seasonality of inversions (inversion frequency ranges from 2% in August/September to 23% in March/April) which may correlate with seasonality in pollution and accidents. Further fixed effects control for variation by weekday and for possible time trends in both accidents and inversions (inversion frequency ranges from 6% in 2014 to 13% in 2012) by allowing for region fixed effects per year. Again, the estimated effect of  $NO_2$  on accidents (0.273) is not significantly different from that in Column 2.

It is important to note that estimates in Columns 3 and 4 of Table 3 do not have a straightforward causal interpretation due to added covariates. However, the fact that these much more detailed specifications produce estimates which are not significantly different from those in Column 2 strengthens our confidence that the baseline IV specification reported in Column 2 produces reliable estimates of the LATE.



## 6. Robustness and Limitations

We have estimated a positive causal effect of air pollution (measured by  $NO_2$ ) on the number of traffic accidents in the United Kingdom between 2009-2014. Below, we present further robustness exercises and discuss limitations of the results.

### 6.1. Measurement Error

A recurring problem when estimating adverse effects from pollution is likely measurement error in attributing pollution levels to units of observation, often times individuals who may move between locations or actively avoid higher levels of pollution. We believe this is less of a concern in this paper for the following reasons.

Firstly, the instrumental variable approach is robust to measurement error given assumptions (1) – (3). Secondly, we focus on large geographic areas ( $1^\circ \times 1^\circ$ ) as the unit of observation and our source of exogenous variation is based on large-scale weather patterns. While we cannot control for all types of avoidance behaviour (such as rolling up car windows), we believe it is reasonable to assume that there is no systematic shift of population or traffic activity between grid cells in response to weather patterns. Finally, attenuation bias from measurement error would result in our estimates being too low, and thus does not threaten the significance of the results.

Table 4 in the Appendix replicates Table 3 using a different methodology of assigning pollution levels to observations. Instead of using unweighted averages of all monitor readings within a grid cell on a given day, we use distance-weighted averages (as proposed by e.g. Currie and Neidell, 2005)<sup>1</sup>. Coefficient estimates appear to be robust (0.283 – 0.285) to such an alternative approach.

### 6.2. Heterogeneity of Treatment Effect

We have so far focused on estimating the local average treatment effect (LATE), under the additional assumption that treatment effects are constant when covariates are included. However, a concern in analyses of adverse effect of pollution is in heterogenous and non-linear effects. It is not a priori established whether the effect of an additional

---

<sup>1</sup>We construct averages where monitor readings are weighted by the inverse distance from the centroid of a given grid cell. Distances are calculated using the Haversine formula to find the great-circle distance in kilometers between two coordinate-points.

unit of pollution on the number of reported accidents should be stable across different regions of the United Kingdom. We report in Table 5 in the Appendix a replication of the results from Table 3 using different sub-samples.

Crucially, one may hypothesise that rather than a linear effect, the effect of pollution may be proportional to a baseline number of accidents. In Column 2 we thus report 2SLS results when the number of accidents is normalised per grid cell (with 1 equal to the long-term average). We find a 2.12% increase in accidents per additional  $\mu\text{g}/\text{m}^3$  of  $\text{NO}_2$ , which is very similar to the 1.8 – 1.9% estimates in Table 3. To assess if there is evidence for proportional effects, we report in Columns 3 and 4 separate results for two samples of grid cells, one with low and one with high average accident counts. While there is some divergence in estimates (0.190 vs. 0.422), we cannot conclude with statistical significance that the coefficients are different.

Another concern are non-linear dose-response functions. It is possible that additional units of pollution are more harmful when concentrations are already high. Columns 5 and 6 report results for separate samples, one with low average  $\text{NO}_2$  concentrations ( $< 25\mu\text{g}/\text{m}^3$ ), one with high ones ( $> 25\mu\text{g}/\text{m}^3$ ). Again, the coefficient estimates diverge somewhat, but we cannot conclude a significant difference in estimates. Finally, there is a high correlation between those regions with higher accident counts and higher pollution levels, which are both likely related to road density. More analysis of differential treatment effects is clearly warranted, but our finding of positive causal effects seem robust across regions.

### 6.3. Pollutant Co-emission

A further problem in the estimation of adverse effects from air pollution comes from the fact that multiple air pollutants, such as  $\text{NO}_2$ ,  $\text{CO}$ , or  $\text{PM}_{10}$ , are often emitted from the same source and are highly correlated. Observed adverse effects from pollution may then be due to any one or several of these pollutants.

Such pollutant co-emission poses a challenge to the causal interpretation of our results, because it seems likely that inversion periods also affect concentrations of pollutants other than  $\text{NO}_2$ . Table 6 in the Appendix shows that night-time inversions do indeed also positively affect the concentration of  $\text{PM}_{10}$  on a given day. Such small particulate matter has been linked to adverse effects on health and productivity and thus poses a threat to

the exclusion restriction (i.e. independence in the potential outcomes framework) when estimating effects of  $NO_2$ .

We believe that our findings demonstrate a causal effect of air pollution (which we measured as  $NO_2$ ) on accident frequency, but we cannot separately identify which pollutant is responsible for which portion of the effect. In order to disentangle effects to a degree, we propose an approach based on the physical properties of different pollutants. In doing so, we follow Schlenker and Walker (2011) and Knittel et al. (2011) who rely on multiple instruments to disentangle effects of multiple pollutants with different physical properties.

As shown in Table 6 in the Appendix, contemporaneous inversions raise levels of both  $NO_2$  and  $PM_{10}$ . In addition,  $PM_{10}$  levels are also positively affected by inversion periods on the previous day, which is not the case for  $NO_2$ . Our interpretation is that  $NO_2$ , as a gaseous substance, is less persistent and does not re-group once dispersed. Meanwhile,  $PM_{10}$ , consisting of solid particles, can accumulate as dust on surfaces and be re-suspended by air motion. We exploit this property in a model specification with two endogenous variables ( $PM_{10}$ ,  $NO_2$ ) and two instruments (contemporaneous inversions and previous-day inversions).

Results of this exercise are presented in Table 7 in the Appendix. Column 1 repeats the specification from Table 3 Column 4 with a smaller sample for which  $PM_{10}$  readings are available. Column 2 shows that  $PM_{10}$ , if used as sole measure of air quality, is attributed a positive effect on accidents. However, Column 3 shows that the effect of  $PM_{10}$  is not significantly different from zero when instrumenting for both  $NO_2$  and  $PM_{10}$ . The effect ascribed to  $NO_2$  is similar in magnitude to estimates in Table 2, albeit at a lower significance level of 90%. We interpret these results as indicative that the causal effect of air pollution on road traffic accidents more likely stems from  $NO_2$  or similar gaseous pollutants, rather than particulate matter.

## 6.4. External Validity

While we believe that our results show strong evidence for a causal effect of air pollution (measured as concentration of  $NO_2$ ) on road safety, external validity is limited in some ways. Firstly, our period of observation are the six years between 2009 and 2014. Any extrapolation to other time periods needs to take into account likely time trends in air quality, automotive technology, road networks and traffic policies.

While we have constructed our sample starting from all reported accidents within the United Kingdom, data availability has likely introduced a certain sampling bias away from grid cells with missing data on air quality, weather conditions, or no reported accidents on any given day. The observed causal effect is thus likely more representative of more densely populated areas. Nevertheless, our sample accounts for 65% of reported accidents between 2009-2014 in much of the United Kingdom.

Extrapolation to other countries is subject to further limitations. We believe that the biological effects of air pollutants on humans, including their ability to drive a car, are likely similar across the world. However, differences in road networks, traffic policies, automotive technologies, and weather conditions would likely result in different magnitudes of effects. Furthermore, different levels of air quality, in particular in emerging economies, might result in different causal effects if one suspects a nonlinear dose-response function from air pollution.

In our analysis, we focus on daily variation in pollution levels, and thus estimate relatively short-term effects. We can neither draw conclusions about long-term exposure effects from living permanently in a polluted area, nor can we infer whether or not the effects observed occur immediately or are of a cumulative nature, resulting from continued exposure to pollution over the span of multiple hours.

Finally, our units of observation are large geographic grid cells ( $1^\circ \times 1^\circ$ ). Extrapolation to other units of observation, such as specific road segments or individual traffic participants, is likely to be erroneous, as both air pollution and road safety vary significantly within regions. For example, we would expect very different effects when comparing a major intersection and a recreational park without accessible roads. Nevertheless, we believe that the robust finding of an average causal effect for larger geographic areas is highly relevant, as such areas can be subject to policies affecting both air quality and road safety.

## 7. Conclusion

This paper assesses the link between pollution and road safety using data from multiple sources, including atmospheric climate data from NASA. The initiating hypothesis was that an increase in ambient air pollution levels impairs drivers' ability to drive safely, thus resulting in an increase in the number of accidents. In order to identify the causal effect of pollution on accident frequency, we have adopted an instrumental variable approach, which relies on plausibly exogenous variation in pollution levels arising from atmospheric temperature inversions.

Our findings indicate a positive and likely causal effect of air pollution in the form of  $NO_2$  on the number of road traffic accidents. Given the identifying assumptions, a  $1\mu g/m^3$  increase in daily average concentration of  $NO_2$  is estimated to cause an increase of 0.3 in the number of daily accidents in the average geographic grid cell ( $1^\circ \times 1^\circ$ ). This represents a significant increase of ca. 2% relative to the mean. This estimate is remarkably robust to different model specifications and robustness tests. To illustrate the magnitude of this effect, we consider the conservative example of the grid cell with the most accidents, containing a large portion of (West) Greater London (centered on  $51^\circ 30'N$   $00^\circ 30'W$ ). This region has a mean level of  $NO_2$  of 48.0 and standard deviation of 14.3. Our estimates suggest that a one standard deviation reduction in  $NO_2$  levels for one day may prevent around 4 accidents in that area alone.

This paper adds to the literature, which has established adverse effects of air pollution on health outcomes and, more recently, cognitive performance. To our knowledge, this paper is the first to assess the effect of air pollution on road safety. We believe that our analysis identifies a causal effect of pollution on road safety, but can only speculate regarding the exact mechanisms involved. Our initial hypothesis has been that pollution impairs drivers' fitness, just as it has previously been shown to affect cognitive performance in other areas. However, we cannot exclude other possible explanations, such as pollution causing physical distractions (e.g. itching nose) or impaired visibility.

Whichever the exact mechanisms involved, the robust finding of a significant effect of air quality on road safety is of strong relevance, given the high cost in damages, well-being and life related to traffic accidents in the United Kingdom every day. These costs may well represent another strong case for the benefits of reducing air pollution levels in addition the costs to health identified by previous research.

## References

- Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113:231–63.
- AIRS Science Team/Joao Teixeira (2015). AIRS/Aqua L3 Daily Standard Physical Retrieval (AIRS-only) 1 degree x 1 degree V006, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). Accessed 11 March 2016 at [10.5067/AQUA/AIRS/DATA303](https://disc.gsfc.nasa.gov/datasets/AQUA/AIRS/DATA303).
- Angrist, J., Imbens, G., and Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91:444–472.
- Arceo, E., Hanna, R., and Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal*, 126(591):257–280.
- Chang, T., Graff Zivin, J., Gross, T., and Neidell, M. (2014). Particulate pollution and the productivity of pear packers. NBER Working Paper 19944, National Bureau of Economic Research.
- Chang, T., Graff Zivin, J., Gross, T., and Neidell, M. (2016). The effect of pollution on worker productivity: Evidence from call-center workers in China. NBER Working Paper 22328, National Bureau of Economic Research.
- Chen, J.-C. and Schwartz, J. (2009). Neurobehavioral effects of ambient air pollution on cognitive performance in us adults. *Neurotoxicology*, 30(2):231–239.
- Currie, J., Heep Ray, S., and Neidell, M. (2011). Quasi-experimental studies suggest that lowering air pollution levels benefits infants’ and children’s health. *Health Affairs*, 30(12):2391–2399.
- Currie, J. and Neidell, M. (2005). Air pollution and infant health: What can we learn from California’s recent experience? *Quarterly Journal of Economics*, 120(3):1003–1030.
- Department for Environment, Food and Rural Affairs (2016). Automatic Urban and Rural Network. Retrieved 11 March 2016 at <https://uk-air.defra.gov.uk/networks/network-info?view=aurn>.

- Department for Transport (2015). Road Accident Data, 2014. UK Data Service. SN: 7752. <http://dx.doi.org/10.5255/UKDA-SN-7752-1>.
- Dockery, D. and Pope, C. (1994). Acute respiratory effects of particulate air pollution. *Annual Review of Public Health*, 15:107–132.
- Freire, C., Ramos, R., Puertas, R., Lopez-Espinosa, M.-J., Julvez, J., Aguilera, I., Cruz, F., Fernandez, M.-F., Sunyer, J., and Olea, N. (2010). Association of traffic-related air pollution with cognitive development in children. *Journal of epidemiology and community health*, 64(3):223–228.
- Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673.
- Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3):689–730.
- Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62:467–476.
- Jans, J., Johansson, P., and Nilsson, J. (2014). Economic status, air quality, and child health: Evidence from inversion episodes. IZA Discussion Paper No. 7929, IZE.
- Knittel, C., Miller, D., and Sanders, N. (2011). Caution, drivers! Children present: Traffic, pollution, and infant health. NBER Working Paper 17222, National Bureau of Economic Research.
- Knittel, C., Miller, D., and Sanders, N. (2016). Caution, drivers! Children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, 98(2):350–366.
- Lay, V., Ebenstein, A., and Roth, S. (2014). The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. NBER Working Paper 20648, National Bureau of Economic Research.
- Met Office (2012). Met Office Integrated Data Archive System (MIDAS) land and marine surface stations data (1853-current). NCAS british atmospheric data centre, retrieved on 16 March 2016.

- Power, M., Weisskopf, M., Alexeeff, S., Coull, B., Spiro, A., and Schwartz, J. (2011). Traffic-related air pollution and cognitive function in a cohort of older men. *Environmental Health Perspectives*, 119(5):682.
- Schlenker, W. and Walker, W. (2011). Airport, air pollution, and contemporaneous health. NBER Working Paper 17684, National Bureau of Economic Research.
- Seaton, A., Godden, D., MacNee, W., and Donaldson, K. (1995). Particulate air pollution and acute health effects. *The Lancet*, 345(8943):176–178.
- Tzivian, L., Winkler, A., Dlugaj, M., Schikowski, T., Vossoughi, M., Fuks, K., Weinmayr, G., and Hoffmann, B. (2015). Effect of long-term outdoor air pollution and noise on cognitive and psychological functions in adults. *International Journal of Hygiene and Environmental Health*, 218(1):1–11.



## A. Measurement Error - Distance-weighting

**Table 4:** Main Results - Effect of pollution on accidents

	(1)	(2)	(3)	(4)
	OLS	LATE	2SLS	2SLS
Mean # Accidents	14.05	14.05	15.10	15.10
NO <sub>2</sub>	0.484***	0.285***	0.288***	0.283***
	(0.00901)	(0.0609)	(0.0799)	(0.107)
Effect	3.44%	2.03%	1.91%	1.88%
Weather Covariates			YES	YES
Region FE			YES	
Region-Year FE				YES
Weekday FE				YES
Month FE				YES
Observations	41,651	41,651	36,274	36,274

Notes: This table replicates Table 3, but where  $NO_2$  levels are constructed using weights according to the inverse distance of an air quality monitor to the grid cell centroid. This table provides estimates of the effect of the average  $NO_2$  concentration on the number of accidents per grid cell per day. Estimates in Column 1 are from a simple OLS estimator. Estimates in Columns 2-4 are from 2SLS estimators using the occurrence of temperature inversions on a given day as instrumental variable for pollution. Robust standard errors are reported in parentheses (\*\*\*)  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ).

## B. Heterogenous Treatment - Sub-sample analysis

**Table 5:** Main Results - Effect of pollution on accidents

	(1)	(2)	(3)	(4)	(5)	(6)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Sample	full	normalised	$\bar{Y} < 10$	$\bar{Y} > 10$	$\bar{D} < 25$	$\bar{D} > 25$
Mean # Accidents	15.10	1 (15.10)	4.99	26.38	6.45	23.60
NO2	0.273***	0.021***	0.190**	0.422**	0.156*	0.361**
	(0.103)	(0.0072)	(0.0768)	(0.197)	(0.0854)	(0.170)
Effect	1.81%	2.12%	3.81%	1.60%	2.42%	1.53%
Weather Covariates	YES	YES	YES	YES	YES	YES
Region-Year FE	YES	YES	YES	YES	YES	YES
Weekday FE	YES	YES	YES	YES	YES	YES
Month FE	YES	YES	YES	YES	YES	YES
Observations	36,274	36,274	19,135	17,139	17,986	18,288

Notes: This table replicates Table 3 with different sub-sets of the sample. Column 1 re-states the results with the full sample. Column 2 states results with the full sample, but where accident counts have been normalised by dividing through the average count of the grid cell over the period 2009-2014. Column 3 reports results for those grid cells with a 2009-2014 daily average accident count below 10 (20 out of 32), while Column 4 reports results for grid cells with more average accidents (12 out of 32). Column 5 reports results for grid cells with a 2009-2014 daily average of  $NO_2$  below  $25\mu g/m^3$  (17 out of 32), while Column 6 reports results for those grid cells with higher average concentrations (15 out of 32). This table provides estimates of the effect of the average  $NO_2$  concentration on the number of accidents per grid cell per day. Estimates in Column 1 are from a simple OLS estimator. Estimates in Columns 2-4 are from 2SLS estimators using the occurrence of temperature inversions on a given day as instrumental variable for pollution. Robust standard errors are reported in parentheses (\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ).

## C. Pollutant Co-emission - Joint Estimation

**Table 6:** Pollutant Co-emission - Effect of inversions on  $NO_2$  and  $PM_{10}$

	(1) OLS (NO2 only)	(2) OLS (PM10 only)
Mean NO2	26.38	
Mean PM10		19.83
Inversion (binary)	1.843*** (0.310)	4.814*** (0.269)
Inversion (binary) [previous day]	-0.295 (0.250)	2.492*** (0.271)
Weather Covariates	YES	YES
Region-Year FE	YES	YES
Weekday FE	YES	YES
Month FE	YES	YES
Observations	20,325	20,325
R-squared	0.483	0.323
F-test	81.32	96.96

Notes: This table shows first stage results when estimating the effect of two instruments, contemporaneous inversions and lagged inversions (one day delayed), on concentrations of  $NO_2$  and  $PM_{10}$  respectively. Robust standard errors are reported in parentheses (\*\* $p < 0.01$ , \* $p < 0.05$ , \* $p < 0.1$ ).

**Table 7:** Pollutant Co-emission - Joint effect of  $NO_2$  and  $PM_{10}$  on accidents

	(1) 2SLS (NO2 only)	(2) 2SLS (PM10 only)	(3) 2SLS (Both)
NO2	0.306*** (0.114)		0.470* (0.253)
PM10		0.0568** (0.0279)	-0.0483 (0.0634)
Weather Covariates	YES	YES	YES
Region-Year FE	YES	YES	YES
Weekday FE	YES	YES	YES
Month FE	YES	YES	YES
Observations	20,325	20,325	20,325

Notes: This table presents results for a joint estimation of the effect of  $NO_2$  and  $PM_{10}$  for a sub-sample for which readings for both pollutants are available. Columns 1-2 repeat the specification from Table 3 Column 4, but with a lagged binary indicator of inversion episodes as second instrument and with  $NO_2$  and  $PM_{10}$  as treatment variable respectively. Column 3 reports 2SLS results from joint estimation with  $NO_2$  and  $PM_{10}$  as treatment variables and both instruments. Robust standard errors are reported in parentheses (\*\* $p < 0.01$ , \* $p < 0.05$ , \* $p < 0.1$ ).

## D. Detailed Regression Tables

**Table 8:** First Stage - Effect of inversions on pollution

	(1)	(2)	(3)
	OLS	OLS	OLS
Mean NO <sub>2</sub>	23.69	23.52	23.52
Inversion (binary)	5.011***	1.525***	1.126***
	(0.272)	(0.426)	(0.246)
Effect	21.15%	6.48%	4.79%
Temperature (ground)		-7.124***	-10.32***
		(2.555)	(1.766)
Temperature squared (ground)		0.0112**	0.0185***
		(0.00457)	(0.00314)
Humidity		-0.142***	-0.0508***
		(0.0118)	(0.00732)
Cloud coverage		-1.571***	-0.699***
		(0.290)	(0.198)
Rainfall		-0.173***	-0.167***
		(0.0542)	(0.0311)
Rainfall squared		0.00623**	0.00775***
		(0.00244)	(0.00157)
Wind speed		-3.033***	-2.776***
		(0.331)	(0.137)
Wind speed squared		0.0829***	0.0659***
		(0.0130)	(0.00528)
Constant	23.21***	1,169***	1,492***
Region FE		YES	
Region-Year FE			YES
Weekday FE			YES
Month FE			YES
Observations	41,651	36,274	36,274
R-squared	0.009	0.333	0.498
F-test	339	34.99	64.76

Notes: This table provides estimates of the effect of inversion episodes on the average daily  $NO_2$  concentration in a grid cell. Robust standard errors are reported in parentheses (\*\*\*)  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ).

**Table 9:** Second Stage - Effect of inversions on accidents

	(1)	(2)	(3)
	OLS	OLS	OLS
Mean Accident Count	14.05	15.10	15.10
Inversion (binary)	1.382*** (0.318)	0.425** (0.181)	0.308*** (0.112)
Inversion Effect	9.84%	2.82%	2.04%
Temperature (ground)		4.052** (1.717)	2.452 (1.573)
Temperature squared (ground)		-0.00704** (0.00300)	-0.00412 (0.00279)
Humidity		0.0183 (0.0132)	0.0160** (0.00707)
Cloud coverage		-1.064*** (0.268)	-1.117*** (0.158)
Rainfall		0.111*** (0.0282)	0.0990*** (0.0192)
Rainfall squared		-0.00422*** (0.00132)	-0.00359*** (0.000855)
Wind speed		-0.170*** (0.0572)	-0.126*** (0.0377)
Wind speed squared		0.00530** (0.00238)	0.00362** (0.00159)
Constant	13.91***	-568.2**	-352.5
Region FE		YES	
Region-Year FE			YES
Weekday FE			YES
Month FE			YES
Observations	41,651	36,274	36,274
R-squared	0.000	0.014	0.107
F test model	18.87	7.797	8.657

Notes: This table provides estimates of the effect of inversion episodes on the number of accidents per grid cell per day. Robust standard errors are reported in parentheses (\*\* $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ ).

**Table 10:** Main Results - Effect of pollution on accidents

	(1)	(2)	(3)	(4)
	OLS	LATE	2SLS	2SLS
Mean # Accidents	14.05	14.05	15.10	15.10
NO2	0.498*** (0.00893)	0.276*** (0.0589)	0.279*** (0.0773)	0.273*** (0.103)
Effect	3.55%	1.96%	1.85%	1.81%
Temperature (ground)			6.040*** (0.967)	5.271*** (1.327)
Temperature squared (ground)			-0.0102*** (0.00166)	-0.00916*** (0.00237)
Humidity			0.0579*** (0.0119)	0.0299*** (0.00690)
Cloud coverage			-0.625*** (0.190)	-0.926*** (0.151)
Rainfall			0.159*** (0.0252)	0.145*** (0.0266)
Rainfall squared			-0.00596*** (0.00122)	-0.00570*** (0.00132)
Wind speed			0.676*** (0.238)	0.631** (0.288)
Wind speed squared			-0.0178*** (0.00661)	-0.0144** (0.00693)
Constant	2.240***	7.516***	-901.5***	-767.3***
Region FE			YES	
Region-Year FE				YES
Weekday FE				YES
Month FE				YES
Observations	41,651	41,651	36,274	36,274

Notes: This table provides estimates of the effect of the average  $NO_2$  concentration on the number of accidents per grid cell per day. Estimates in Column 1 are from a simple OLS estimator. Estimates in Columns 2-4 are from 2SLS estimators using the occurrence of temperature inversions on a given day as instrumental variable for pollution. Robust standard errors are reported in parentheses (\*\*\*)  $p < 0.01$ , (\*\*)  $p < 0.05$ , (\*)  $p < 0.1$ ).