// NO.22-013 | 05/2022

DISCUSSION PAPER

// YASEMIN KARAMIK AND KATHRINE VON GRAEVENITZ

Gone With the Wind: The Effect of Air Pollution on Crime – Evidence From Germany

Leibniz Association

ZEW

Gone with the Wind: The Effect of Air Pollution on Crime - Evidence from Germany

Yasemin Karamik* Kathrine von Graevenitz †

May 12, 2022

Abstract

Recent evidence suggests a positive impact of air pollution on crime in large cities. We provide first evidence on the potential effect of air pollution on criminal activity using a broader set of geographical regions with lower air pollution levels. We use a unique combination of daily crime data with weather and emission records for the states of Baden-Wuerttemberg (BW) and Rhineland-Palatinate (RLP) in Germany from 2015 until 2017. We exploit the variation in air pollution which is attributable to changes in daily wind direction. We find that an increase of one standard deviation of PM10 leads to an increase in crime of 4.6%.

Keywords: Air Pollution, Crime.

JEL Classification: K42, Q53

*ZEW Mannheim, e-mail: yasemin.karamik@zew.de

†ZEW Mannheim and University of Mannheim, e-mail: kathrine.vongraevenitz@zew.de

1 Introduction

Quantifying the costs of air pollution is a massive undertaking. Air pollution impacts the environment as well as individuals through a plethora of channels. According to the German Federal Environment Agency - *Umweltbundesamt* (UBA) -, the economic cost of air pollution purely related to health amounts to 153 billion euro annually for the case of Germany (Kallweit and Bünger, 2015). This number gives a first indication of the monetary value of the overall costs of air pollution. Recent research has shown that air pollution not only affects areas such as health (Currie et al., 2014) or housing prices (Chay and Greenstone, 2005) but also reduces e.g., productivity (Zivin and Neidell, 2012), labor market participation (Hanna and Oliva, 2015) and student test scores (Ebenstein, Lavy and Roth, 2016). The policy relevance of understanding the impact of air pollution is apparent - both in terms of the scale of the impacts but also the level of air pollution at which detrimental effects occur.

The link between air pollution and crime has also been studied (see e.g., Bondy, Roth and Sager (2020); Herrnstadt et al. (2021)). Channels through which pollution may affect crime are established in the literatures of medicine, biology and psychology. For instance, air pollution can alter hormone levels in human beings as they inhale these pollutants (Li et al., 2017). This, in turn, can lead to changes in brain chemistry which impacts the way individuals behave. Riis-Vestergaard et al. (2018), for example, show that individuals alter their time preferences if exposed to elevated levels of stress hormones like cortisol. Using the rational choice model of Becker (1968), altered time preferences might modify the expected cost of crime for an individual through the channel of the discount rate. This could lead to an increase in the number of premeditated crimes as well as instantaneous crimes. Further, behavior of individuals could be affected through exacerbated morbidity induced by pollution. Discomfort or physical pain can induce aggressive behavior (Anderson and Bushman, 2002). Consequently, air pollution might lead to an increase in violent crimes.

To assess this relationship, we study whether daily variation in air pollution affects crime in the states of Baden-Wuerttemberg (BW) and Rhineland-Palatinate (RLP) in Germany following the proposed approach of Bondy, Roth and Sager (2020). We employ a fixed effect estimation and, additionally, use daily wind direction as an Instrumental Variable (IV) to capture the exogenous variation in PM10. We focus on the effect of PM10 as it is the most consistently measured pollutant for the case of Germany. PM10 belongs to the group of fine particulate matter with an aerodynamic diameter less or equal to 10 micrometer (μm). Fine particulate matter are solid dust particles which do not immediately sink to the ground when set free, but linger in the atmosphere for a certain period of time.[1] Using daily wind direction as an IV, we are able to circumvent concerns of endogeneity of air pollution due to time-varying unobserved correlated factors.

For the analysis, we created a unique data set which combines administrative records of crimes, weather and air pollution at the daily level. The data set allows us to shift the focus of analysis from major cities like London (Bondy, Roth and Sager (2020)) and Chicago (Herrnstadt et al. (2021)) to less metropolitan areas. In the study by Bondy, Roth and Sager (2020) for the 2004-

[1] Definition according to UBA.

2005 period, London had a crime rate of 34 per 100,000 inhabitants at the ward level on average and PM10 levels of 28 $\mu g/m^3$. In comparison, the average geographical region in our sample, has an annual mean PM10 of 15.69 $\mu g/m^3$ and the mean crime rate across our regions is approximately 9 per 100,000 inhabitants. Our data allows us to investigate the relationship between air pollution and crime in a much less urban setting. A challenge is that the coverage of emission and weather stations is more sporadic in comparison to the London or Chicago setting. We employ a buffer zone approach to define the regions of interest where data is available.

Past research has primarily been based on very disaggregated crime data, e.g. daily crime at the ward level in London (Bondy, Roth and Sager (2020)) or in neighborhoods divided by major highways in Chicago (Herrnstadt et al. (2021)) or police administrative areas in Los Angeles or Houston (Herrnstadt et al. (2019)). Sarmiento (2020) takes it a step further and analyses effects of air pollution at the hourly level in Los Angeles, Mexico City, New York and Toronto. Our design is less disaggregated, but not necessarily less accurate. As a result of the aggregation, our research design is less vulnerable to concerns about the exact location (and time) of the crime. Research designs based on small geographic entities within a city may be invalid if the location of the crime is not accurately registered or if exposure occurs in one location but the crime somewhere else. For violent crime discovery and reporting is likely to be fairly accurate, whereas pick-pocketing may be discovered with a time lag. Therefore this concern can be addressed to some extent by comparing effects of air pollution across crime types. With regard to exposure to air pollution it is more difficult to assess the effect of accuracy in the exposure measure. In principle it is possible to assess whether there are spillover effects, e.g. between wards, but in practice air pollution levels are likely to be strongly correlated across space at short distances making such analysis difficult. Both issues of exact reporting of the location of a crime and exposure to air pollution are less likely to be a concern in a more aggregated research design like ours.

Our main finding is that a 10 $\mu g/m^3$ increase in daily PM10 (approximately one standard deviation) leads to an increase of crime by 4.6%. Our main findings are robust to various specifications. In addition, a placebo exercise as well as a weak IV test support the validity of the instruments. Our heterogeneity analysis does not show much variation across crime type in the point estimates, but the regression is underpowered. Only sexual crimes show a significant effect.

The paper is structured as follows. Section 2 describes the data and applied definitions. The estimation strategy is explained in Section 3 and Section 4 presents the results. Section 5 discusses potential biases and caveats as well as limitations concerning the analysis. Section 6 concludes.

2 Data

The final data set is a combination of three administrative data sources that contain information on crime, emissions and weather. For information on weather, the German Weather Agency - *Deutscher Wetterdienst* (DWD)-, provides data on hourly means of temperature, relative humidity, wind direction, wind speed, atmospheric pressure and precipitation. In the empirical analysis, we use daily measures of weather variables whereas wind direction serves as the instrument. To assess the effect of air pollution, daily data can be obtained from files provided by the UBA. These files contain information on hourly means of several pollutants that include carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO2), sulfur dioxide (SO2), total suspended particles (TSP), and particulate matter (PM10 and PM2.5). In the empirical analysis, we focus on PM10 as the coverage of stations measuring this pollutant is most consistent for the case of Germany. As several air pollutants, e.g. NOx, PM2.5 and PM10, tend to correlate strongly, we consider PM10 as a proxy for multiple air pollutants.

In order to asses the causal link between air pollution and crime by using the daily wind direction IV, we collaborated with the State Criminal Investigations Department - *Landeskriminalamt* (CID) of BW and RLP.[2] We created a unique data set on daily crimes for these two states with observations for the years 2003 until 2019 for BW (9,859,183 observations) and 2015 until 2018 for RLP (1,031,346 observations). The data set provides variables on date of crime, crime type description in text format, crime key and location of crime (name of locality and, for the case of RLP, postcode). For cases in BW, the data set additionally includes a municipality key and an estimated time frame of crime expressed in time and date.[3] Details on the harmonization procedure of the two CID data sets and further data preparation can be found in Appendix 9.1.

We restrict the time frame of analysis to 2015 until 2017 to ensure the most consistent and precise measurements for environmental, weather and crime data.[4] With these restrictions, the number of observations reduce to 1,767,547 for BW and 792,039 for RLP.

2.1 Area Definition

The selection of geographic areas based on buffer zones is inspired by Banzhaf and Walsh (2008) and is conducted in the following way: First, stations that consistently measure all mentioned weather variables or PM10 over the time frame of analysis are selected (Appendix Figure A.3 depicts regional distribution of weather stations before (top) and after (bottom) the consistency restriction is applied). Afterwards, stations are assigned to municipalities based on shape files for the year 2019 provided by the Federal Agency for Cartography and Geodesy - *Bundesamt für*

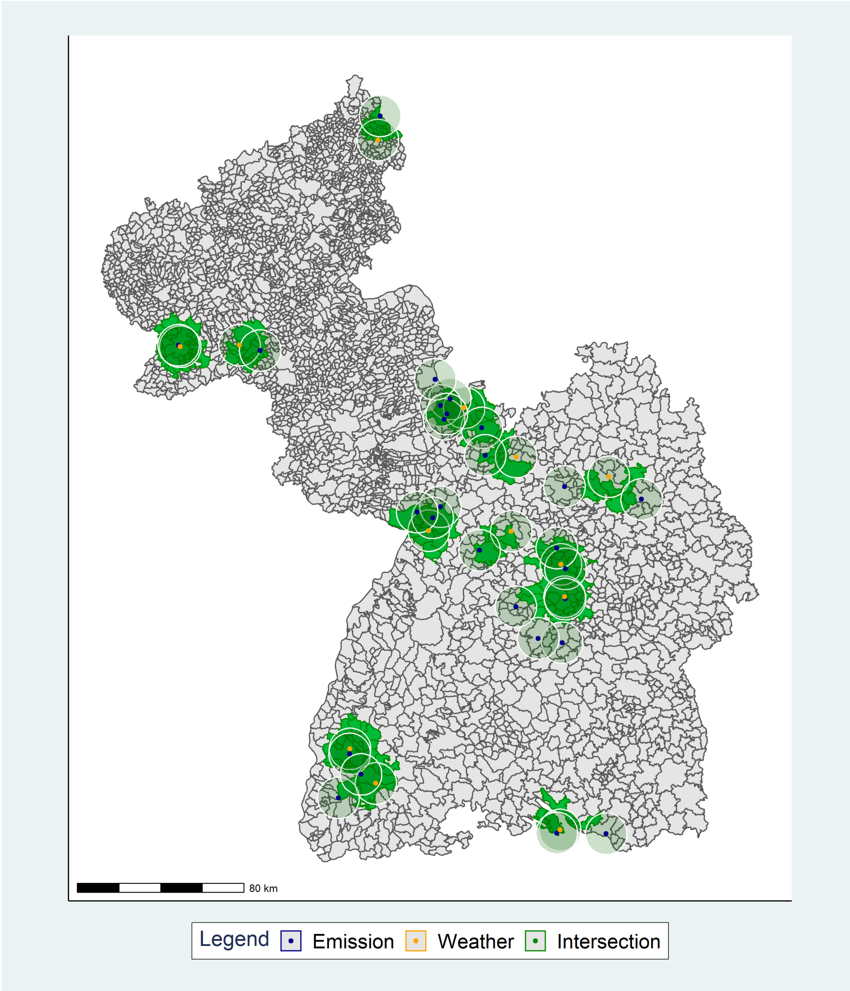
[2] The decision for choosing these states depended on the willingness of the CIDs in Germany to provide data access.

[3] A municipality key is a sequence of numbers assigned to each politically independent city, municipality and unincorporated regions for identification purposes.

[4] The time frame is defined by considering following conditions: First, the number as well as the distribution of environmental stations are more consistent for recent years starting from 2010. In addition, the incidence of stations that consistently report measured days throughout the year is higher. Third, to combine the crime data sets of BW and RLP, the beginning of the time frame is set to 2015. Finally, we have refrained from including the year of 2018 into the data set due to legislative changes on the definition of certain crime types that induced a change at the first level of crime classification (see Appendix 9.1). For an overview on legislative and institutional changes of crime types and their corresponding crime key, please have a look at *PKS 2020 – Straftatenkatalog (4-stellig) - Historie bis 2020* available on the website of Federal Criminal Police Office - *Bundeskriminalamt* (CPO).

Kartographie und Geodäsie (GeoBasis-DE).

Figure 1: Geographic Groups based on 10 km Buffer Zones with Station Locations



Notes: Resulting geographic groups base on 10 km buffer zones (circles) around emission and weather stations in BW and RLP. For the marginal polygons version please have a look on Appendix Figure A.4.
Source: Own figure based on environmental data provided by the DWD and UBA. Shape files base on © GeoBasis-DE / BKG 2019.

We create buffer zones of 10 km around each station.[5] An assignment is successful if the buffer zone of a weather station and an emission station fall within the borders of a municipality. In this way, ten geographical groups emerge. Figure 1 illustrates the buffer zone approach and Figure 2 shows in more detail the resulting geographic groups.[6] For geographical groups with more than one of each station type assigned, the average is taken and used for the whole group.

The days of observation are harmonized within each group by only counting days for which all assigned stations within each group report entries for the same date. Missing hourly entries for weather and emission data for days that have five or less missing entries per day are imputed using a linear regression imputation.[7] All days with more than five missing entries per day are treated as

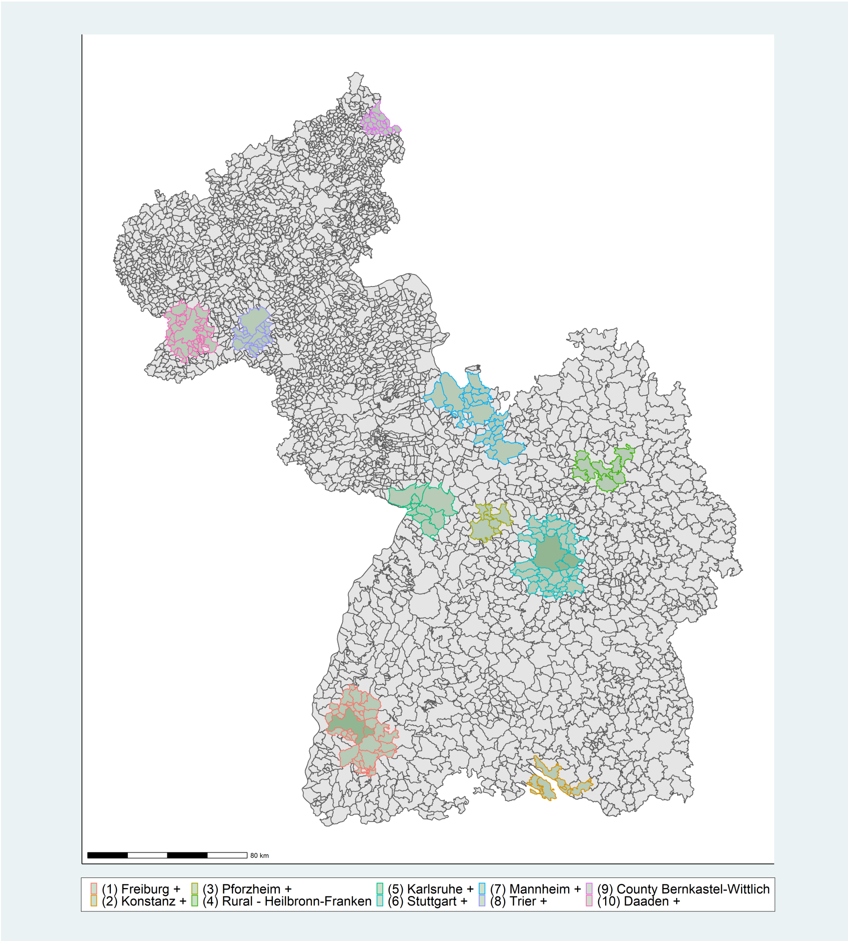
[5]The buffer zone of 10 km is set to ensure the creation of enclosed geographic regions and, in addition, to guarantee that at least one of each station type can be assigned to the emerged geographic groups.

[6]A polygon of a municipality is dropped if its intersection with the buffer zones of each station type is marginal. This is conducted via a visual inspection with the software QGIS. For a detailed version with exact station location and dropped municipalities please have a look at Appendix Figure A.4.

[7]As the number of missing values are mostly reported in consecutive hours, we have decided to restrict the maximum number of missing entries to five. Further, we are able to partly circumvent the concern on the non-randomness of missing entries as most of the missing values are reported at night-time (see Appendix Figure A.6 and Appendix Figure A.7 for heatmaps and frequencies of environmental variables.)

missing observations. In this manner, the influence of measurement errors as well as the influence of strategic placement and monitoring of stations is reduced.[8]

Figure 2: Geographic Groups based on 10 km Buffer Zones



Notes: Resulting geographic groups base on 10 km buffer zones around emission and weather stations in BW and RLP. The plus sign stand for *and outer area*.
Source: Own figure based on environmental data provided by the DWD and UBA. Shape files base on © GeoBasis-DE / BKG 2019.

2.2 Crime Data

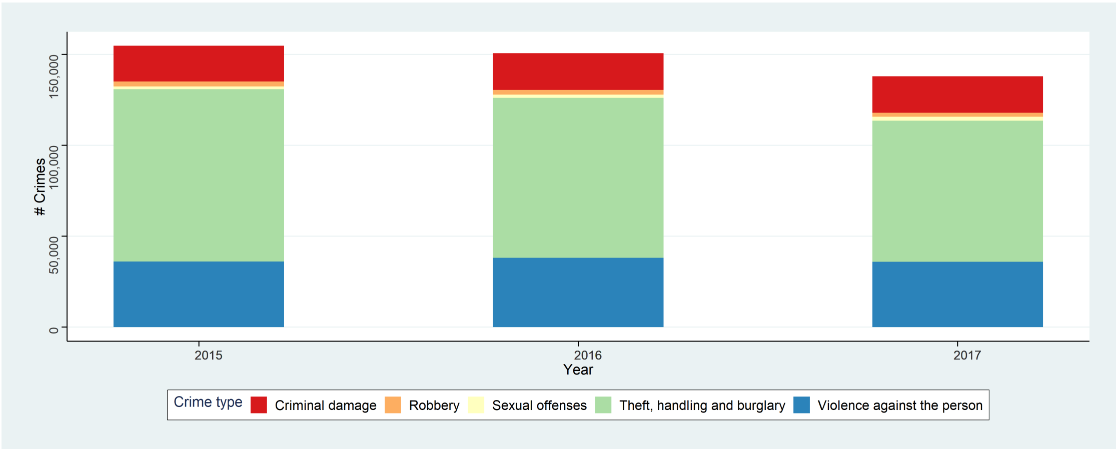
In order to be able to compare the estimates obtained from the German setting to the London setting of Bondy, Roth and Sager (2020), we define overall crime as the sum of six crime groups: (1) Criminal damage, (2) Robbery, (3) Sexual offenses, (4) Violence against the person, (5) Theft and handling and (6) Burglary. This division follows the structure of the crime reports of the London Metropolitan Police Service. As the crime groups (5) Theft and handling and (6) Burglary cannot be clearly distinguished in the German crime reports these subdivisions will be redefined as (5) Theft, handling and burglary in our setting.[9] A detailed description on the German crime reports and the definition of crimes as well crime types can be found in Appendix 9.2 and Appendix Table

[8] Zou (2021) shows that strategic monitoring can and does occur in the US. However, there is no evidence for strategic monitoring for the case of Germany.

[9] The German crime reports follow the structure of the Police Criminal Statistics - *Polizeiliche Kriminalstatistik* (PCS) in which crime is divided broadly into eight categories: (0) Crimes against life, (1) Crimes against sexual self-determination total, (2) Act of brutality and crimes against personal liberty, (3) Theft without aggravating circumstances, (4) Theft with aggravating circumstances, (5) Property and forgery offenses, (6) Other criminal offenses and (7) Criminal ancillary laws.

A.1. After imposing the restriction on crime types as well as the geographic subset, the final data set contains 485,547 offenses (430,934 BW and 54,613 RLP). Figure 3 depicts the overall trend of crimes for the selected geographical groups of both states. The number of all crimes seems to follow a decreasing trend whereby crimes defined as (5) Theft, handling and burglary account for the greatest share of conducted crimes, followed by (4) Violence against the person and (1) Criminal damage. The smallest share is attributed to (2) Robbery and (3) Sexual offenses.

Figure 3: Trend Type of Crimes for the 10 Selected Groups



Notes: Overall trend of crimes for the 10 selected geographical groups in the states of RLP and BW differentiated among crime types as describes in text.
Source: Own graph based on the data set provided by CID of BW and RLP.

2.3 Descriptive Statistics

Table 1 presents the summary statistics for the main estimation sample consisting of 10,243 geographic group-by-day observations. The mean daily concentration of PM10 is 15.69 micrograms per cubic meter ($\mu g/m^3$) with an average population of 374,063 per group, whereas the number of crimes per 100,000 people equals 8.76. Appendix Table A.2 presents the allocation of observed days per group. For the case of London for the years of 2004 until 2005, the mean daily concentration of PM10 per ward is 28.05 $\mu m/m^3$ with an average population of 11,900. The number of crimes per 100,000 people per ward equals 34.06 (numbers according to Bondy, Roth and Sager (2020, p.562)).

These descriptive statistics illustrate the differences between major cities like London compared to the urban areas included in our sample for Germany. The mean daily concentration of PM10 in our sample is approximately half of the concentration measured in London. Moreover, crime is much less common. Summing up, the analysis in this study shifts the focus on less populated and less polluted cities. Appendix Figure A.5 depicts annual PM10 emissions for the selected geographical groups which follow a slight decreasing trend.

The overall distribution of the total number of crimes and PM10 are plotted in Figure 4. Two important features of the dependent variable have to be considered when proceeding with the estimation. First, the number of crimes is a count variable including zeros as outcomes and, second, it roughly follows a Poisson distribution which corresponds to the findings of Bondy, Roth and Sager (2020). Hence, the second stage of the IV approach is adjusted by employing a

Table 1: Descriptive Statistics

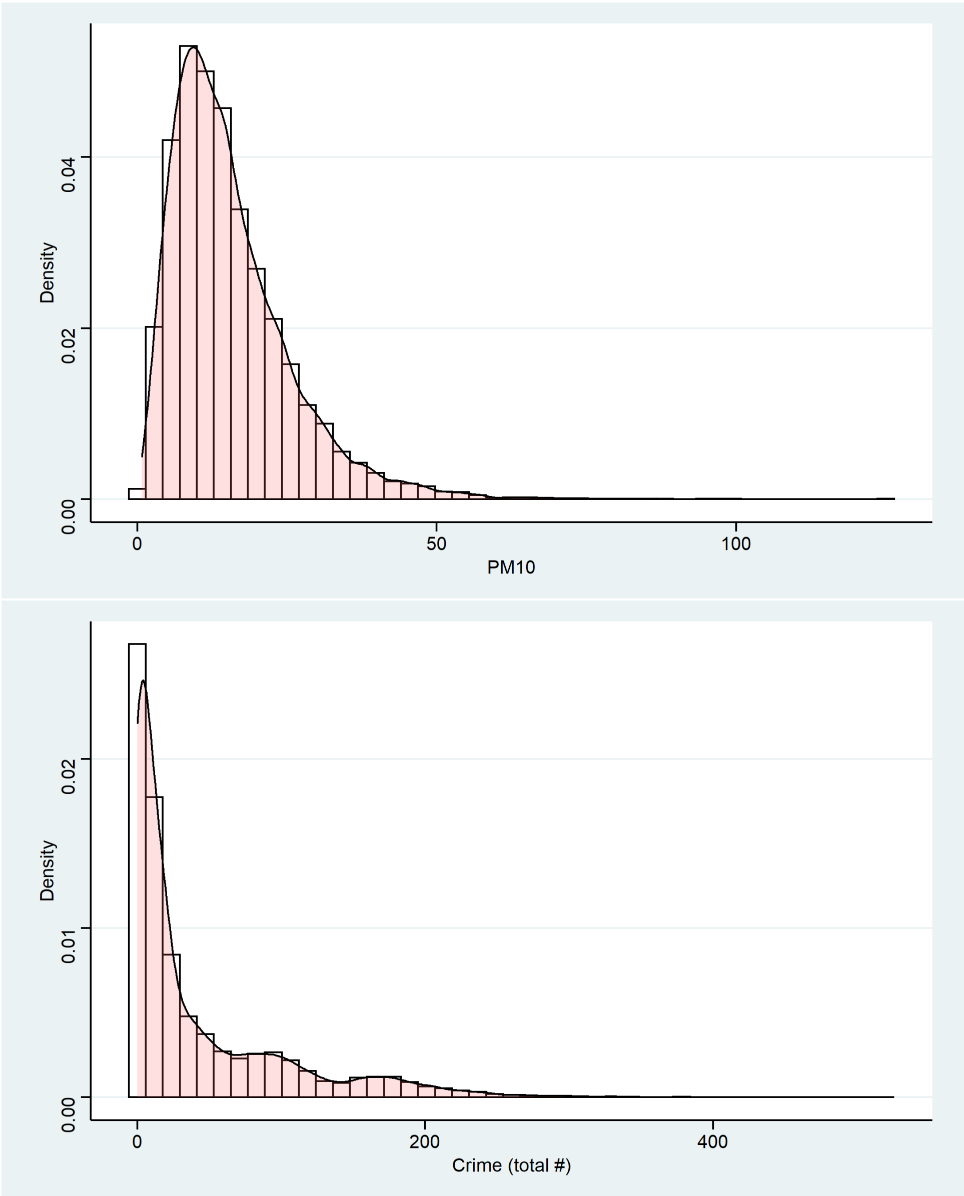
Variable	Mean	Std. Dev.	Median	Min.	Max.	N
Crime (no. per 100K population)	8.76	5.61	8.71	0.00	50.89	10,243
Wind Speed (m/s)	3.16	1.80	2.69	0.12	15.94	10,243
PM10 (mg/m3)	15.69	10.56	13.38	0.71	125.70	10,243
Relative Humidity (%)	77.80	12.27	79.08	35.62	100.00	10,243
Temperature (Celsius)	10.13	7.31	9.79	-11.36	30.15	10,243
Rainfall (mm)	2.21	4.92	0.00	0.00	76.20	10,243
Crime (log)	2.90	1.62	3.00	0.00	6.25	9,187
Population Density (per km2)	608.71	442.31	649.74	60.04	1,559.09	10,243
Population Total	374,063.52	443,338.63	178,693.00	19,060.00	1,495,276.00	10,243

Note: Descriptive statistics base on main estimation sample including days with imputed values if a maximum of five entries per day were missing as well as New Year‘s Eve and the day after. The observation unit is geographic group per day.
Source: Crime Data from CID BW and RLP. Environmental variables from UBA and BW. .

Poisson Pseudo-Maximum Likelihood (PPML) estimation for which the offset variable is defined as population per group.[10] Section 3 describes the estimation procedure in more detail.

[10]The PPML is beneficial in this setting as it does not constrain the conditional variance to be equal to the mean as the standard Poisson estimation imposes.

Figure 4: Density Plots for PM10 and Crime



Notes: Density plots for PM10 (top) and total number of crimes (bottom) by day by geographical group for the main estimation sample. Appendix Figure A.10 depict density plots differentiated among crime types.
Source: Own graph based on the crime data provided by CID of BW and RLP and emission data by UBA.

3 Estimation Strategy

Our estimation strategy is similar to the strategy used in Bondy, Roth and Sager (2020). We briefly summarize, first, the reasons for using an IV approach and second, the rationale behind instrumenting air pollution by wind direction.

There are several challenges that prevent the unbiased identification of the causal relationship between air pollution and crime. For instance, the number of crimes can be driven by unobserved correlated factors such as neighborhood characteristics and the presence of police stations. If such factors are not taken into account, the obtained estimates will be biased. Fixed effects can account for unobserved time-invariant confounders. However, unobserved time-varying confounders remain a concern. For example, high levels of activity in a city may increase air pollution through traffic and at the same time provide more ample opportunity for crime. We use the IV-method to capture exogenous variation in local air quality. For the instrument we follow the recent literature based on Deryugina et al. 2019 and use variation in wind direction. This is based on two identifying assumptions: (1) The relevance of the instrument -, i.e., *wind direction* must be correlated with the endogenous variable air pollution, and (2) the exclusion restriction -, i.e., wind direction should not be correlated with the error term. In other words, wind direction should affect the outcome of interest only through the variation it induces in air pollution.

With regard to relevance of the instruments, the first stage F-test rejects the null hypothesis that the coefficients on the instruments are jointly zero (See Appendix Section 10 Table A.8). For further evidence, Appendix Figure A.8 depicts the relationship between daily PM10 concentrations and average wind direction for each geographical group. A further requirement, the monotonicity condition, applies in order for us to interpret the results as local average treatment effects (LATE) (Athey and Imbens, 2017). In other words, the effect of wind direction on air pollution should affect all geographic regions in the same way. In our setting, the condition would be violated if a particular wind direction is transporting high pollution levels into one geographic area, but, at the same time, systematically blows away air pollution from another geographic group. We address this concern by allowing the effect of wind direction to differ between regions. To assess whether the condition is violated within geographic regions, we vary the bin size for our wind direction from a 90 degree interval to 60 (See Section 3.1 for the definition of wind direction bin sizes). If the variation changes our results we should be concerned about violations of the monotonicity assumption. Section 4.3 discusses the test results for monotoniticy and weak IV.

The exclusion restriction, in contrast, cannot be tested statistically. To the best of our knowledge, there is no evidence supporting the idea of a direct effect of wind direction on criminal activity.[11] Alternative instruments based on other weather variables that affect the variation in air pollution like temperature or rainfall fail in terms of the exclusion restriction (see, e.g., Ranson 2014).

[11] Except for sailing pirates for which the wind direction might play a role. However, this is not a concern in the German context.

3.1 Model Specification

The focus of our analysis lies on the causal effect of air pollution captured by PM10 on daily crimes. We define the estimation equation as follows:

$$\begin{aligned} Crime_{at} = & exp\,\{\,\beta PM10_{at} + f(Temp_{at}, RH_{at}) + \tau Wind_{at} + \\ & \omega Rain_{at} + \gamma_a + \boldsymbol{\mu_t}\,\} + \varepsilon_{at} \end{aligned} \tag{1}$$

Here the dependent variable is the number of crimes in area a on day t. β is the parameter of interest, capturing the effect of air pollution ($PM10_{at}$). For interpretation purposes, $PM10_{at}$ is divided by 10. This changes the effect size of an increase in crime per PM10 $\mu g/m^3$ to per 10 $\mu g/m^3$ and corresponds approximately to one standard deviation. The estimates for Equation (1) are obtained by a PPML estimation using group population as the offset unless otherwise noted. We include a vector of weather controls using wind speed and precipitation and a flexible function of temperature and relative humidity (RH_{at}) to account for weather conditions that could bias β.[12] For the main analysis, the weather controls are defined as in Bondy, Roth and Sager (2020). A robustness check is conducted in Section 4.3 by varying the functional form. The estimation equation includes two kinds of fixed effects (FE): (1) area FE (γ_a) to account for time invariant geographic differences in crime and pollution and (2) time FE ($\boldsymbol{\mu_t}$). With regard to time, we control for day-of-the-week and month-year using fixed effects. With these controls we capture a variety of time-invariant confounders and general time trends. For instance, weekday FE captures the differences between busy weekdays from quiet ones in terms of pollution levels and criminal activity. This is especially important as the *Sunday and National Holiday Observance Act (Feiertagsgesetz FTG)* applies in Germany. Therefore, Sundays and national holidays are subject to special protection that prohibits "publicly noticeable work on these days" (Service-bw, n.d.).

However, the presence of unobserved time-varying factors related to pollution and crime remains a concern. Therefore, we adopt the proposed IV approach by Bondy, Roth and Sager (2020). We rely on changes in daily wind direction in the defined area, that serve as an exogenous shock to local pollution levels. Formally, we define the first stage as follows:

$$\begin{aligned} PM10_{at} = & \boldsymbol{\rho_a}\mathbf{WindDir_{at}} + \delta f(Temp_{at}, RH_{at}) + \phi Wind_{at} + \varphi Rain_{at} \\ & + \theta a + \boldsymbol{\alpha_t} + \upsilon_{at} \end{aligned} \tag{2}$$

Here the instrument, $\mathbf{WindDir_{at}}$, is a binary indicator for the average wind direction in region a belonging to one of the four wind directions. The four wind directions are defined by dividing the 360 degrees into 90 degree bins, e.g., North-East is equivalent to [0° till 90°), etc..[13] The base category of the instruments is wind blowing from the South-West, as this is the most common wind direction in Germany.[14] Appendix Figure A.8 illustrates the wind direction behavior graphically.

[12]More precisely, we include a dummy variable indicating five equally sized temperature bins ($Tempbin$), relative humidity (RH), an interaction between temperature bins and relative humidity, RH^2 and ($Temp^2 * RH^2$).

[13]The DWD reports wind direction as direction the wind is blowing from, in clockwise degrees starting off by 0° representing north.

[14]A similar definition was successfully employed by Isphording and Pestel (2021).

Our research design accounts for the scarcity of emission stations in order to ensure a precise measurement of air pollution exposure. Further, we are able to divide the area of analysis into spatial groups in order to allow for geographic differences concerning wind patterns. Specifically, we allow $\boldsymbol{\rho_a}$ to differ between regions (see, e.g., Deryugina et al. 2019). The remaining variables, controls and fixed effects are defined as in Equation (1).

We use the control function approach to account for endogeneity of air pollution as proposed by Wooldridge (2014). The estimation equation in the second stage changes to:

$$\begin{aligned} Crime_{at} =& exp\{\,\beta PM10_{at} + \alpha PM10\widehat{\_residual}_{at} + f(Temp_{at}, RH_{at}) + \tau Wind_{at} + \\ & \omega Rain_{at} + \gamma_a + \boldsymbol{\mu_t}\,\} + \varepsilon_{at} \end{aligned} \tag{3}$$

The second stage PPML estimation includes the actual pollution variable depicted by $PM10_{at}$ and the residual, $PM10\widehat{\_residual}_{at}$, from the linear (OLS) first stage which is estimated via Equation (2). The estimation procedure is calculated manually in Stata. The standard errors therefore have to be adjusted to account for the uncertainty in the first stage.[15] We bootstrap the standard errors using the group variable as a resampling cluster with 1,000 replications for both stages.[16]

[15]For the first stage the Stata package *reghdfe* is used (Correia, 2014). The second stage is calculated with the Stata package *ppmlhdfe* (Correia, Guimarães and Zylkin, 2019).

[16]Preferably, standard errors should be clustered at the group and year-month level. However, the number of groups is too low for asymptotic theory to apply (e.g. Cameron and Miller (2015) suggest at least 40 to 50 clusters are necessary). For comparison to the clustered bootstrap, we also calculated heteroscedasticity-consistent standard errors (Huber/White/sandwich estimators) using *vce(robust)*. Results are available from the authors upon request.

4 Results

4.1 Main Results

Table 2 reports the main results of the regressions of the daily number of crimes on PM10. We report results from a pooled PPML (column (1)) followed by fixed effect models (columns (2) to (4)) and finally the results using our instrumental variables with a control function approach (column (5)) and manual instrumentation (column (6)). Standard errors are bootstrapped using the group variable as a resampling cluster.[17] The first estimate including only weather controls, but no fixed effects suggests a positive relationship between PM10 and crime. However, the coefficient estimate is insignificant. Adding fixed effects as we move from left (column (2)) to right (column (4)) the coefficient becomes larger and significant at the 1% level ranging from a 2.8% to 3.6% increase in crime per 10 $\mu g/m^3$ increase in daily PM10. Column (5) is our preferred specification using a control function approach with changes in the wind direction as instruments. We find an increase of crime by 4.6% per 10 $\mu g/m^3$ increase in daily PM10. With manual instrumentation for PM10 based on the same instruments we get a very similar point estimate of 4.3% per 10 $\mu g/m^3$ increase in PM10.

Table 2: Main Results

	(1) All Crimes	(2) All Crimes	(3) All Crimes	(4) All Crimes	(5) All Crimes	(6) All Crimes
PM10	0.029	0.030***	0.028**	0.036***	0.046***	0.043***
	(0.0222)	(0.0088)	(0.0095)	(0.0095)	(0.0104)	(0.0115)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	No	Yes	Yes	Yes	Yes	Yes
DOW FE	No	No	Yes	Yes	Yes	Yes
Year-Month FE	No	No	No	Yes	Yes	Yes
N	10243	10243	10243	10243	10243	10243

Notes: Throughout column (1) until (4) a PPML is estimated with group population as the offset. Column (6) is estimated via manual instrumentation for which the second stage is run with PPML. The full regression table for these columns can be found in Appendix Table A.5. In column (5) the IV estimation is run using the control approach in which the second stage is run via PPML. The full regression table for this column can be found in Appendix Table A.6. Appendix Table A.7 shows the results for the manual instrumentation with boostrapped standard errors. First stage results are reported in Appendix Table A.8. Bootstrapped standard errors are reported in parantheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2 Heterogeneity Analysis

We conduct two types of heterogeneity analyses to check whether the effect of air pollution on crime is non-linear and whether the estimate is driven by any particular type of crime. To examine the potential nonlinear relationship of air pollution and crime, we substitute the continuous measure of PM10 with dummy variables in the fixed effects estimation (column (4) in Table 2). We do not use variation in wind direction in this analysis, as the IV is weaker once we convert the continuous measure of PM10 into discrete dummies. The results reveal, as shown in Table 3, that the positive effect of air pollution on crime can be found even with PM10 concentration levels well below the current EU regulatory standards (air quality directive (2008/EC/50)) which sets a limit of 50 $\mu g/m^3$ per day.

[17] We also calculated robust standard errors, which are slightly smaller than their bootstrapped counterparts. These results are available from the authors upon request.

Table 3: Nonlinear Model

	All Crimes
PM10_>=15_<30	0.030*
	(0.0129)
PM10_>=30_<45	0.046**
	(0.0173)
PM10_>=45	0.138*
	(0.0543)
Controls	Yes
Group FE	Yes
DOW FE	Yes
Year-Month FE	Yes
N	10243

Notes: Nonlinear estimates obtained via fixed effect PPML estimation and an indicator for each PM10 level. Bootstrapped standard errors are reported in parantheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

We differentiate the number of crimes into its five components: Theft, handling and burglary (*Theft*), Violence against the person (*Violence*), Criminal damage, Sexual offenses (*Sexual*), and Robbery. We estimate the effects on each type using the control function approach with bootstrapped standard errors and the corresponding crime type as the dependent variable. Figure 5 displays the results for each crime type, The point estimates are very similar to the main estimate, but imprecisely estimated. For four out of five crime types, the effect is insignificant, with only sexual crimes having a positive and statistically significant effect. Appendix Table A.2 lists the number of days a crime occurred differentiated per year per group. *Sexual* and *Robbery* have the highest number of zero occurrences (lowest number of observed days). *Theft* occurs most often, followed by *Violence* and *Criminal Damage*. This pattern is partly caused by our restrictive definition of crime and the corresponding crime types.[18]

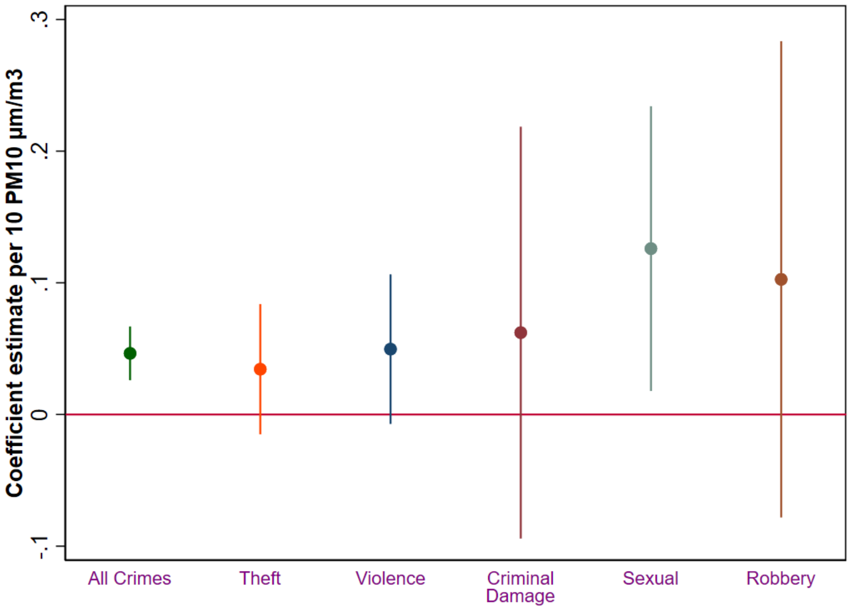
4.3 Robustness Checks

One potential concern that may persist is the occurrence of special 'incidents' that drive both variables, air pollution and crime. For instance, New Year's Eve impacts the overall level of air pollution due to the heavy usage of fireworks.[19] Additionally, the event itself might affect the number of crimes through raised opportunity of, e.g., pickpocketing due to crowdedness. Therefore, we test whether the results hold if the day of New Year's Eve, as well as the first of January, are excluded from the data set. Column (1) of Table 4 reports the result. The size and significance of the coefficient remain similar compared to the preferred specification in column (5) in Table 2. Further, we test the robustness of the results by varying the functional form of weather variables (column (3)) and using the restrictive data set instead of the imputed one (column (5)). In both cases, the estimates remain significant, though the magnitudes differ slightly. Next, we estimate an alternative model specification with a linear estimation model using $ln(y+1)$ as the dependent variable. This variation is able to circumvent the issue on zero outcomes of the count variable and

[18] As mentioned in Appendix Section 9.1, for instance, we exclude several crime keys within sexual crimes, though the number of sexual crimes is comparatively low also in the raw data.

[19] Appendix Figure A.11 illustrates the daily variation of PM10 for which especially the day after New Year's Eve report high PM10 concentrations.

Figure 5: Heterogeneity Analysis



Notes: Coefficients based on estimates using bootstrapped standard errors with Control Function (CF). Each estimate is obtained by running the estimation with the depicted crime type as the dependent variable. Crimes types *Theft* and *Violenge* stand for *Theft, handling and burglary* and *Violence against the person*, respectively. Estimates for bootstrapped standard errors with manual IV are illustrated in Appendix Figure A.9.
Source: Own graph. Data as described in text.

can be interpreted in the same way as the log transformation of the count if the data contains only a few zeros. Again, the results remain positive and significant, though smaller (column (2)) indicating that the handling of zeros matters. Moreover, the coefficient is not directly comparable to our main specification as the dependent variable is not adjusted for population size. Finally, column (4) reports results based on wind instruments within a 60 degree interval (instead of 90) to assess the validity of the monotonicity assumption (see Section 3).

Table 4: Robustness Checks

	(1) NY Excluded	(2) l(y+1) 2SLS	(3) Alt. Weather-C	(4) Wind Bin 60	(5) Restrictive Data	(6) Placebo IV
Control Function:						
PM10	0.046***	0.042*	0.050***	0.051***	0.034***	-0.017
	(0.0108)	(0.0163)	(0.0106)	(0.0075)	(0.0089)	(0.0447)
	(1)	(2)	(3)	(4)	(5)	(6)
Manual IV:						
PM10	0.043***	0.042*	0.045***	0.047***	0.031**	-0.017
	(0.0111)	(0.0163)	(0.0113)	(0.0090)	(0.0101)	(0.0439)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes	Yes
DOW FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes
N	10186	10243	10243	10243	8788	10243

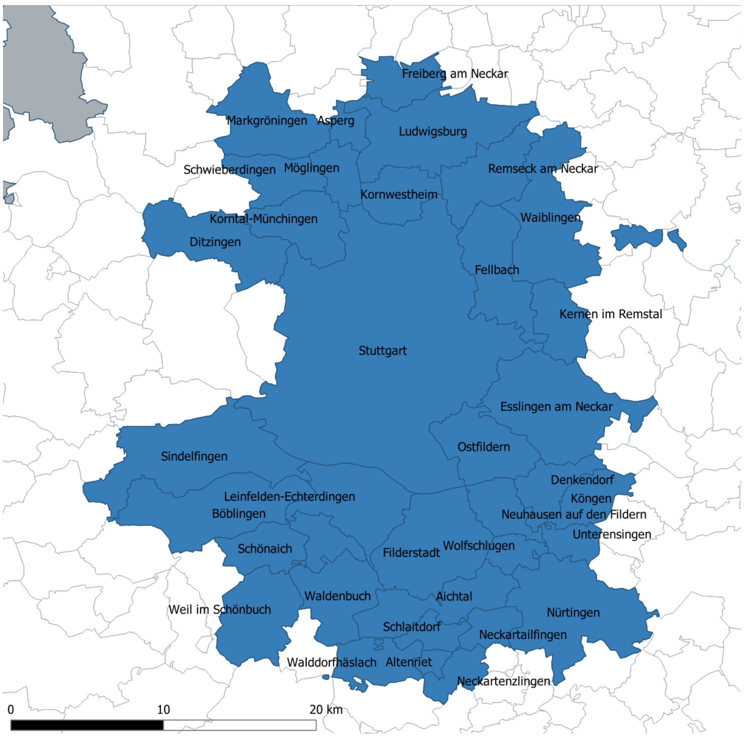
Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) and column(3) until (6) a PPML is estimated with group population as the offset. In column (2), the model is estimated with 2SLS using log crime plus 1 as dependent variable. The full regression table can be found in Appendix Table A.9. The results for the manual IV with bootstrapped standard errors is shown in Appendix Table A.10.First stage results of the placebo estimation (column (6)) are reported in Appendix Table A.11.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, we carry out a placebo test to assess validity of the instrument. In this setup, the IV procedure is calculated by using randomly generated instruments instead of the true instruments. If the estimation with randomly generated instruments leads to similar results as the ones obtained with the true instruments, it raises doubt about the validity of the instruments. The results reported in column (6) are all insignificant.

5 Discussion

The accuracy of the reporting of crimes is a major concern for an analysis such as ours. There are at least four possible ways in which inaccurate reporting might affect the estimation. First, time and location of the reported crime might differ from the actual time and location of the crime. For instance, the offender might have stolen a cell phone in city B, however, the victim noticed this in city A. Second, there is the possibility of spillover effects between our regions. For instance, the offender might be affected by air pollution in city A but conducts the crime in city B. Our geographic regions are larger than e.g. wards in London or neighborhoods in Chicago, making this less of a concern in our setting than in past research. In Figure 6, the geographic region *Stuttgart +* is plotted in more detail. The region not only includes the municipality of Stuttgart, that dominates the number of crimes in this region, but also surrounding municipalities.

Figure 6: Group 6 - Stuttgart and Surrounding Municipalities



Notes: Areas shaded in blue depict municipalties belonging to group *Stuttgart +*. Lines depict overall municipality borders. Names of municipalities are printed within the corresponding polygon.
Source: Own graph using QGIS 3.18 Zürich software and shape files of © GeoBasis-DE.

Third, the presence of legislative changes and/or changes in the definition of crimes might affect the number of reported offences. The most recent comprehensive legislative change affected the definition of sexual offences introduced in 2016 in response to the New Year's incident 2015 in Cologne (Bosen, 2020). This change extended the definition of sexual offences based on the "no-means-no"-model and, additionally, added sexual offences conducted by groups. As depicted in Figure 3, the number of sexual offences as well as the remaining crime types seems to be quite stable over time, suggesting that the impact of these changes has been minor.

Finally, crimes might be under-reported, e.g., if the expected payoff of reporting the crime is negative. Most German insurance policies require the customer to directly report a crime upon becoming aware of it. As long as under-reporting is not systematic, the estimated effect of air pollution on crime will still reflect a lower bound on the true effect. To the best of our knowledge, there is no evidence that reporting itself is correlated with air pollution, which would give greater cause for concern.

Other factors like unemployment, the exact number of police officers deployed in municipalities (see e.g. Blesse and Diegmann (2022)) and crowdedness could affect criminal activity.[20] By including various time dummies as explained in Section 3, we can at least partly account for these effects.

Avoidance behavior due to pollution may affect opportunities for crime. For instance, Graff Zivin and Neidell (2009) show that individuals tend to adopt avoidance behavior concerning physical activities on polluted days. Consequently, criminal opportunities for crime types as pickpocketing might decrease if individuals are aware of current pollution levels in their close environment and respond accordingly. We focus on the short-term (daily) fluctuations of air pollution based on wind direction patterns, whereas Graff Zivin and Neidell (2009) examine behavior changes in days after a smog alert occurred. Avoidance behavior is less likely to play an important role in our setting, though we cannot rule it out. We note that by reducing the opportunity for crime, avoidance behavior would imply that our estimate is a lower bound on the true effect of air pollution on crime.

[20]For example, Bondy, Roth and Sager (2020) use *tube activity* as a proxy for crowdedness.

6 Conclusion

We analyse the impact of air pollution on crime in Germany by exploiting daily variation in PM10. Our analysis is based on three administrative data sets on weather, emissions and daily crimes. We find a postive and significant effect of air pollution on crime both in our fixed effects model and our preferred specification using instrumental variables based on wind direction. Our preferred estimate implies that an increase of PM10 $\mu m/m^3$ (approximately one standard deviation) leads to an increase in overall crime of 4.6%. This result is robust to an extensive set of tests and specifications. The effect does not seem to be driven by a specific type of crime. Although there is an indication that effects may be larger for sexual crimes and robbery, these are imprecisely estimated. We find some evidence of nonlinear effects, as the impact on crime increases with the level of PM10, though it remains statistically significant even at very low levels of pollution.

Our results are obtained using data for regions in which the average PM10 concentration level is generally well below the current EU regulatory standards (air quality directive (2008/EC/50)) of 50 $\mu m/m^3$ per day. Our findings suggest, that there may be considerable benefits even beyond health effects to lowering air pollution limits further.

7 Acknowledgements

We thank Sebastian Blesse for helpful discussions and Sebastian Siegloch for his support on the project. Sefi Roth generously provided additional information on Bondy, Roth and Sager (2020) in email correspondence. We also thank the State Criminal Investigations Departments of Baden-Wuerttemberg and Rhineland-Palatinate for kindly providing data on crimes. Daniel Nietzer provided excellent assistance in cleaning the weather data and the pollution data provided by the German Federal Environment Agency. Seminar audiences at the ZEW provided great feedback. All remaining errors are our own.

8 Bibliography

Anderson, Craig A, and Brad J Bushman. 2002. "Human Aggression." *Annual Review of Psychology*, 53(1): 27–51.

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives*, 31(2): 3–32.

Banzhaf, H. Spencer, and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, 98(3): 843–863.

Becker, Gary S. 1968. "Crime and Punishment : An Economic Approach Author." *Journal of Political Economy*, 76(2): 169–217.

Blesse, Sebastian, and André Diegmann. 2022. "The place-based effects of police stations on crime: Evidence from station closures." *Journal of Public Economics*, 207: 104605.

Bondy, Malvina, Sefi Roth, and Lutz Sager. 2020. "Crime Is in the Air: The Contemporaneous Relationship between Air Pollution and Crime." *Journal of the Association of Environmental and Resource Economists*, 7(3): 555–585.

Bosen, Ralf. 2020. "New Year´s Eve n Cologne: 5 years after the mass assaults." *https://www.dw.com/en/new-years-eve-in-cologne-5-years-after-the-mass-assaults/a-56073007, Date Accessed 2021-04-29.*

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, 50(2): 317–372.

Chay, Kenneth Y., and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2): 376–424.

Correia, Sergio. 2014. "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects." *Statistical Software Components, Boston College Department of Economics.*

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2019. "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects."

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics*, 6(1): 217–247.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review*, 109(12): 4178–4219.

Draca, Mirko, Stephen Machin, and Robert Witt. 2011. "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks." *American Economic Review*, 101(5): 2157–2181.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution." *American Economic Journal: Applied Economics*, 8(4): 36–65.

Graff Zivin, Joshua, and Matthew Neidell. 2009. "Days of haze: Environmental information disclosure and intertemporal avoidance behavior." *Journal of Environmental Economics and Management*, 58(2): 119–128.

Hanna, Rema, and Paulina Oliva. 2015. "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City." *Journal of Public Economics*, 122: 68–79.

Herrnstadt, Evan, Anthony Heyes, Erich Muehlegger, and Soodeh Saberian. 2021. "Air Pollution and Criminal Activity: Microgeographic Evidence from Chicago." *American Economic Journal: Applied Economics*, 13(4): 70–100.

Herrnstadt, Evan, Erich Muehlegger, Anthony Heyes, and Soodeh Saberian. 2019. "Air Pollution as a Cause of Violent Crime." *Unpublished Manusscript*, 1–43.

Isphording, Ingo E., and Nico Pestel. 2021. "Pandemic meets pollution: Poor air quality increases deaths by COVID-19." *Journal of Environmental Economics and Management*, 108(102448).

Kallweit, Dagmar, and Björn Bünger. 2015. "Feinstaub macht krank und kostet Leben – Berechnung jährlich entstehender Kosten durch die Feinstaubbelastung in Deutschland." *UMID: Umwelt und Mensch–Informationsdienst*, 2: 69–72.

Li, Huichu, Jing Cai, Renjie Chen, Zhuohui Zhao, Zhekang Ying, Lin Wang, Jianmin Chen, Ke Hao, Patrick L. Kinney, Honglei Chen, and Haidong Kan. 2017. "Particulate matter exposure and stress hormone levels: A randomized, double-blind, crossover trial of air purification." *Circulation*, 136(7): 618–627.

Ranson, Matthew. 2014. "Crime, weather, and climate change." *Journal of Environmental Economics and Management*, 67(3): 274–302.

Riis-Vestergaard, Michala Iben, Vanessa van Ast, Sandra Cornelisse, Marian Joëls, and Johannes Haushofer. 2018. "The effect of hydrocortisone administration on intertemporal choice." *Psychoneuroendocrinology*, 88: 173–182.

Sarmiento, Luis. 2020. "I Am Innocent: Hourly Variations in Air Pollution and Crime Behavior." *DIW Discussion Papers*, No. 1879.

Service-bw. n.d.. "How to apply for an exemption from the Sunday and National Holiday Observance Act." *https://www.service-bw.de/en_US/web/guest/leistung/-/sbw/How+to+apply+for+an+exemption+from+the+Sunday+and+National+Holiday+Observance+Act-522-leistung-0, Date Accessed 2021-04-12.*

Wooldridge, Jeffrey M. 2014. "Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables." *Journal of Econometrics*, 182(1): 226–234.

Zivin, Joshua Graff, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." *American Economic Review*, 102(7): 3652–3673.

Zou, Eric Yongchen. 2021. "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality." *American Economic Review*, 111(7): 2101–2126.

9 Appendix A: Further Details on Data Preparation

9.1 Harmonization of CID data sets

In the empirical analysis, the following variables are harmonized for the final crime data set: (1) municipality key, (2) definition of crime and (3) definition of date of crime. Before starting with the harmonization of the municipality key, the crime data set of RLP is adjusted in order that the locality of crime could be stated via municipality key. To be able to assign the locality of crime to a municipality key, we have matched the missing municipality key using postcode and name of the municipality with the extensive data base on municipalities in Germany provided by the German Federal Office of Statistics (Statistisches Bundesamt). In this process, the number of observations are reduced by 29 due to non uniqueness in municipality key matching.[21]

The need for the harmonization of the municipality key bases on the fact that municipality keys as well as borders are subject to changes in Germany, even though the incidence for the states of BW and RLP are marginal. This way, the municipality keys are recoded to match municipality borders that base on the year of 2019. For instance, if municipality A in year 2015 is merged to municipality B in year 2016 and not changed afterwards, the municipality key for municipality A in 2015 is recoded to match municipality key of municipality B in 2019. During this procedure, the number of observations for RLP remain, whereby for BW 86 observations are dropped due to non-uniquess of the matching procedure. Concerning the definition of crime types and the date of crime we follow the structure of PCS whereby further restrictions are put on the set up of these two variables following a conservative approach. Chapter 9.2 discusses this in more detail.

9.2 Definition of Crime

The definition of crimes and crime keys follows the structure of the PCS. The PCS is conducted anually by the CPO based on data provided by the 16 CIDs in Germany.[22] The PCS accounts for offenses that are known to the police excluding crimes against the state, traffic offenses, administrative offenses, offenses that do not fall within the jurisdiction of the police (e.g., tax offenses) and offenses that are directly reported to the public prosecutor's office (exceptions apply). The remaining offenses are assigned to unique crime keys (*Erfassungsschlüssel*) which can be summed up to broader crime groups depending on the sequence of numbers in the corresponding crime key. For example, *Murder related to sexual offenses* with crime key 012000 belongs to the group of *Crimes against life* found in key 000000. These broader keys are called *Summenschlüssel*. In this way, the PCS divides crimes in eight broader groups of crimes depending on their first number in the key sequence: (0) Crimes against life, (1) Crimes against sexual self-determination total, (2) Act of brutality and crimes against personal liberty, (3) Theft without aggravating circumstances, (4) Theft with aggravating circumstances, (5) Property and forgery offenses, (6) Other criminal offenses and (7) Criminal ancillary laws.

At this point it is important to consider the potential caveat of legislative changes that might

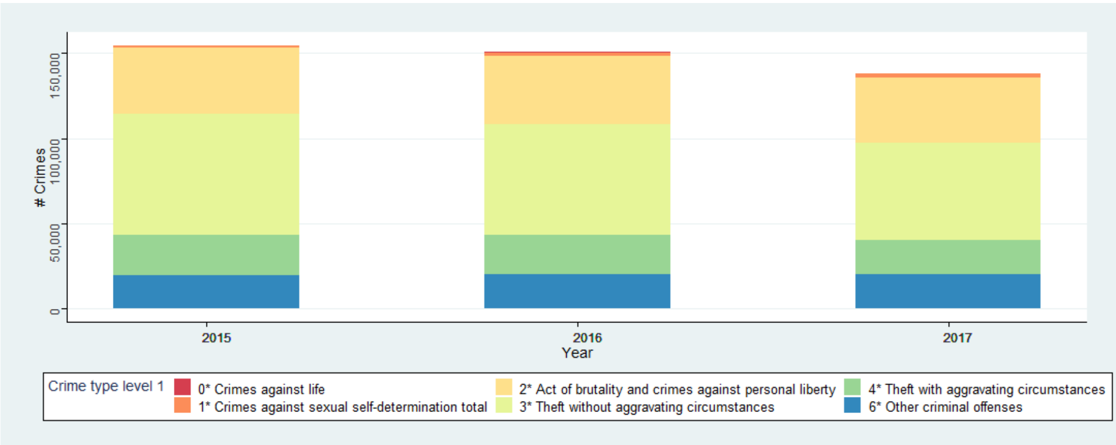
[21]See Table A.1 for an overview the selection process induced losses in observation numbers.

[22]For a detailed description on the role and purpose of the PCS and BKA please have a look at https://www.bka.de/DE/Home/home_node.html.

drive the number of crimes and/or how the conducted crimes are assigned to their corresponding crime key. For the time frame of analysis, the definition of sexual offences are adapted as a response to the New Year's incident of 2015 in Cologne (Bosen, 2020) (see Section 5 for a discussion on this topic). However, most of these adaptations took place on the second level and lower (e.g., change of key sequence starting from second entry). As we differentiate crime types based on the first number in the key sequence, the definition of crime and crime types remains constant. However, the legislative changes might lead to a change in the number of crimes reported in that category.

As shown in Figure A.1, which depicts the overall trend of crimes for the selected geographical groups in the states of RLP and BW, none of the crime types show large spikes. Therefore, the issue of the potential impact of legislative changes on the number of recorded crimes over the analyzed time period is discarded. Furthermore, one can see that the trend concerning the number of crimes is slightly decreasing, which is in line with the overall trend of crimes in the whole state of RLP and BW (Appendix Figure A.2).

Figure A.1: Trend Type of Crimes for the 10 Selected Groups



Notes: Overall trend of crimes for the 10 selected geographical groups in the states of RLP and BW differentiated among first level crime type. First level crime type is defined by the first number in the crime key sequence according to the PCS. *Source:* Own graph based on the data set provided by CID of BW and RLP.

In addition, the definition of crime and the corresponding crime types are subject to further restrictions. First, crime keys for abortion, human trafficking, prostitution and pornography are excluded. All of these crimes, except abortion, which belongs to crimes against life, belong to the category of sexual crimes. We exclude these crimes because the reported day in which the crime is potentially conducted incorporates high measurement errors.[23] Appendix Table A.4 depicts the definition of crime and crime types and the corresponding German crime key in more detail. Furthermore, we follow the guideline of the PCS by defining the date of crime as the end date of the estimated time frame of crime. Consequently, for the case of BW, the data is further restricted to observations for which the estimated start and end of the crime is reported to be on the same date. This way, only offenses for which the estimated time frame can be attributed to one exact date are accounted for, as we investigate the daily short-term variation in air pollution on crime. As the data

[23] In email correspondence with the CIDs, we have been informed that offenses can be added to the PCS in retrospective. For instance, if a victim reports the experienced sexual crime to the police after ten years, the offense will be added to the PCS in the year of reporting. This way, the time of the offense relies on the testimony of the victim.

for RLP only provides the date of crime as defined in PCS, there is no further restriction possible which constitute a potential caveat. However, the aforementioned crimes with high possibility of measurement error concerning reporting are excluded and, in addition, crime data of RLP only represent 11% percent of our overall data in the analysis. Additionally, we exclude all entries for which the locality of the crime cannot be defined or be attributed to a direct municipality key.

10 Appendix B: Additional Tables

Table A.1: Overview Overall Number of Observations and Observations Lost per Restriction

State	BW	RLP
Raw*	9859183	1031346
Between<40	827424	0
Year >2000	8328	0
Matching AGS	32104	4917
Year>2004	1177199	0
Recode AGS	86	0
Time Frame	6155499	237819
Between == 0	367390	0
Crimetypes	591054	369288
Geo subset	267544	361936
Group subset	1621	2773
Final**	430934	54613

Notes: **Raw* depicts the number of observations as delivered from the CID´s of the states. **Final* depicts the number of observations after all restrictions apply. The numbers inbetween represent the obsevations lost due to that restriction in comparison to the previous line. *Between* is defined as the estimated time frame a crime has occured counted in days.

Table A.2: Average Crime and Number of Days Observed per Year per Group (per Crime Type)

year	group	Average	N All Crimes	N Theft	N Violence	N Criminal Damage	N Sexual	N Robbery
2015	1	75.75	281	281	281	275	84	135
2016	1	72.64	301	301	300	298	98	137
2017	1	68.99	335	335	335	329	124	149
2015	2	11.46	346	341	284	261	31	39
2016	2	10.74	366	366	303	268	40	42
2017	2	10.91	354	346	293	271	55	30
2015	3	13.86	348	346	316	275	43	62
2016	3	13.18	359	356	335	302	43	67
2017	3	11.88	359	355	327	302	50	57
2015	4	1.58	352	199	92	93	9	4
2016	4	1.46	362	174	93	98	7	6
2017	4	1.40	336	142	82	104	15	6
2015	5	45.29	356	356	355	349	108	195
2016	5	39.10	364	364	363	354	125	158
2017	5	33.51	360	360	357	354	125	154
2015	6	196.68	348	348	348	348	259	297
2016	6	186.30	346	346	346	346	267	296
2017	6	175.86	333	333	333	333	281	261
2015	7	102.29	307	307	307	307	175	251
2016	7	100.17	331	331	331	331	216	264
2017	7	103.97	269	269	269	269	187	211
2015	8	0.91	357	107	57	63	12	0
2016	8	0.96	342	115	71	62	14	4
2017	8	1.02	345	111	57	72	10	3
2015	9	1.59	365	168	101	104	3	2
2016	9	1.64	357	187	93	104	3	4
2017	9	1.44	324	154	80	85	8	6
2015	10	20.11	360	360	340	334	71	63
2016	10	19.88	344	344	322	322	79	54
2017	10	20.33	336	336	317	316	83	69

Notes: The table depicts the average number of crimes per year per group in column called *Average*. The remaining columns show the number of days in which at least one crime is reported e.g., maximum number of days at least one crime can be observed is 365 or 366 in a leap year.
Source: Crime Data from CID BW and RLP.

Table A.3: Total Number of Crimes per Group per Type

year	group	All Crimes	Theft	Violence	Criminal Damage	Sexual	Robbery
2015	1	21285	13422	5175	2158	173	357
2016	1	21865	13440	5523	2346	209	347
2017	1	23112	13656	5918	2871	283	384
2015	2	3965	2608	722	552	34	49
2016	2	3932	2399	867	572	46	48
2017	2	3862	2380	784	608	58	32
2015	3	4824	2870	1185	654	46	69
2016	3	4730	2536	1253	823	46	72
2017	3	4264	2150	1285	711	52	66
2015	4	557	285	129	130	9	4
2016	4	528	251	139	125	7	6
2017	4	471	198	111	141	15	6
2015	5	16122	11478	2608	1606	140	290
2016	5	14233	9523	2700	1636	158	216
2017	5	12065	7693	2464	1536	159	213
2015	6	68444	38636	18418	9332	796	1262
2016	6	64461	33813	19242	9247	854	1305
2017	6	58560	29366	17992	9072	1135	995
2015	7	31403	20693	6021	3855	284	550
2016	7	33156	21491	6797	3935	363	570
2017	7	27969	17756	5867	3561	353	432
2015	8	324	153	80	79	12	0
2016	8	327	149	90	70	14	4
2017	8	353	163	84	91	11	4
2015	9	581	252	182	141	4	2
2016	9	584	298	154	124	4	4
2017	9	466	226	125	101	8	6
2015	10	7239	4392	1496	1194	88	69
2016	10	6837	4004	1391	1287	92	63
2017	10	6832	4022	1326	1308	98	78

Notes: Table depicts total number of crimes for each variable shown in the header.
Source: Crime Data from CID BW and RLP.

Table A.4: Overview Crime Type Definition

Number	Major	Minor	PCS Key	Description as in PCS
1	Criminal Damage:	CD to other Buildings	640***	Brandstiftung und Herbeiführen einer Brandgefahr §§ 306-306d, 306f StGB
		CD to a dwelling	641***	(Vorsätzliche) Brandstiftung und Herbeiführen einer Brandgefahr §§ 306-306c, 306f Abs. 1 und 2 StGB
		Other CD	674***	Sachbeschädigung §§ 303-305a StGB
		CD to a motor vehicle		
2	Robbery:	Business Property	210***	Sonstiger Raub § 249 StGB
		Personal Property	211***	Raub, räuberische Erpressung auf/gegen Geldinstitute, Postfilialen und -agenturen
			212***	Raub, räuberische Erpressung auf/gegen sonstige Zahlstellen und Geschäfte
			213***	Raub, räuberische Erpressung auf/gegen Geld- und Werttransporte
			214***	Räuberischer Angriff auf Kraftfahrer § 316a StGB
			217***	Sonstige Raubüberfälle auf Straßen, Wegen oder Plätzen
			216***	Handtaschenraub
			218***	Raub zur Erlangung von Betäubungsmitteln
			219***	Raubüberfälle in Wohnungen
3	Sex Offences:	Rape	111***	Vergewaltigung, sexuelle Nötigung und sexueller Übergriff im besonders schweren Fall einschl. mit Todesfolge §§ 177, 178 StGB
		Other Sexual	112***	Sexueller Übergriff und sexuelle Nötigung § 177 Abs. 1, 2, 4, 5, 9 StGB
			113***	Sexueller Missbrauch von Schutzbefohlenen pp., unter Ausnutzung einer Amtsstellung oder eines Vertrauensverhältnisses §§ 174, 174a-c StGB
			114***	Sexuelle Belästigung § 184i StGB
			115***	Straftaten aus Gruppen § 184j StGB
			131***	Sexueller Missbrauch von Kindern §§ 176, 176a, 176b StGB
			132***	Exhibitionistische Handlungen und Erregung öffentlichen Ärgernisses §§ 183, 183a StGB
			133***	Sexueller Missbrauch von Jugendlichen § 182 StGB
			134***	Sexueller Missbrauch Widerstandsunfähiger § 179 StGB
4	Violence againts the Person:	ABH; assault with bodily harm	221***	Körperverletzung mit Todesfolge §§ 227, 231 StGB
		GBH; grievous bodily harm	222***	Gefährliche und schwere Körperverletzung, Verstümmelung weiblicher Genitalien §§ 224, 226, 226a, 231 StGB
		Harrasment	223***	Misshandlung von Schutzbefohlenen § 225 StGB
		Murder	224***	Vorsätzliche einfache Körperverletzung § 223 StGB
		Offensive Weapon	225***	Fahrlässige Körperverletzung § 229 StGB
		Other Violence	01****	Mord § 211 StGB
			02****	Totschlag und Tötung auf Verlangen §§ 212, 213, 216 StGB
			03****	Fahrlässige Tötung § 222 StGB - nicht i.V.m. Verkehrsunfall -
5	Theft and Handling:	Handling stolen goods	3*****	Sonstiger einfacher Diebstahl §§ 242, 247, 248a-c StGB
		Other Theft	4*****	Sonstiger schwerer Diebstahl insgesamt §§ 243 - 244a StGB
		Picking Pockets		
		Theft from a MV		
		Theft of pedal cycles		
		Snatches		
		Motor Vehicles interference - and tampering		
		Theft from Shops		
		Theft / Taking of Motor Vehicles		
6	Burglary:	Burglary in Other Buildings		
		Burglary in a Dwelling		

Notes: Major and minor crime and crime type definition according to Bondy, Roth and Sager (2020) which base on the files of Draca, Machin and Witt (2011). The equivalent PCS key and German description is presented on the right hand side of the table.
Source: Own Table base on information from Bondy, Roth and Sager (2020), Draca, Machin and Witt (2011) and PCS.

Table A.5: Pooled and Fixed Effect Model: Bootstrapped Standard Errors

	(1) All Crimes	(2) All Crimes	(3) All Crimes	(4) All Crimes	(5) All Crimes
PM10	-0.002	0.029	0.030***	0.028**	0.036***
	(0.0352)	(0.0222)	(0.0088)	(0.0095)	(0.0095)
Rainfall		-0.004	-0.004***	-0.005***	-0.003***
		(0.0021)	(0.0004)	(0.0004)	(0.0003)
Temperature_bin_1		0.000	0.000	0.000	0.000
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2		0.366	0.458**	0.542***	0.258
		(0.2772)	(0.1617)	(0.1430)	(0.1338)
Temperature_bin_3		0.262	0.406**	0.529***	0.312*
		(0.3077)	(0.1464)	(0.1433)	(0.1291)
Temperature_bin_4		0.480	0.550**	0.613**	0.256
		(0.3504)	(0.2082)	(0.2064)	(0.1612)
Temperature_bin_5		0.944**	0.704*	0.650*	0.144
		(0.3302)	(0.3111)	(0.2634)	(0.2095)
RelativeHumidity		-0.009	-0.007	-0.008	-0.007*
		(0.0108)	(0.0051)	(0.0049)	(0.0033)
Temperature_bin_1*RelativeHumidity		0.000	0.000	0.000	0.000
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2*RelativeHumidity		-0.003	-0.004	-0.004**	-0.002
		(0.0029)	(0.0020)	(0.0015)	(0.0016)
Temperature_bin_3*RelativeHumidity		-0.001	-0.002	-0.004*	-0.002
		(0.0032)	(0.0020)	(0.0016)	(0.0015)
Temperature_bin_4*RelativeHumidity		-0.004	-0.004	-0.005	-0.001
		(0.0036)	(0.0027)	(0.0026)	(0.0022)
Temperature_bin_5*RelativeHumidity		-0.011**	-0.006	-0.005	0.001
		(0.0037)	(0.0042)	(0.0035)	(0.0030)
RelativeHumidity_Quadratic		0.000	0.000**	0.000*	0.000**
		(0.0001)	(0.0000)	(0.0000)	(0.0000)
Temperature_Quadratic		-0.001*	-0.000	-0.000	-0.000
		(0.0002)	(0.0003)	(0.0002)	(0.0001)
RelativeHumidity_Quad*Temperature_Quad		0.000*	0.000	0.000	0.000
		(0.0000)	(0.0000)	(0.0000)	(0.0000)
_cons	-9.061***	-9.011***	-9.158***	-9.149***	-8.979***
	(0.1269)	(0.5145)	(0.2911)	(0.2264)	(0.2194)
Controls	No	Yes	Yes	Yes	Yes
Group FE	No	No	Yes	Yes	Yes
DOW FE	No	No	No	Yes	Yes
Year-Month FE	No	No	No	No	Yes
one	Yes	Yes	No	No	No
N	10243	10243	10243	10243	10243

Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) until (5) a PPML is estimated with group population as the offset. In column (6), the model is estimated with linear OLS using crime rate per 100,000 people as the dependent variable. The constant *one* is included if no fixed effect is inlcuded when employing PPML

Source: Data as described in Section 2.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: Intrumental Variable Model: Control Function (CF) with Bootstrapped Standard Errors

	(1) All Crimes	(2) All Crimes	(3) All Crimes	(4) All Crimes	(5) All Crimes
PM10	0.036***	0.011	0.040**	0.042***	0.046***
	(0.0095)	(0.1063)	(0.0124)	(0.0099)	(0.0104)
Rainfall	-0.003***	-0.004	-0.004***	-0.004***	-0.003***
	(0.0003)	(0.0021)	(0.0005)	(0.0005)	(0.0004)
Temperature_bin_1	0.000	0.000	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2	0.258	0.339	0.448**	0.526***	0.238
	(0.1338)	(0.3044)	(0.1645)	(0.1467)	(0.1371)
Temperature_bin_3	0.312*	0.224	0.407**	0.529***	0.296*
	(0.1291)	(0.3601)	(0.1471)	(0.1419)	(0.1351)
Temperature_bin_4	0.256	0.443	0.566**	0.635**	0.250
	(0.1612)	(0.4079)	(0.2063)	(0.1984)	(0.1661)
Temperature_bin_5	0.144	0.901*	0.728*	0.684**	0.141
	(0.2095)	(0.3843)	(0.3066)	(0.2529)	(0.2124)
RelativeHumidity	-0.007*	-0.010	-0.007	-0.008	-0.007*
	(0.0033)	(0.0122)	(0.0050)	(0.0048)	(0.0031)
Temperature_bin_1*RelativeHumidity	0.000	0.000	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2*RelativeHumidity	-0.002	-0.003	-0.003	-0.004*	-0.002
	(0.0016)	(0.0032)	(0.0021)	(0.0016)	(0.0016)
Temperature_bin_3*RelativeHumidity	-0.002	-0.001	-0.002	-0.003	-0.002
	(0.0015)	(0.0037)	(0.0021)	(0.0016)	(0.0017)
Temperature_bin_4*RelativeHumidity	-0.001	-0.004	-0.004	-0.005	-0.001
	(0.0022)	(0.0042)	(0.0027)	(0.0026)	(0.0023)
Temperature_bin_5*RelativeHumidity	0.001	-0.010*	-0.006	-0.005	0.001
	(0.0030)	(0.0040)	(0.0042)	(0.0035)	(0.0031)
RelativeHumidity_Quadratic	0.000**	0.000	0.000**	0.000*	0.000**
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Temperature_Quadratic	-0.000	-0.001**	-0.000	-0.000	-0.000
	(0.0001)	(0.0002)	(0.0003)	(0.0002)	(0.0001)
RelativeHumidity_Quad*Temperature_Quad	0.000	0.000**	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Residual_PM10		0.021	-0.010	-0.015	-0.011
		(0.0970)	(0.0103)	(0.0099)	(0.0132)
_cons	-8.979***	-8.919***	-9.184***	-9.186***	-8.985***
	(0.2194)	(0.7357)	(0.2875)	(0.2307)	(0.2148)
Controls	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	No	Yes	Yes	Yes
DOW FE	Yes	No	No	Yes	Yes
Year-Month FE	Yes	No	No	No	Yes
one	No	Yes	No	No	No
N	10243	10243	10243	10243	10243

Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) until (5) a PPML is estimated with group population as the offset. In column (6), the model is estimated with 2SLS using crime rate per 100,000 people as the dependent variable. The constant *one* is included if no fixed effect is inlcuded when employing PPML.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.7: Intrumental Variable Model: Bootstrapped Standard Errors

	(1) All Crimes	(2) All Crimes	(3) All Crimes	(4) All Crimes	(5) All Crimes
PM10	0.036***	-0.004	0.036**	0.039***	0.043***
	(0.0095)	(0.1011)	(0.0119)	(0.0106)	(0.0115)
Rainfall	-0.003***	-0.005*	-0.004***	-0.005***	-0.003***
	(0.0003)	(0.0021)	(0.0006)	(0.0007)	(0.0005)
Temperature_bin_1	0.000	0.000	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2	0.258	0.479	0.528***	0.585**	0.321
	(0.1338)	(0.3246)	(0.1442)	(0.2002)	(0.1815)
Temperature_bin_3	0.312*	0.346	0.495***	0.593**	0.387
	(0.1291)	(0.3735)	(0.1439)	(0.2047)	(0.2069)
Temperature_bin_4	0.256	0.490	0.626*	0.675**	0.313
	(0.1612)	(0.4068)	(0.2529)	(0.2553)	(0.2239)
Temperature_bin_5	0.144	0.902*	0.765*	0.705*	0.175
	(0.2095)	(0.3801)	(0.3589)	(0.2921)	(0.2476)
RelativeHumidity	-0.007*	-0.010	-0.007	-0.008	-0.007*
	(0.0033)	(0.0128)	(0.0043)	(0.0043)	(0.0029)
Temperature_bin_1*RelativeHumidity	0.000	0.000	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_bin_2*RelativeHumidity	-0.002	-0.005	-0.005**	-0.005*	-0.003
	(0.0016)	(0.0035)	(0.0017)	(0.0024)	(0.0023)
Temperature_bin_3*RelativeHumidity	-0.002	-0.003	-0.003	-0.004	-0.003
	(0.0015)	(0.0039)	(0.0017)	(0.0024)	(0.0027)
Temperature_bin_4*RelativeHumidity	-0.001	-0.005	-0.005	-0.005	-0.002
	(0.0022)	(0.0045)	(0.0035)	(0.0035)	(0.0032)
Temperature_bin_5*RelativeHumidity	0.001	-0.011*	-0.007	-0.005	0.000
	(0.0030)	(0.0044)	(0.0050)	(0.0040)	(0.0035)
RelativeHumidity_Quadratic	0.000**	0.000	0.000*	0.000*	0.000*
	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Temperature_Quadratic	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.0001)	(0.0002)	(0.0003)	(0.0002)	(0.0001)
RelativeHumidity_Quad*Temperature_Quad	0.000	0.000*	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
_cons	-8.979***	-8.949***	-9.211***	-9.199***	-9.018***
	(0.2194)	(0.6548)	(0.2277)	(0.1925)	(0.2071)
Controls	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	No	Yes	Yes	Yes
DOW FE	Yes	No	No	Yes	Yes
Year-Month FE	Yes	No	No	No	Yes
one	No	Yes	No	No	No
N	10243	10243	10243	10243	10243

Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) until (5) a PPML is estimated with group population as the offset. In column (6), the model is estimated with 2SLS using crime rate per 100,000 people as the dependent variable. The constant *one* is included if no fixed effect is inlcuded when employing PPML.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.8: Intrumental Variable Model: First Stage

	(1)	(2)	(3)	(4)	(5)
	PM10	PM10	PM10	PM10	PM10
WindDirection_90_1	0.362**	0.341*	0.392**	0.464**	0.464**
	(0.1351)	(0.1358)	(0.1384)	(0.1499)	(0.1499)
WindDirection_180_1	0.120	0.098	0.097	0.103	0.103
	(0.0890)	(0.0920)	(0.0922)	(0.0844)	(0.0844)
WindDirection_360_1	-0.128	-0.142	-0.126	-0.060	-0.060
	(0.1382)	(0.1415)	(0.1411)	(0.1106)	(0.1106)
WindDirection_90_2	0.259*	0.292**	0.303**	0.343***	0.343***
	(0.1043)	(0.1091)	(0.1099)	(0.1024)	(0.1024)
WindDirection_180_2	0.152**	0.189**	0.190**	0.156**	0.156**
	(0.0509)	(0.0596)	(0.0591)	(0.0526)	(0.0526)
WindDirection_360_2	-0.130	-0.104	-0.093	-0.031	-0.031
	(0.0917)	(0.0964)	(0.0959)	(0.0883)	(0.0883)
WindDirection_90_3	0.547***	0.709***	0.717***	0.725***	0.725***
	(0.1411)	(0.1451)	(0.1461)	(0.1306)	(0.1306)
WindDirection_180_3	0.159**	0.311***	0.313***	0.303***	0.303***
	(0.0532)	(0.0619)	(0.0618)	(0.0591)	(0.0591)
WindDirection_360_3	-0.003	0.150	0.159	0.148	0.148
	(0.0892)	(0.0951)	(0.0950)	(0.0981)	(0.0981)
WindDirection_90_4	0.972***	0.801***	0.806***	0.753***	0.753***
	(0.1044)	(0.1089)	(0.1089)	(0.0971)	(0.0971)
WindDirection_180_4	0.717***	0.548***	0.544***	0.491***	0.491***
	(0.0648)	(0.0715)	(0.0714)	(0.0657)	(0.0657)
WindDirection_360_4	0.197*	0.022	0.033	0.060	0.060
	(0.0832)	(0.0886)	(0.0853)	(0.0889)	(0.0889)
WindDirection_90_5	0.622***	0.463***	0.473***	0.476***	0.476***
	(0.0696)	(0.0740)	(0.0745)	(0.0666)	(0.0666)
WindDirection_180_5	0.606***	0.445***	0.443***	0.391***	0.391***
	(0.0708)	(0.0747)	(0.0750)	(0.0700)	(0.0700)
WindDirection_360_5	-0.018	-0.166***	-0.131***	0.004	0.004
	(0.0215)	(0.0331)	(0.0387)	(0.0478)	(0.0478)
WindDirection_90_6	0.546***	0.559***	0.569***	0.531***	0.531***
	(0.1352)	(0.1397)	(0.1396)	(0.1232)	(0.1232)
WindDirection_180_6	0.553***	0.571***	0.570***	0.545***	0.545***
	(0.0627)	(0.0712)	(0.0709)	(0.0660)	(0.0660)
WindDirection_360_6	0.193**	0.206*	0.215**	0.287***	0.287***
	(0.0730)	(0.0807)	(0.0798)	(0.0784)	(0.0784)
WindDirection_90_7	1.095***	0.711**	0.719**	0.567**	0.567**
	(0.2303)	(0.2336)	(0.2293)	(0.1953)	(0.1953)
WindDirection_180_7	0.870***	0.480***	0.474***	0.434***	0.434***
	(0.0586)	(0.0672)	(0.0669)	(0.0626)	(0.0626)
WindDirection_360_7	0.167	-0.213*	-0.206*	-0.168	-0.168
	(0.0885)	(0.0945)	(0.0923)	(0.0893)	(0.0893)
WindDirection_90_8	-0.309***	0.124	0.127	0.105	0.105
	(0.0797)	(0.0817)	(0.0822)	(0.0835)	(0.0835)
WindDirection_180_8	-0.300***	0.135*	0.143*	0.093	0.093
	(0.0562)	(0.0595)	(0.0596)	(0.0588)	(0.0588)
WindDirection_360_8	-0.480***	-0.053	-0.054	0.023	0.023
	(0.0705)	(0.0737)	(0.0741)	(0.0820)	(0.0820)
WindDirection_90_9	-0.286**	0.132	0.129	0.137	0.137
	(0.0905)	(0.0946)	(0.0943)	(0.1074)	(0.1074)
WindDirection_180_9	-0.265***	0.145**	0.148**	0.097	0.097
	(0.0462)	(0.0521)	(0.0525)	(0.0497)	(0.0497)
WindDirection_360_9	-0.240***	0.163*	0.173**	0.191**	0.191**
	(0.0587)	(0.0633)	(0.0632)	(0.0605)	(0.0605)
WindDirection_90_10	0.811***	0.490***	0.497***	0.469***	0.469***
	(0.0752)	(0.0796)	(0.0791)	(0.0722)	(0.0722)
WindDirection_180_10	0.592***	0.276***	0.280***	0.245***	0.245***
	(0.0745)	(0.0790)	(0.0785)	(0.0728)	(0.0728)
WindDirection_360_10	-0.227*	-0.535***	-0.516***	-0.412***	-0.412***
	(0.1034)	(0.1065)	(0.0987)	(0.1081)	(0.1081)
Rainfall	-0.025***	-0.025***	-0.026***	-0.020***	-0.020***
	(0.0020)	(0.0020)	(0.0020)	(0.0019)	(0.0019)
Temperature_bin_1	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)
Temperature_bin_2	1.923	1.734	1.810	2.517*	2.517*
	(1.1597)	(1.1629)	(1.1719)	(1.2128)	(1.2128)
Temperature_bin_3	1.051	0.874	0.994	2.348	2.348
	(1.1469)	(1.1510)	(1.1594)	(1.2057)	(1.2057)
Temperature_bin_4	-0.691	-0.774	-0.638	1.487	1.487
	(1.1754)	(1.1778)	(1.1862)	(1.2301)	(1.2301)
Temperature_bin_5	-1.405	-1.470	-1.346	1.220	1.220
	(1.2505)	(1.2497)	(1.2572)	(1.2933)	(1.2933)
RelativeHumidity	0.050**	0.030	0.032*	0.056***	0.056***
	(0.0160)	(0.0158)	(0.0158)	(0.0163)	(0.0163)
Temperature_bin_1*RelativeHumidity	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)
Temperature_bin_2*RelativeHumidity	-0.039**	-0.037*	-0.038*	-0.042**	-0.042**
	(0.0148)	(0.0149)	(0.0150)	(0.0152)	(0.0152)
Temperature_bin_3*RelativeHumidity	-0.032*	-0.030*	-0.031*	-0.041**	-0.041**
	(0.0147)	(0.0148)	(0.0149)	(0.0151)	(0.0151)
Temperature_bin_4*RelativeHumidity	-0.012	-0.011	-0.013	-0.032*	-0.032*
	(0.0152)	(0.0152)	(0.0153)	(0.0155)	(0.0155)
Temperature_bin_5*RelativeHumidity	0.003	0.004	0.002	-0.026	-0.026
	(0.0164)	(0.0164)	(0.0165)	(0.0166)	(0.0166)
RelativeHumidity_Quadratic	-0.000	0.000	0.000	-0.000	-0.000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Temperature_Quadratic	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
RelativeHumidity_Quad*Temperature_Quad	-0.000***	-0.000***	-0.000***	-0.000**	-0.000**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	0.057	0.856	0.705	-1.236	-1.236
	(1.1319)	(1.1303)	(1.1369)	(1.1896)	(1.1896)
Controls	Yes	Yes	Yes	Yes	Yes
Group FE	No	Yes	Yes	Yes	Yes
DOW FE	No	No	Yes	Yes	Yes
Year-Month FE	No	No	No	Yes	Yes
one	Yes	No	No	No	No
F-test	38.881	20.725	19.597	15.817	15.817
N	10243	10243	10243	10243	10243

Notes: First stage results of IV estimation in Table 2 with robust standard errors. First number in wind direction dummies indicate the upper limit of the wind bin. The second number indicates the group number.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.9: Robustness Checks: CF with Bootstrapped Standard Errors

	(1) NY Excluded	(2) l(y+1) 2SLS	(3) Alt. Weather-C	(4) Wind Bin 60	(5) Restrictive Data
PM10	0.046***	0.042*	0.050***	0.051***	0.034***
	(0.0108)	(0.0163)	(0.0106)	(0.0075)	(0.0089)
Residual_PM10	-0.011	-0.020	-0.013	-0.016	0.001
	(0.0112)	(0.0161)	(0.0124)	(0.0110)	(0.0104)
Rainfall	-0.003***	-0.002	-0.003***	-0.003***	-0.003***
	(0.0005)	(0.0011)	(0.0004)	(0.0004)	(0.0008)
Temperature_bin_1	0.000	0.000		0.000	0.000
	(0.0000)	(0.0000)		(0.0000)	(0.0000)
Temperature_bin_2	0.238	0.640**		0.229	0.085
	(0.2376)	(0.2373)		(0.1361)	(0.5090)
Temperature_bin_3	0.296	0.630**		0.289*	0.106
	(0.2256)	(0.2296)		(0.1323)	(0.4443)
Temperature_bin_4	0.250	0.871**		0.246	0.072
	(0.2456)	(0.3195)		(0.1651)	(0.4676)
Temperature_bin_5	0.141	0.780		0.140	0.029
	(0.2818)	(0.4048)		(0.2141)	(0.4867)
RelativeHumidity	-0.007	0.001	-0.006*	-0.007*	-0.009
	(0.0036)	(0.0042)	(0.0025)	(0.0030)	(0.0059)
Temperature_bin_1*RelativeHumidity	0.000	0.000		0.000	0.000
	(0.0000)	(0.0000)		(0.0000)	(0.0000)
Temperature_bin_2*RelativeHumidity	-0.002	-0.007*		-0.001	0.000
	(0.0028)	(0.0031)		(0.0016)	(0.0061)
Temperature_bin_3*RelativeHumidity	-0.002	-0.007*		-0.002	0.001
	(0.0027)	(0.0030)		(0.0016)	(0.0053)
Temperature_bin_4*RelativeHumidity	-0.001	-0.010*		-0.001	0.001
	(0.0031)	(0.0042)		(0.0023)	(0.0055)
Temperature_bin_5*RelativeHumidity	0.001	-0.009		0.001	0.002
	(0.0038)	(0.0057)		(0.0031)	(0.0059)
RelativeHumidity_Quadratic	0.000***	0.000	0.000	0.000**	0.000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_Quadratic	-0.000*	-0.000	-0.001**	-0.000	-0.000
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
RelativeHumidity_Quad*Temperature_Quad	0.000	0.000**	0.000	0.000	0.000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature			0.014*		
			(0.0060)		
Temperature*RelativeHumidity			-0.000		
			(0.0001)		
_cons	-8.985***	2.276**	-8.914***	-8.986***	-8.832***
	(0.3083)	(0.6959)	(0.1446)	(0.2177)	(0.4855)
Controls	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes
DOW FE	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes
N	10186	10243	10243	10243	8788

Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) and column(3) until (5) a PPML is estimated with group population as the offset. In column (2), the model is estimated with 2SLS using log crime plus 1 as the dependent variable.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.10: Robustness Checks: Boostrapped Standard Errors

	(1) NY Excluded	(2) l(y+1) 2SLS	(3) Alt. Weather-C	(4) Wind Bin 60	(5) Restrictive Data
PM10	0.043***	0.042*	0.045***	0.047***	0.031**
	(0.0111)	(0.0163)	(0.0113)	(0.0090)	(0.0101)
Rainfall	-0.003***	-0.002	-0.003***	-0.003***	-0.004***
	(0.0005)	(0.0011)	(0.0005)	(0.0005)	(0.0009)
Temperature_bin_1	0.000	0.000		0.000	0.000
	(0.0000)	(0.0000)		(0.0000)	(0.0000)
Temperature_bin_2	0.321	0.640**		0.311	0.013
	(0.2270)	(0.2373)		(0.1805)	(0.5213)
Temperature_bin_3	0.387	0.630**		0.378	0.049
	(0.2150)	(0.2296)		(0.2041)	(0.4500)
Temperature_bin_4	0.313	0.871**		0.308	-0.006
	(0.2395)	(0.3195)		(0.2225)	(0.4818)
Temperature_bin_5	0.175	0.780		0.172	-0.071
	(0.2776)	(0.4048)		(0.2487)	(0.5080)
RelativeHumidity	-0.007*	0.001	-0.007*	-0.007*	-0.011
	(0.0035)	(0.0042)	(0.0027)	(0.0029)	(0.0066)
Temperature_bin_1*RelativeHumidity	0.000	0.000		0.000	0.000
	(0.0000)	(0.0000)		(0.0000)	(0.0000)
Temperature_bin_2*RelativeHumidity	-0.003	-0.007*		-0.003	0.001
	(0.0027)	(0.0031)		(0.0023)	(0.0061)
Temperature_bin_3*RelativeHumidity	-0.003	-0.007*		-0.003	0.001
	(0.0026)	(0.0030)		(0.0027)	(0.0052)
Temperature_bin_4*RelativeHumidity	-0.002	-0.010*		-0.002	0.002
	(0.0030)	(0.0042)		(0.0032)	(0.0055)
Temperature_bin_5*RelativeHumidity	0.000	-0.009		0.000	0.003
	(0.0037)	(0.0057)		(0.0035)	(0.0060)
RelativeHumidity_Quadratic	0.000***	0.000	0.000*	0.000*	0.000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature_Quadratic	-0.000	-0.000	-0.001**	-0.000	-0.000
	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.0001)
RelativeHumidity_Quad*Temperature_Quad	0.000	0.000**	0.000	0.000	0.000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Temperature			0.018*		
			(0.0082)		
Temperature*RelativeHumidity			-0.000		
			(0.0001)		
_cons	-9.018***	2.276**	-8.883***	-9.018***	-8.729***
	(0.3010)	(0.6959)	(0.1538)	(0.2096)	(0.5288)
Controls	Yes	Yes	Yes	Yes	Yes
Group FE	Yes	Yes	Yes	Yes	Yes
DOW FE	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes
N	10186	10243	10243	10243	8788

Notes: Each column in the table represents a separate regression. The dependent variable is shown in the header. Throughout column (1) and column(3) until (5) a PPML is estimated with group population as the offset. In column (2), the model is estimated with 2SLS using log crime plus 1 as the dependent variable.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

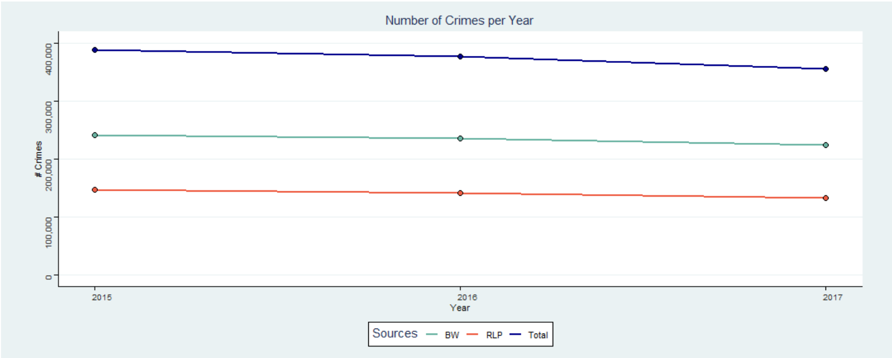
Table A.11: Placebo IV: First Stage

	(1) Placebo-IV
WindDirection_90_1	-0.051
	(0.0684)
WindDirection_180_1	-0.044
	(0.0700)
WindDirection_360_1	-0.118
	(0.0676)
WindDirection_90_2	-0.052
	(0.0681)
WindDirection_180_2	0.061
	(0.0672)
WindDirection_360_2	0.009
	(0.0687)
WindDirection_90_3	0.110
	(0.0800)
WindDirection_180_3	0.091
	(0.0769)
WindDirection_90_4	0.032
	(0.0815)
WindDirection_180_4	-0.096
	(0.0731)
WindDirection_360_4	0.028
	(0.0770)
WindDirection_360_3	0.062
	(0.0735)
WindDirection_90_5	-0.075
	(0.0727)
WindDirection_180_5	-0.119
	(0.0727)
WindDirection_360_5	-0.145*
	(0.0731)
WindDirection_90_6	-0.064
	(0.0720)
WindDirection_180_6	0.048
	(0.0753)
WindDirection_360_6	0.088
	(0.0800)
WindDirection_90_7	-0.057
	(0.0903)
WindDirection_180_7	-0.096
	(0.0800)
WindDirection_360_7	-0.062
	(0.0815)
WindDirection_90_8	0.032
	(0.0578)
WindDirection_180_8	-0.042
	(0.0573)
WindDirection_360_8	0.029
	(0.0589)
WindDirection_90_9	0.027
	(0.0639)
WindDirection_180_9	0.084
	(0.0567)
WindDirection_360_9	-0.035
	(0.0582)
WindDirection_90_10	0.050
	(0.0701)
WindDirection_180_10	0.035
	(0.0678)
WindDirection_360_10	0.104
	(0.0670)
Rainfall	-0.022***
	(0.0020)
Temperature_bin_1	0.000
	(.)
Temperature_bin_2	2.012
	(1.2299)
Temperature_bin_3	1.621
	(1.2227)
Temperature_bin_4	0.608
	(1.2461)
Temperature_bin_5	0.126
	(1.3041)
RelativeHumidity	0.035*
	(0.0167)
Temperature_bin_1*RelativeHumidity	0.000
	(.)
Temperature_bin_2*RelativeHumidity	-0.037*
	(0.0154)
Temperature_bin_3*RelativeHumidity	-0.034*
	(0.0153)
Temperature_bin_4*RelativeHumidity	-0.023
	(0.0157)
Temperature_bin_5*RelativeHumidity	-0.014
	(0.0167)
RelativeHumidity_Quadratic	0.000
	(0.0001)
Temperature_Quadratic	0.003***
	(0.0004)
RelativeHumidity_Quad*Temperature_Quad	-0.000***
	(0.0000)
Constant	0.268
	(1.2145)
Controls	Yes
Group FE	Yes
DOW FE	Yes
Year-Month FE	Yes
F-test	1.072
N	10243

Notes: First stage results for using randomly generated intruments as placebo test. Robust standard errors are reported in parantheses. First number in wind direction dummies indicate the upper bound of the wind bin. The second number indicates the group number.
Source: Data as described in Section 2.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

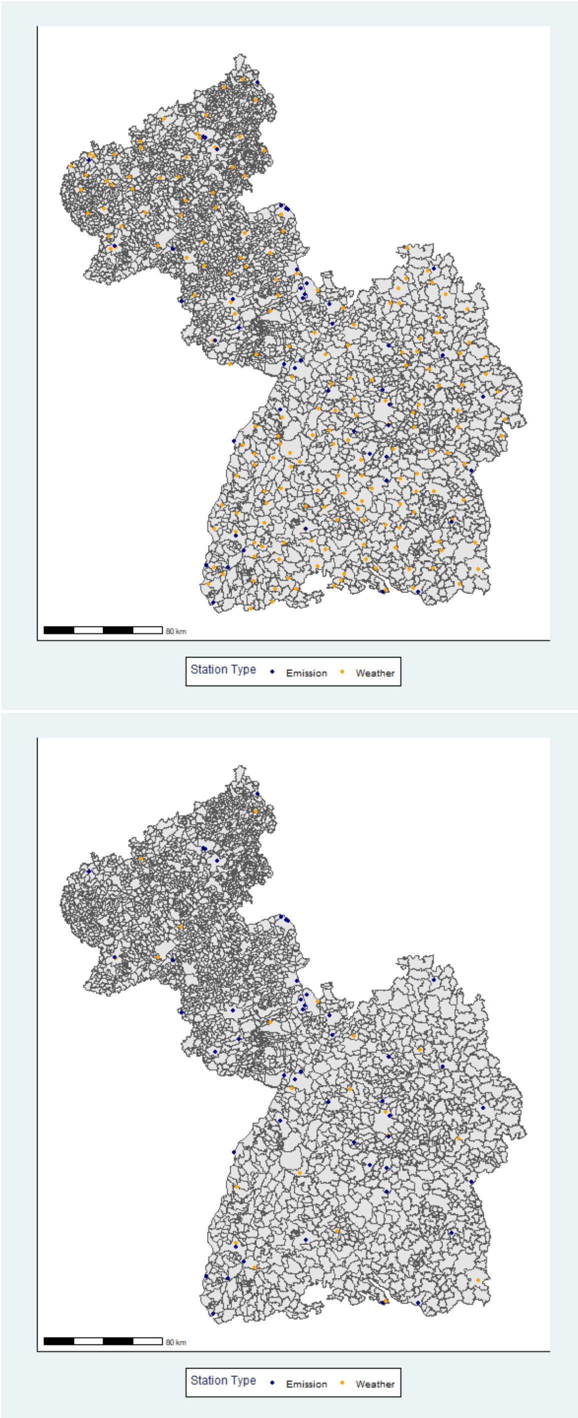
11 Appendix C: Additional Figures

Figure A.2: Trend Number of Crimes for the States of Baden-Wuerttemberg (BW) and Rhineland-Palatinate (RLP)



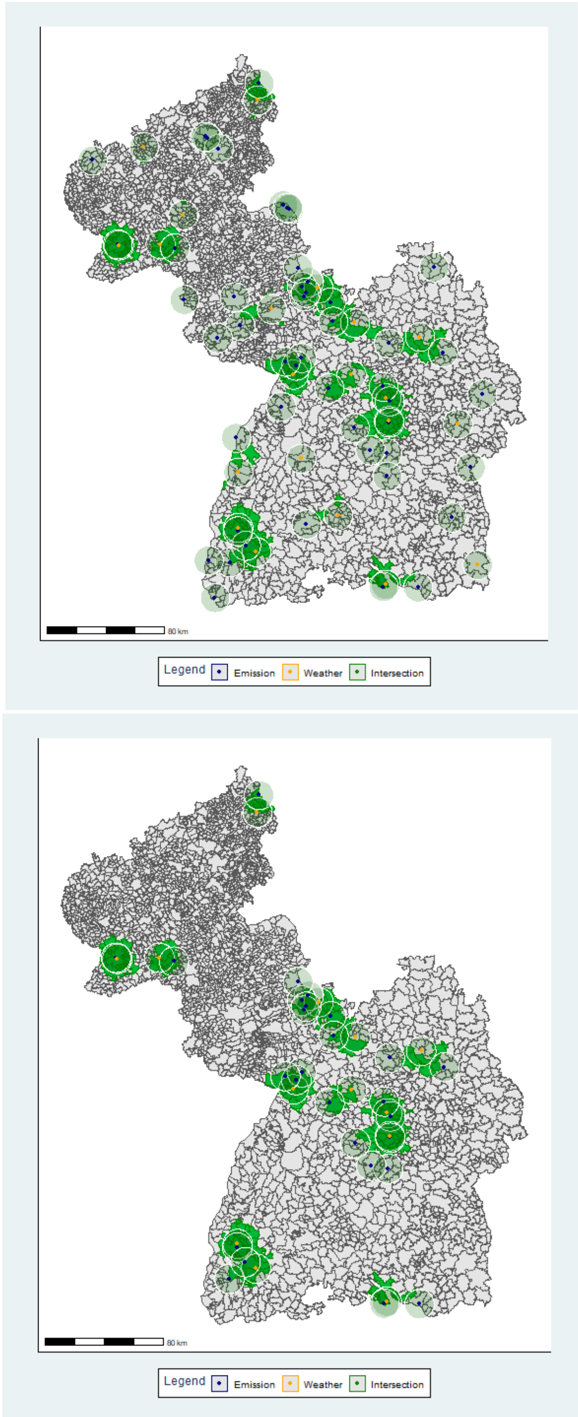
Notes: Overall trend of the number of crimes in BW and RLP.
Source: Own figure based on crime data from CID BW and RLP.

Figure A.3: Distribution of Stations before and after Consistency Restriction.



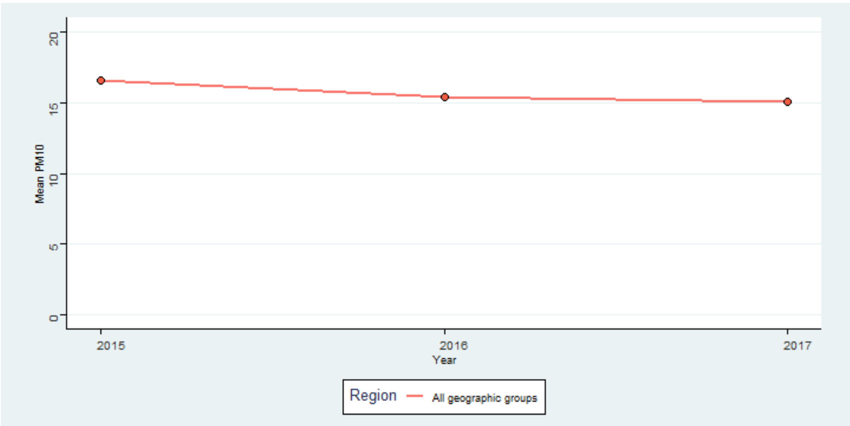
Notes: Distribution of stations before (top) and after (bottom) consistency restriction applied. The station number is based on the year 2015.
Source: Own figure based on environmental data provided by the DWD and UBA. Shape files base on © GeoBasis-DE / BKG 2019.

Figure A.4: Geographic Groups based on 10 km Buffer Zones with Station Locations with and without Marginal Municipalites



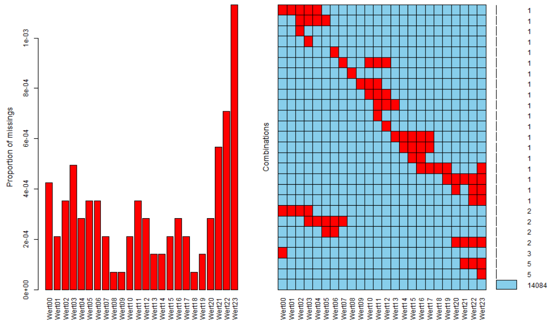
Notes: Resulting geographic groups based on 10 km buffer zones (circles) around emission and weather stations in BW and RLP. Before adjustment (top) and after (bottom) adjustment concerning polygon intersection.
Source: Own figure based on environmental data provided by the DWD and UBA. Shape files base on © GeoBasis-DE / BKG 2019.

Figure A.5: Yearly Trend of PM10 Concentrations

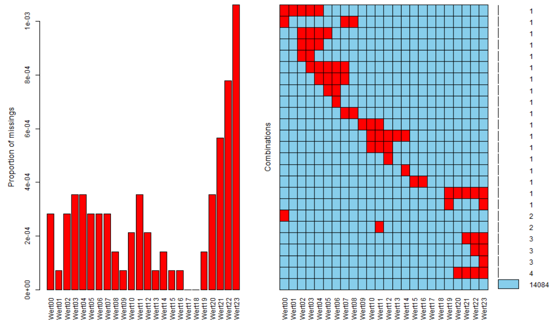


Notes: Trend line depcits the annual mean concentration of PM10 for the selected geographical groups.
Source: Own figure based on emission data provided by the UBA.

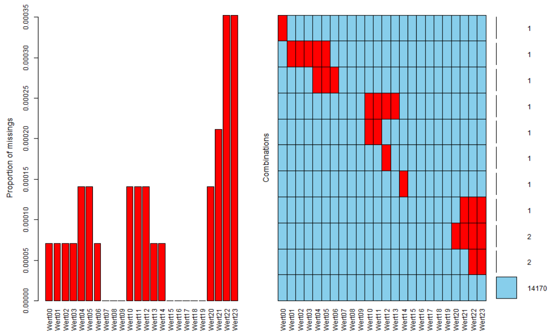
Figure A.6: Heatmap and Frequency of Missing Entries of Weather Variables



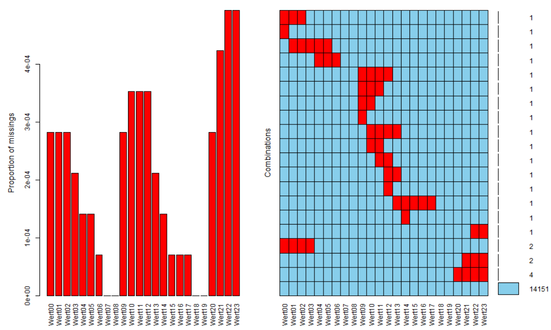
(a) Wind Direction



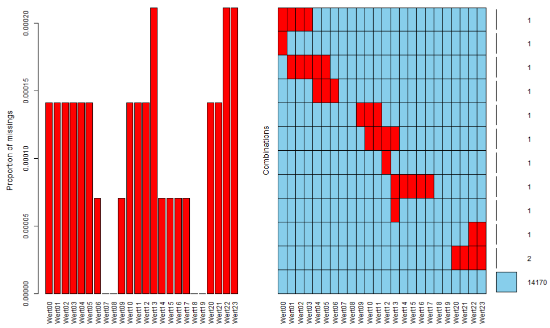
(b) Wind Speed



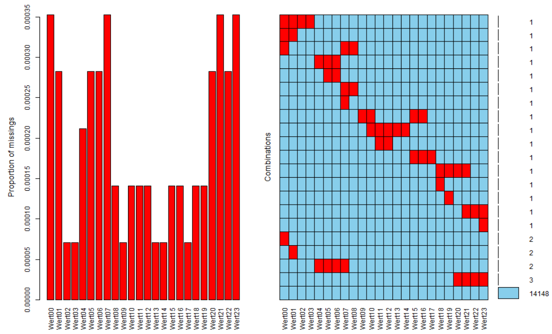
(c) Air Pressure



(d) Relative Humidity



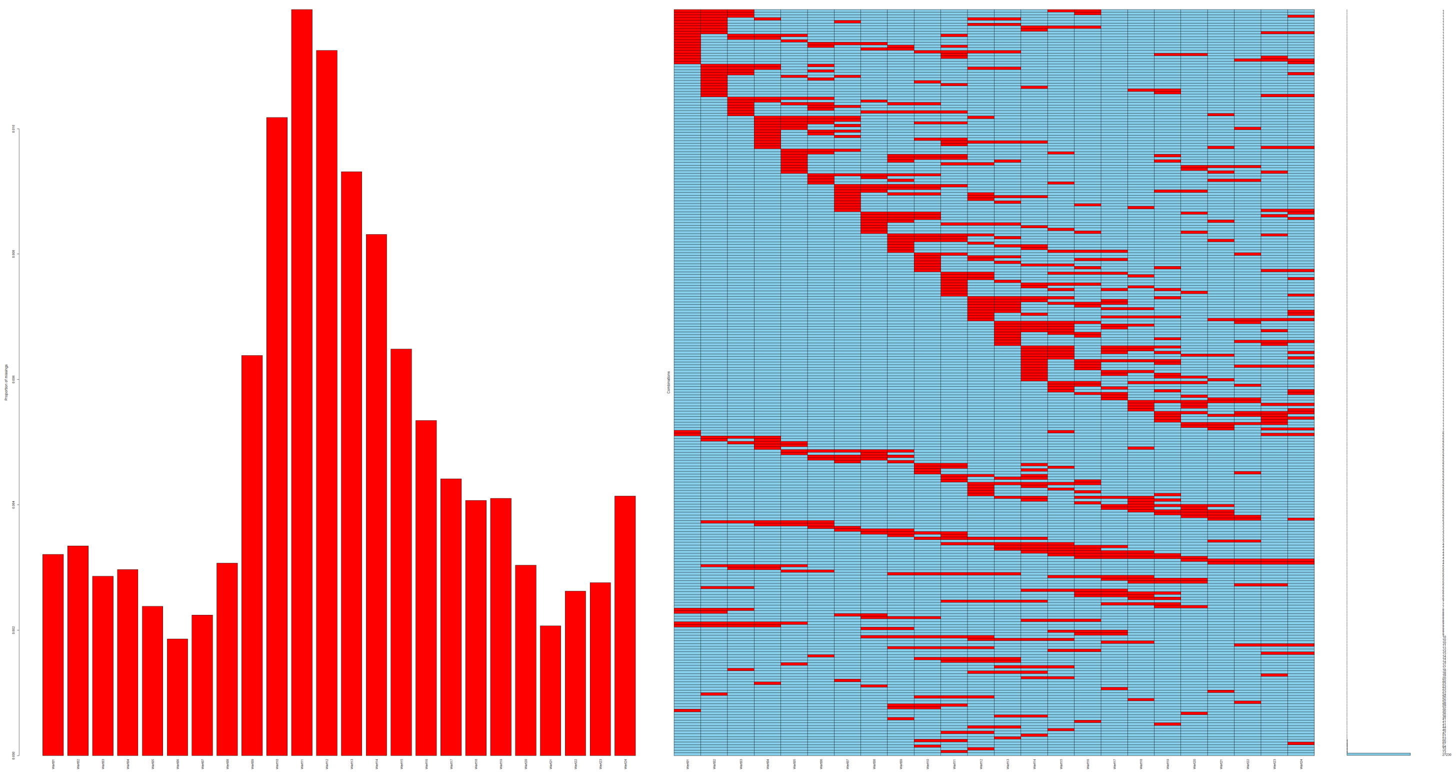
(e) Temperature



(f) Rainfall

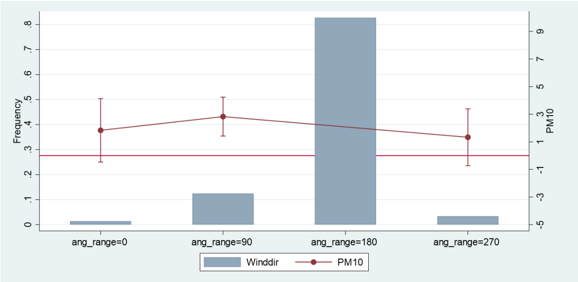
Notes: Each plots depicts a frequency plot (left) and heatmap (right) of missing entries for the mentioned weather variables. The x-axis of the frequency plot lists all 24 hour entries for which *Wert00* stands for midnight, whereby the y-axis show the proportion of missings in each hour entry. The heatmap depicts in each row a certain combination of missing and non-missing entries for which the number of occurences for the displayed combination is shown on the very right of each row. The numbers on occurences are sorted from lowest (at the top) to highest (at the bottom). The red colored column depict missing entries for the certain hour entry, whereby blue means entry observed. The frequency plot shows that most missing entries occured during night time but there is no systemic missing structure among all weather variables. A similar conclusion can be derived from the heatmaps, as no sequence of missing and non-missing entries occur systematically for each of the weather variables as well as among all weather variables.
Source: Own graph based on the data set provided by DWD.

Figure A.7: Heatmap and Frequency of Missing Entries PM10



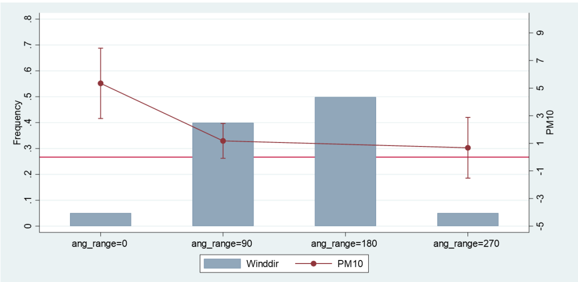
Notes: Frequency (left)and heatmap (right) of missing entries for PM10. The x-axis of the frequency plot lists all 24 hour entries for which *Wert00* stands for midnight, whereby the y-axis show the proportion of missings in each hour entry. The heatmap depicts in each row combinations of missing and non-missing entries for which the number of occurences for the displayed combination is shown on the very right of each row. The numbers on occurences are sorted from lowest (at the top) to highest (at the bottom). Due to high number of combinations for the heatmap, legends are automatically adjusted to small font. The red colored column depict missing entries for the certain hour entry, whereby blue means entry observed. The frequency plot shows that most missing entries occured during midday. However, the heatmap shows that there is little to none systematic combinations of missing and non-missing entries. The highest number combinations including missings is around 70 for which each of the combination only has one to two missing entries.
Source: Own graph based on the data set provided by UBA.

Figure A.8: Frequency of Wind Direction and First Stage Estimates of Equation 2

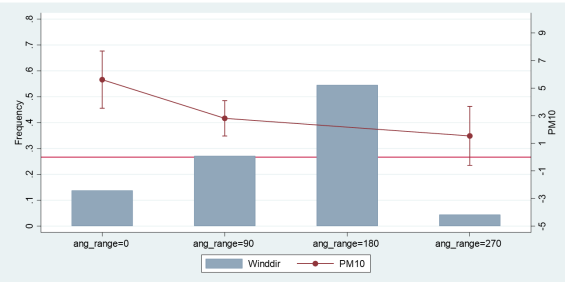


(a) Group 1

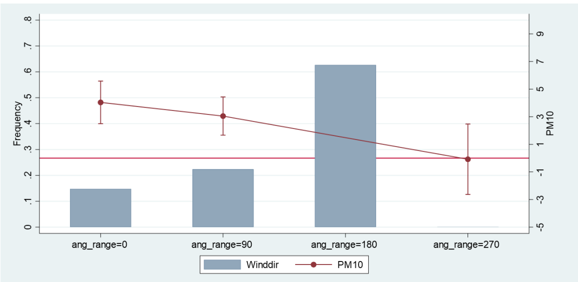
(b) Group 2



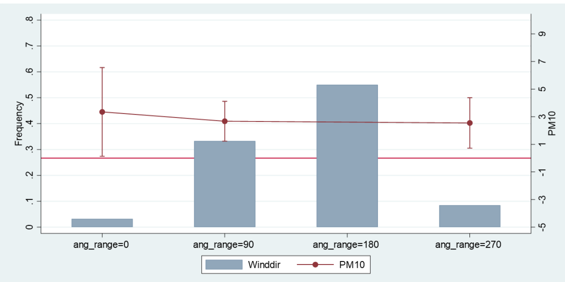
(c) Group 3



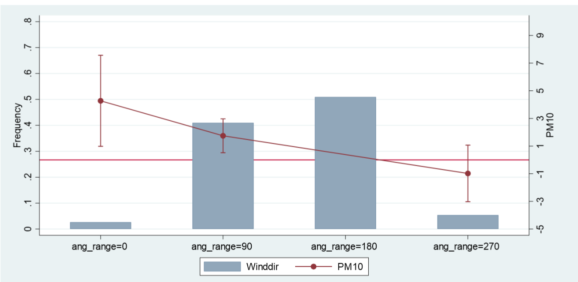
(d) Group 4



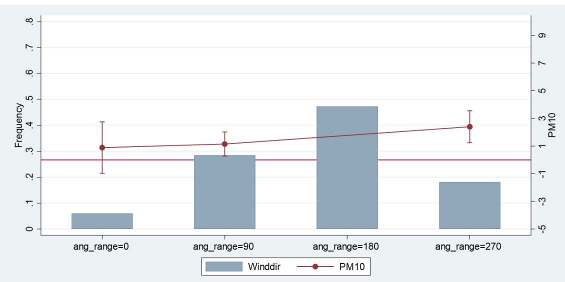
(e) Group 5



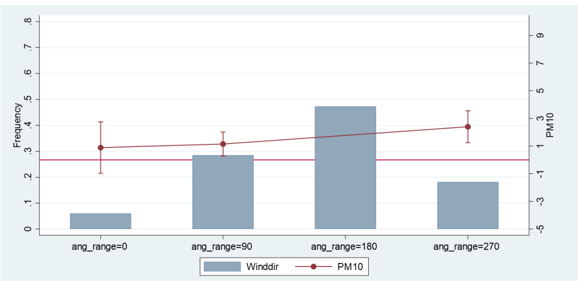
(f) Group 6



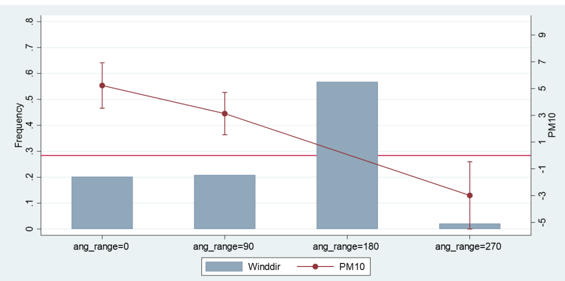
(g) Group 7



(h) Group 8



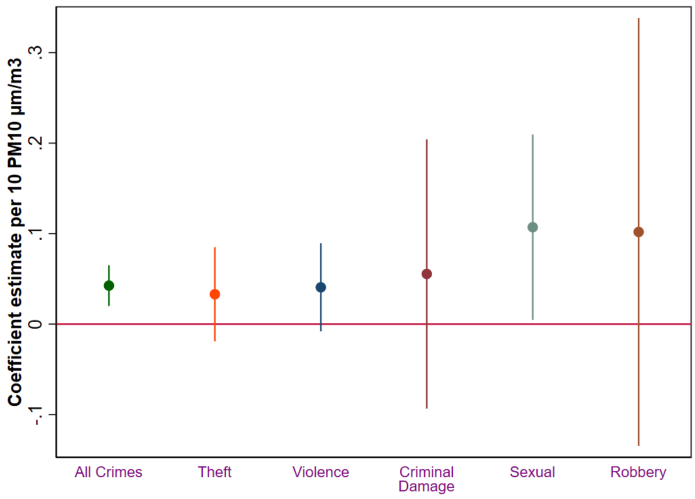
(i) Group 9



(j) Group 10

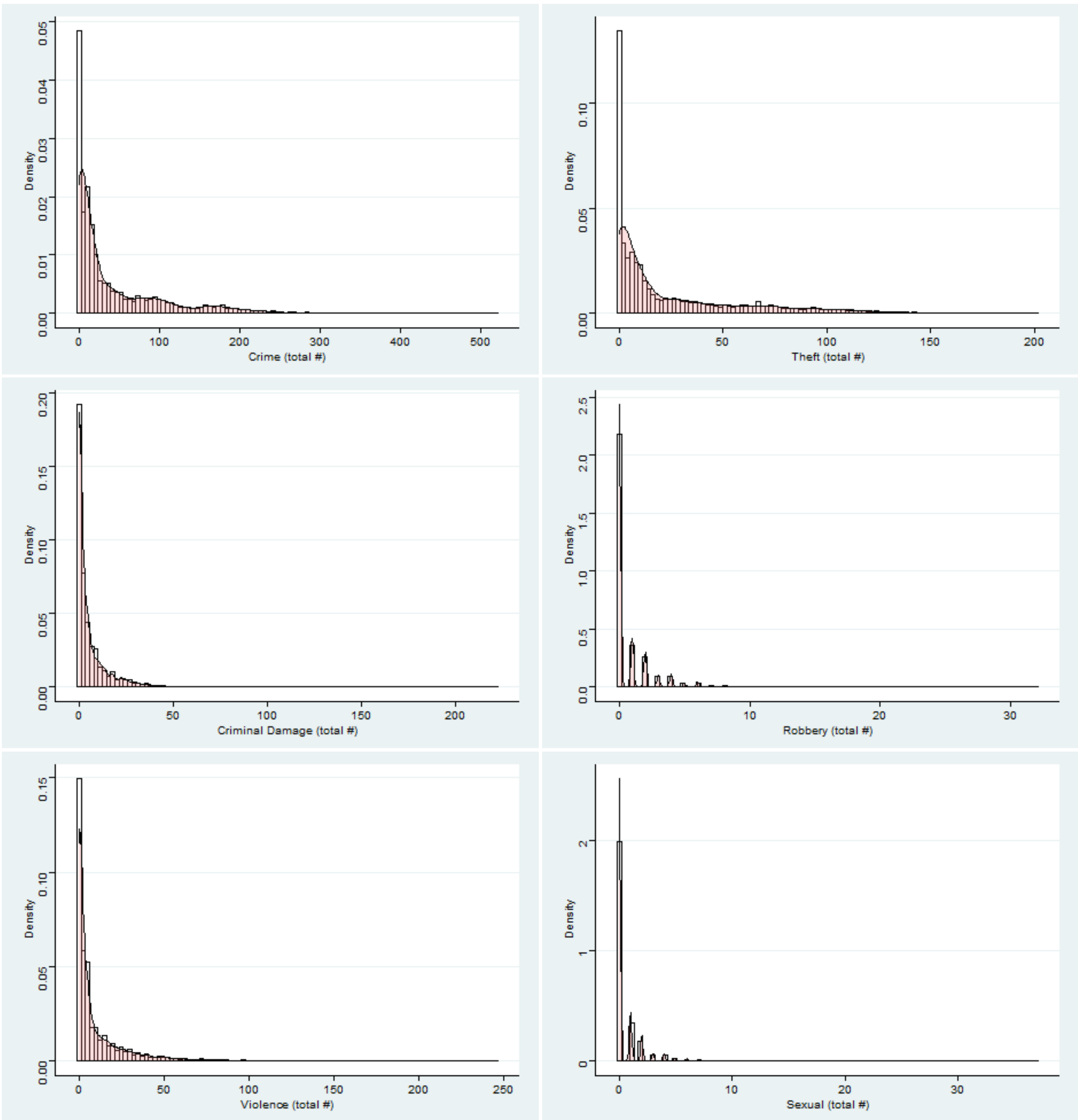
Notes: Each plots depicts the frequency of winddirection in the mentioned geographical group (bars). The relationship between PM10 and wind direction is obtained by a first stage estimation as depicted by Equation 2 using robust standard errors. First stage estimates and confidence intervals are illustrated by the red line. Group numbers base on Figure 2. The number of the angle range shows the lower limit of the wind direction bin.
Source: Own graph based on the data set provided by UBA and DWD.

Figure A.9: Heterogeneity Analysis: Bootstrapped Standard Errors



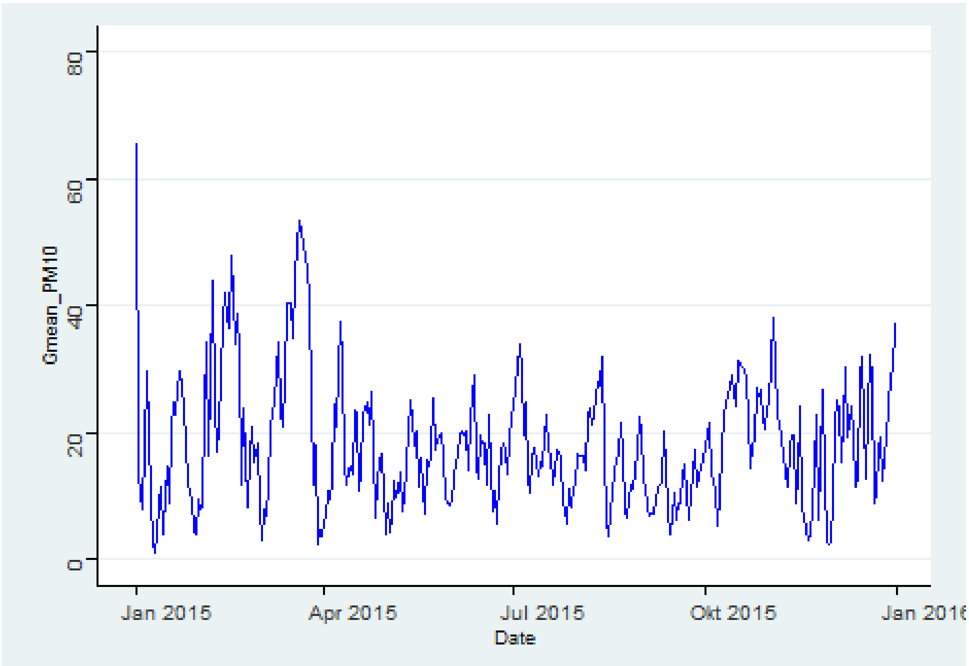
Notes: Coefficients based on estimates using bootstrapped standard errors. Each estimate is obtained by running the estimation with the depicted crime type as the dependent variable.
Source: Own graph. Data as describe in text.

Figure A.10: Density Plots for Overall Crime and Crime Types



Notes: Density plots for overall crime and each crime type by day by geographical group for the main estimation sample.
Source: Own graph based on the crime data provided by CID of BW and RLP.

Figure A.11: Daily Variation in PM10



Notes: Mean daily PM10 levels for group *(6) Stuttgart +'* in the year 2015.
Source: Own graph based on emission data by UBA.

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de