HEDG
HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 13/34

# The Short-Term Population Health Effects Of Weather and Pollution: Implications of Climate Change

Nicolas R. Ziebarth, Maike Schmitt and Martin Karlsson

December 2013

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

# THE SHORT-TERM POPULATION HEALTH EFFECTS OF WEATHER AND POLLUTION: IMPLICATIONS OF CLIMATE CHANGE ‡

Nicolas R. Ziebarth
Cornell University, DIW Berlin, and IZA Bonn*

Maike Schmitt
TU Darmstadt**

Martin Karlsson
University of Duisburg-Essen††

December 20, 2013

**Abstract**

This study comprehensively assesses the immediate effects of extreme weather conditions and high concentrations of ambient air pollution on population health. For Germany and the years 1999 to 2008, we link the universe of all 170 million hospital admissions, along with all 8 million deaths, with weather and pollution data reported at the day-county level. Extreme heat significantly increases hospitalizations and deaths. Extreme cold has a negligible effect on population health. High ambient $PM_{10}$, $O_3$ and $NO_2$ concentrations are associated with increased hospitalizations and deaths, particularly when *ignoring* simultaneous weather and pollution conditions. We find strong evidence for "harvesting", and that the instantaneous heat-health relationship is only present in the short-term. We calculate that one "Hot Day" with a temperature higher than 30 °C (86 °F) triggers short-term adverse health effects valued between \$0.10 and \$0.68 per resident.

**Keywords**: register data, hospital admissions, mortality, weather and pollution, climate change

**JEL classification**: I12; I18; Q51; Q53; Q54; Q58

‡We thank the German Federal Statistical Office (*Statistisches Bundesamt (destatis)*), the German Meteorological Service (*Deutscher Wetterdienst (DWD)*) and the German Federal Environmental Office (*Umweltbundesamt (UBA)*) that provided the data basis for the study as well as Joerg Blankenback for his great support in the interpolation of the geodesic data. Moreover, we thank Daniel Baumgarten, Antonio Bento, Damon Clark, Peter Eibich, Maria Fitzpatrick, Rick Geddes, Albrecht Glitz, Dan Grossman, Don Kenkel, Michael Kvasnicka, Mike Lovenheim, Dean Lillard, Sean Lyons, Alan Mathios, Jordan Matsudaira, Vincent Pohl, Emily Owens, Steve Stillmann, Hanna Wielandt, Robert Williams III, Will White, Martina Zweimüller and participants at the 2013 meeting of the American Economic Association (AEA), the $2^{nd}$ Workshop on Energy Policy and Environmental Economics at Cornell, the 2013 Conference of the European Society for Population Economics (ESPE) in Aarhus, the 2013 UK Health Economists' Study Group (HESG) Meeting at Warwick, the 2013 Canadian Health Economists' Study Group (CHESG) Meeting in Winnipeg, the Economics of Disease Conference in Darmstadt 2013 as well as seminar participants at the Institute on Health Economics, Health Behaviors and Disparities (IHEHBD) at Cornell University, the German Institute for Economic Research (DIW Berlin), and the Berlin Network of Labour Market Researchers (BeNA) for their helpful comments and discussions. In particular, we thank Evelyn Forget, Silviya Nikolova, Seiro Ito, and Reed Walker for outstanding discussions of this paper. We also thank Felix Heinemann (TU Darmstadt) and Peter Eibich (DIW Berlin) for excellent research assistance. All remaining errors or shortcomings of the article are our own. The research reported in this paper is not the result of a for-pay consulting relationship. Our employers do not have a financial interest in the topic of the paper which might constitute a conflict of interest. Funding from the Cornell Institute for Social Science (ISS) *Small Grant Program* as well as the Cornell Population Center (CPC) *Seed Grant Program* are gratefully acknowledged.

*Corresponding author: Nicolas R. Ziebarth, Cornell University, Department of Policy Analysis and Management (PAM), 106 Martha Van Rensselaer Hall, Ithaca, NY 14850, USA, phone: +1-(607)255-1180, fax: +1-(607)255-4071, e-mail: nrz2@cornell.edu

††University of Duisburg-Essen, Chair of Health Economics, Schützenbahn 70, 45117 Essen, Germany e-mail: martin.karlsson@uni-due.de

**Department of Law and Economics, Residenzschloss, Marktplatz 15, D-64283 Darmstadt, Germany e-mail: schmitt@vwl.tu-darmstadt.de

# 1 Introduction

Over the last decade, the economics literature has seen a staggering rise in the number of studies that empirically estimate the impact of air pollution on health outcomes. Certainly one reason for this boom lies in increasingly sophisticated data collection and availability. In addition, researchers and policymakers understand that state-of-the-art empirical investigations may produce evidence-based causal relationships that help to design effective and welfare increasing policy regulation. Air pollution that harms humans represents negative externalities that could be internalized by their producers via optimal policy regulation.

Although thorough cost-benefit analyses are still scant, almost all existing studies find that pollution negatively affects, or is negatively associated with, population health. This finding has been shown to hold particularly for vulnerable subgroups such as newborns (Currie and Schmieder, 2009; Currie and Walker, 2011), children (Chay and Greenstone, 2003; Nilsson, 2009; Currie et al., 2009), or the elderly (Villena et al., 2008; Schlenker and Walker, 2011; Karlsson and Schmitt, 2011), but also for whole populations (Almond et al., 2009). Outcome measures are typically mortality statistics (Knittel et al., 2011; Clay et al., 2013), but some studies also rely on hospitalization data (Neidell, 2009; Lagravinese et al., 2013), school absence data (Currie et al., 2009), specific diagnoses (Hammitt and Zhou, 2006), or even self-reported health data (Evans and Smith, 2005; Edwards and Langpap, 2012). Most existing studies use data from industrialized countries, although there has been an upswing in the work on developing countries in recent years (Quah and Boon, 2003; Greenstone and Hanna, 2011; Hanna et al., 2012; Greenstone and Jack, 2013).

By study design or due to data availability, the limitations of existing studies are *(i)* very narrowly defined outcome measures, *(ii)* very narrowly defined geographic foci, or *(iii)* single pollutant or weather measures, which limit the ability to model air pollution comprehensively—in particular the various interactions between the multiple pollutants and weather conditions.

A closely related subfield of the economics literature studies the impact of weather conditions on population health (Deschênes and Moretti, 2009; Deschênes et al., 2009; Deschênes and Greenstone, 2011; Deschênes, 2012; Barreca et al., 2013).[1] The relatively sparse literature on the impact of weather on health is surprising, particularly when considering the heated debates surrounding the causes and consequences of climate change. The famous Stern (2006) report states that the world's average temperature has risen by 0.74 $^\circ$C (1.33 $^\circ$F) over the past 100 years. It

[1] The epidemiological literature on this topic is older and, thus, more diverse (Curriero et al., 2002; Basu and Samet, 2002)

projects this trend will continue into the future. For the US, the predicted temperature change ranges between a 2.2 and 6.1 °C (4 and 11 °F) increase by the end of the $21^{st}$ century (United States Global Change Research Program, 2009). Moreover, climate scientists project a significant increase in the number of hot days with temperatures above 30 °C (86 °F), as well as in the number of heat waves. The Intergovernmental Panel on Climate Change (IPCC) projects *"warmer and fewer cold days and nights"* and states: *"It is very likely that hot extremes, heat waves and heavy precipitation events will continue to become more frequent."* (Intergovernmental Panel on Climate Change (IPCC) (2007), p. 46, 53).

This study aims to comprehensively assess the joint population health effects of pollution and weather for an entire nation over one decade. We base our findings on various high-quality administrative datasets from Germany—the most populous European country and fourth largest industrialized nation in the world— between 1999 and 2008. The study considers a battery of pollution and weather indicators to thoroughly model specific weather conditions and nonlinear associations between weather and pollution. It relies on 11 weather and 11 pollution measures collected by up to 2,358 ambient monitors on a daily basis. This very dense, high-quality network of stations covers the entire 357,121 km$^2$ (138,000 mi$^2$) surface of Germany.

In addition to the rich administrative pollution and weather measures, this paper bases its health outcome findings on two high-quality health register datasets: First, the *German Mortality Census*, containing all deaths occurring on German territory from 1999 to 2008. Second, the *German Hospital Admission Census*, containing more than 170 million hospital admissions from 1999 to 2008. Most previous studies primarily focused on mortality effects. However, solely relying on deaths only allows the researcher to capture a fraction of the total population health effects of weather and pollution. In contrast, relying on both the universe of deaths and hospital admissions should capture most serious population health effects. We aggregate and link the *German Mortality Census*, the *German Hospital Admission Census*, the weather data, and the pollution data at the county-day level over ten years, resulting in 1.5 million county-day observations.

Methodologically, the paper exploits the exogenous nature of day-to-day climatic variation along with rich sets of spatial and temporal fixed effects. As such it identifies the immediate effects of spikes in temperature and pollution levels on hospital admissions and deaths. In other words, the econometric models use the random de-seasonalized within-county variation in extreme temperature and pollution levels over 365 days and 10 years to identify their short-run impact on residents' health.

The setup of the German health care system is particularly well-suited for our research objective since institutional and geographic access barriers to hospitals are very low. Germany has one of the highest densities of hospital beds worldwide, universal health care coverage, and virtually no access barriers for inpatient care (cf. OECD, 2013). In addition, Germany's climatic conditions are ideal to empirically study and identify the effects of extreme temperatures and high pollution levels. Like most countries in the North Temperature Zone, Germany has four seasons, hot summers and cold winters. For example, during the 10 years that we study, daily maximum temperatures range from -14 $^\circ$C (7 $^\circ$F) to 39 $^\circ$C (102 $^\circ$F). Partly due to the fact that weather conditions determine pollution levels, the tempo-spatial pollution variation is likewise very rich. For example, the daily maximum $O_3$ concentration ranges from 1.2 to 192$\mu$g/$m^3$. Lastly, the findings are likely to have external validity for most industrialized countries in the North Temperate Zone, where the majority of the world's population resides.

Overall, this paper makes four main contributions to the literature. First of all, to our knowledge, this paper represents the most comprehensive attempt to assess the population health effects of weather and pollution for an entire country over a long time horizon. The approach provides credible externally valid estimates that can be interpreted as policy-relevant "intention-to-treat" estimates. It is not the purpose of this paper to identify full exposure estimates that we believe can be better and more cleanly assessed in laboratory experiments. Rather, we stress that we use real-world data for an industrialized country over a decade to analyze the real-world impact of extreme temperature and pollution conditions. We show that extreme heat events have a highly significant and largely adverse impact on both hospitalizations and deaths, whereas extreme cold seems to have a negligible real-world impact on population health. Furthermore, particularly when *not* controlling for simultaneous weather and pollution conditions, high levels of outdoor air pollution are associated with significant adverse health effects. Note that these levels of air pollution lie significantly below the regulatory threshold levels in the US. The EU alert thresholds are two to three times lower than their US counterparts. In addition, the actual average $PM_{10}$ and $O_3$ concentrations in the US are 2 to 3 times higher than in Germany; however, the geographical variation in concentrations is also huge in the US (Environmental Protection Agency (EPA), 2013). Thus, our findings have direct implications not only for European countries, but also for the US. For example, the EPA is currently discussing a petition that foresees tighter air pollution regulation on nine "Rust Belt" and "Appalachian" states with high coal emissions. To the degree that regulators can actively bring down spikes in pollution levels, the findings of this paper suggest that a more stringent environmental regulation would be very effective in improving population health.

Second, we present estimates separately for all-cause admissions and deaths along with estimates for five different groups of diseases. This allows researchers and policymakers to better pinpoint vulnerable disease groups, to further investigate the relationships studied, and to develop plans to ameliorate the adverse health effects of temperature and pollution spikes. For example, a *Hot Day*—defined as a day with a maximum temperature of more than 30 $^\circ$C (86 $^\circ$F)—moderately increases almost all cause-specific hospitalizations and deaths by between 2 and 5%. Ongoing heat—e.g., at least four consecutive *Hot Days*—leads to a strong increase in these adverse health effects, particularly for admissions and deaths due to respiratory, infectious, and metabolic reasons.

Third, thanks to the extraordinarily rich palette of weather and pollution measures, the paper also advances the literature methodologically. Although it is generally accepted that short-run climatic variations are exogenous to humans and out of individuals' influence, it is also clear that the climate represents a very complex system with thousands of mostly unobserved factors playing a role. For example, it is well documented that the different available weather measures are non-linear functions of one another (cf. Arya, 1998). In addition, different pollutants form non-linear relationships with each other—some pollutants are necessary chemical input factors for the development of other pollutants (cf. Potter, 2002). To complicate it further, specific weather conditions can be seen as necessary conditions for the formation of high pollution concentrations, whereas there is less evidence that pollution affects weather (cf. Arya, 1998; Seinfeld, 2006; Li et al., 2011). Despite its richness, the economic literature has been relatively silent as how to philosophically and econometrically interpret these relationships and interactions. We hope to provide a first step towards a better understanding by making use of unique and extraordinarily rich data.

To keep the analysis tractable, the empirical portion focuses on two main approaches. Henceforth, *Approach I* is defined as the "Unconditional Approach." This means that the underlying models only consider one single weather or pollution measure as the variable of interest. The identified effect of this single indicator on population health yields the "overall" effect of this indicator on health. For example, consider the effect of a *Hot Day*. In *Approach I*, we do not net out any contemporaneous weather and pollution conditions that typically prevail on a *Hot Day*, e.g., high ozone levels or sunshine. In contrast, *Approach II*—the "Conditional Approach"—always controls for a full set of simultaneous weather and pollution conditions as well as their own and cross-interactions in order to estimate the net effect of a *Hot Day*, i.e., the pure heat effect net of higher pollution levels, less rain, and more sunshine. We show that this distinction makes a crucial difference: When comprehensively considering contemporaneous weather and pollution conditions, the impact of extreme temperature on health shrinks dramatically, by at least 50%. Maybe even more surprisingly, the highly significant and large impact of high concentration levels of single pollutants vanishes almost entirely when fully considering other pollutants and weather conditions that prevail simultaneously. This finding suggests that it is more the overall *combination* of various highly elevated concentrations of pollutants—along with the mostly extreme weather conditions that prevail on high pollution days—that causes humans' physical condition to deteriorate. This is in line with well-established epidemiological and medical experimental evidence suggesting that very high concentrations of single pollutants are required before significant physical health effects can be detected in the lab (cf. Stewart et al., 1970; Anderson et al., 1973; Hackney et al., 1975; Kerr et al., 1979; Horstman et al., 1988; Lippmann, 1989; Jäppinen et al., 1990; Dye et al., 2001). However, for environmental regulators who rely on only unconditional measures from ambient monitors in order to develop alert thresholds and policy action, the unconditional estimates seem to be the relevant ones.

As a final contribution, we provide a first step to better assess and understand the health costs associated with extreme climatic conditions that are very likely to increase in the future due to climate change. Systematically and comprehensively monetizing different subcategories of climate change-related costs is a necessary first step towards cost-benefit analyses. Solid cost-benefit analyses are crucial for a welfare increasing evidence-based climate change management. The most concrete and reliable climate change prediction of the Intergovernmental Panel on Climate Change (IPCC) (2007) is an increase in the number of extreme heat events. Thus, we attempt to monetize the health loss associated with one additional *Hot Day* for an entire nation. One main conclusion from this exercise is that two factors crucially drive the estimates: *(i)* the choice between the *Unconditional* and the *Conditional Approach*, and *(ii)* whether one considers "harvesting" or not. In line with the literature, we find strong evidence in line with the harvesting hypothesis, according to which heat mostly adversely affects humans in bad health who, in the absence of heat, would have likely died shortly after. Empirical tests demonstrate that heat does not lead to a *permanent* increase in hospitalizations and deaths. Depending on the underlying assumptions, the last part of the paper estimates that one additional *Hot Day* triggers a monetized health loss of between €6m and €43m for Germany or between €0.07 ($0.10) and €0.52 ($0.68) per German resident.

The next section describes the datasets used as well as the rich tempo-spatial variations in the main variables of interest. Section 3 describes the estimation strategy and discusses the identification of the effects. Section 4 contains the empirical findings. In Section 5, we monetize the health loss of one additional *Hot Day*. Section 6 concludes.

## 2 Datasets, Main Variables, and Identifying Variation

### 2.1 *Hospital Admission Census*: The Universe of all German Hospital Admissions 1999-2008

The first dataset used is the *Hospital Admission Census*. Access is provided by the German Federal Statistical Office. It comprises all German hospital admissions from 1999 to 2008. Germany has about 82 million inhabitants and registers about 17 million hospital admission per year. We observe every single hospital admission from 1999 to 2008, i.e., a total of more than 170 million hospitalizations.[2] To obtain our working dataset, we aggregate the individual-level data at the day-county level and normalize admissions per 100,000 people using official population counts (see Appendix F).[3]

As seen in Appendix A, along with other admission characteristics, the *Hospital Admission Census* provides information on the age and gender of the patient, the day of admission, the length of stay, the county of residence as well as the primary diagnosis in form of the ICD-10 code (10$^{\text{th}}$ revision of the *International Statistical Classification of Diseases and Related Health Problems*).

#### Construction of Main Dependent Variables

Using the information on the primary diagnosis, we generate a series of dependent variables. The dependent variables represent different groups of diagnoses, generated by extracting the letter and digits of the ICD-10 code, e.g., J00-J99 refers to "diseases of the respiratory system." In some cases, the second and third ICD-10 digits are helpful to identify more specific conditions. In addition to *all-cause hospitalizations*, which is simply the sum of all admissions, we examine five

[2] By law, German hospitals are required to submit depersonalized information on every single hospital admission. This excludes military hospitals and hospitals in prisons. The 16 German states collect the information and the German Federal Statistical Office (*Statistische Ämter des Bundes und der Länder*) provides restricted data access for researchers.

[3] The remote access servers of the *Statistische Ämter des Bundes und der Länder* only provide a memory of 18 gigabytes per computer. The individual admission data is provided in files by calendar years. The memory capacities only allow to merge and analyze one calendar year of hospitalizations on the individual admission level. Therefore, one has to restrict the working dataset to patients who were admitted after January 1$^{\text{st}}$ of a given year. In other words, one has to delete all admissions that led to hospital stays over New Year. This is because we first aggregate admissions at the day-county level and then merge the files by calendar years, resulting in duplicate observations for counties and days with admissions in $t_0$ and discharges in $t_1$. In a robustness checks, we run the analysis using only one calendar year, but include stays over New Year's Day. The results are robust to excluding people who stay in hospitals over the change in years and are available upon request.

specific subgroups: *(i) cardiovascular hospitalizations* (I00-I99), *(ii) respiratory hospitalizations* (J00-J99), *(iii) infectious hospitalizations* (A00-B99), *(iv) metabolic hospitalizations* (E00-E89), and *(v) neoplastic hospitalizations* (C00-D49).

We also exploit the death and length of stay information. Following up on the example from above, *cardiovascular death* identifies people who died after they were admitted to a hospital due to a cardiovascular disease. *Cardiovascular hospital days* includes the number of nights that a patient spent in a hospital after a cardiovascular admission.

After having summed the total of admissions—as well as the cause-specific admissions—at the day-county level, we normalize the dependent variables per 100,000 population using official population data at the year-county level (Federal Institute for Research on Building, Urban Affairs and Spatial Development, 2012). Appendix A displays the summary statistics of all normalized dependent hospital admission variables.[4] For example, on a given day, we observe 58 hospital admissions per 100,000 residents. This admission rate varies substantially over days and across counties; the standard deviation is 26. On average, a day triggers 489 hospital days, i.e., the 58 admissions have an average length of stay of 8.4 days. The largest single group of diseases is *cardiovascular hospitalizations*. Nine cardiovascular admissions per 100,000 pop. represent 16% of all admissions.

### 2.2 *Mortality Census*: The Universe of all German Deaths 1999-2008

The second dataset employed is the *Mortality Census* which is also provided by the German Federal Statistical Office. The *Mortality Census* includes every death that occurred on German territory. Per year, one observes approximately 800,000 deaths, i.e, about 8 million deaths from 1999 to 2008. To obtain the working dataset, we aggregate the individual-level data at the day-county level and generate the mortality rate per 100,000 population.

Appendix B shows all raw measures included in the *Mortality Census*. It contains information on age, gender, day of death, county of residence as well as the primary cause of death in ICD-10 form.

[4] Note that the German data protection laws prohibit us from reporting min. and max. values.

### Construction of Main Dependent Variables

Analogously to the *Hospital Admission Census*, we generate dependent variables that indicate the all-cause mortality rate, as well as the cause-specific mortality rates for five specific categories: the *(i) cardiovascular mortality rate*, *(ii) respiratory mortality rate*, *(iii) infectious mortality rate*, *(iv) metabolic mortality rate*, and *(v) neoplastic mortality rate*. The total daily mortality rate is 2.99 deaths per 100,000 pop.—1.38 or almost 50% of which are caused by cardiovascular health issues. The summary statistics of the all-cause—as well as cause-specific—mortality rates are displayed in Appendix B.

## 2.3 Official Daily Weather Data from 1,044 stations 1999-2008

The weather data is provided by the German Meteorological Service (*Deutscher Wetterdienst (DWD)*), a publicly funded federal institution. Weather measures are collected from 1999 to 2008 from up to 1,044 meteorological monitors which are distributed all over Germany. Figure 1 shows the distribution of all ambient monitors along with county borders. It is easy to see that the German weather station network is very dense.

**[Insert Figure 1 about here]**

The paper uses official measurement data from all existing weather stations in a given year. As described in Section 2.5, we interpolate the point measures into county space on a daily basis using Inverse Distance Weighting (IDW).

### Weather Variation Across Space and Time

Summary statistics of all raw weather measures at the day-county level are given in Panel A of Table C1 in Appendix C. The mean daily air temperature is 9.6 $^\circ$C (49.2 $^\circ$F), averaged over the whole time period and all counties. Note the extremely rich variation in the average daily temperature: it ranges from -19 $^\circ$C (-2.2 $^\circ$F) to 30.6 $^\circ$C (87.1 $^\circ$F). Equally rich is the variation in the minimum and maximum temperature, hours of sunshine and other weather measures.

**[Insert Figure 2 about here]**

Figure 2a shows a boxplot of the mean temperature over the twelve months of a year (averaged over all ten years). The graph illustrates the large cross-county as well as cross-seasonal variation in weather. One observes a clear increase in average temperatures and sunshine duration during the summer months. Figure 2b shows the daily cross-county temperature, sunshine, and precipitation variation over ten years. One observes the typical seasonal trends along with a lot of spikes in the high-frequency data. The empirical models will exploit the rich positive and negative weather shocks across space and over time. Deviations in daily weather variations are plausibly exogenous to individuals' health.

Figure 10 in Appendix C displays a scatter matrix which shows, illustratively, the associations between some raw weather measures. Not surprisingly, one finds a strong positive association between the hours of sunshine and the temperature, as well as a strong negative association between the hours of sunshine and the precipitation level.

### Construction of Extreme Temperature Indicators & Identifying Variation

**Construction of Extreme Temperature Indicators.** At the beginning of the Results section, this study employs the raw continuous weather measures and provides nonparametric evidence on the (non-linear) relationship between temperature and health effects. As a next step, we employ semiparametric variants that net out seasonal effects, but let a series of temperature regressors flot flexibly. In the main parametric models, however, we mostly employ a single binary indicator to measure extreme heat and cold for the following reasons:

*(i)* The binary measures generated refer to the official definitions of *Hot* and *Cold Days*, e.g., the German Meteorological Service defines a *Hot Day* as a day with a maximum temperature above 30 ° C (86 ° F). In addition, the previous literature has employed these binary measures, which facilitates the comparison of results (cf. Deschênes and Moretti, 2009; Barreca et al., 2013).[5]

*(ii)* Defining a binary indicator to measure *Hot* and *Cold Days* greatly simplifies the empirical analysis, provides the reader with a better intuition, and makes it easier to follow the thought experiment wherein we ask, "What are the health effects of one additional *Hot Day*?"

*(iii)* As we will demonstrate in the Results section, there is empirical evidence that most adverse health effects kick in when temperatures exceed 30 ° C (86 ° F). Thus we define the binary measures:

- ***Hot Day*** = 1 if the max. temperature >30 ° C (86 ° F), 0 else.
- ***Heat Wave Day*** = 1 if *Hot Day*=1 and the 3 previous days were also *Hot Days*, 0 else.
- ***Cold Day***= 1 if the min. temperature < -10 ° C (14 ° F).

[5] To be precise, the US studies by Deschênes and Moretti (2009) and Barreca et al. (2013) define a *Hot Day* as a day with the max. temperature >90 ° F (32.2 ° C).

- ***Cold Wave Day***= 1 if *Cold Day*=1 and the 3 previous days were also *Cold Days*, 0 else.

**Identifying Variation.** Panel B of Table C1 in Appendix C shows the descriptive statistics for the generated extreme temperature indicators. As seen, 1.97% of all days are *Hot Days*. This translates into roughly seven *Hot Days* per year. However, between 1999 and 2008, the number of *Hot Days* varied between 4 (1999, 2004, 2007) and 18 (2003). Note that the variation in *Hot Days* between counties is even larger. The number of *Hot Days* varies between 0 and 40 per year, depending on county (Figure 3a). In one empirical specification, we aggregate the data at the year-county level and exploit the variation in the annual number of *Hot Days*.

On average, there is about one *Heat Wave Day* per year (see Panel B of Table C1). Between 1999 and 2008, the number of *Heat Wave Days* varied between 0.03 (2005) and more than 6 (2003, not shown in Table C1).[6]

Figure 3a plots the distributions of *(i)* the annual mean of the maximum daily county-level temperatures, and *(ii)* the annual number of *Hot Days* per county. This is to show that *(a)* the annual mean maximum temperatures follow a normal distribution with the mass point around 14 $^\circ$C (57 $^\circ$F), and *(b)* that the annual number of *Hot Days* is skewed to the right and exhibits substantial variation with many counties showing more than 10 *Hot Days* per year. Overall, Figure 3a illustrates that the identifying variation stems from the majority of counties and not just a small subset of "hot" counties. Thus extrapolation and out-of-sample predictions are largely avoided.

[**Insert Figure 3 about here**]

Turning to the other temperature extreme, 1.24% or about 20,000 of all county-day observations are *Cold Days* with minimum temperatures below -10 $^\circ$C (14 $^\circ$F). This translates into 4.4 days per year, but the variation ranges from an average of 1 *Cold Day* in 2008 to 10 *Cold Days* in 2003. The annual county-level variation in *Cold Days* ranges from 0 to 41 (see Figure 3b). Rarer are *Cold Wave Days* with more than three consecutive *Cold Days*—between one and zero occur per year. However, over 10 years, we still count 2,870 county-level *Cold Wave Days*.

Figure 3b shows the distributions of *(i)* the annual mean of the minimum daily county-level temperature, and *(ii)* the annual number of *Cold Days* per county. As in the extreme heat case, the minimum temperature distribution is about normal (mean 5.5 $^\circ$C (42 $^\circ$F)), while the county-level *Cold Day* distribution is skewed to the right (mean 4.4 days). Again, the empirical models

[6] Note that 6 heat wave days could be the result of a very long heat wave, lasting 9 days, or 6 short heat waves of 4 consecutive *Hot Days*, or a combination of the two.

largely avoid out-of-sample predictions since the identifying variation occurs in the large majority of German counties.

### 2.4 Official Daily Pollution Data from 1,314 stations 1999-2008

The pollution data are provided by the German Federal Environmental Office (*Umweltbundesamt (UBA)*), a publicly funded federal agency. From 1999 to 2008, pollution measures are collected from up to 1,314 ambient monitors (Figure 1). As with the weather measures and as described in Section 2.5, we interpolate the monitor point measures into the county space. Panel A of Table D1 in Appendix D shows all raw pollution measures on a daily county-level basis.

Appendix D also discusses in detail the chemical composition of the five pollutants investigated, as well as their health hazards. Moreover, Appendix D describes and graphically illustrates the tempo-spatial variation of the pollutants and their association with weather conditions: All pollutants have in common that they *(i)* exhibit some seasonal pattern, *(ii)* exhibit strong (non-)linear associations with the weather indicators—in particular the temperature—and most importantly for identification purposes: *(iii)* exhibit strong daily variation across counties and over time.

#### Construction of Non-Compliance Pollution Indicators, Comparison to US Thresholds, & Identifying Variation

**Construction of Non-Compliance Pollution Indicators.** One objective of this paper is to assess the effects of high pollution concentrations on human health. As in the temperature case, we first demonstrate the nonparametric nonlinear relationship between pollution concentrations and health graphically. Then we employ nonparametric models that net out seasonal impacts but explore the pollution-health relationship via flexibly varying pollution concentration regressors. Finally, the main models make use of the official EU alert thresholds to assess the health impact of crossing these thresholds.[7] Appendix D discusses the different thresholds in detail and also the policy action required when counties violate these thresholds. Henceforth, we call a day during which the pollution concentration exceeds its EU threshold a "non-compliance" day. Accordingly, the following binary indicator variables are generated (European Environment Agency, 2012).[8] The descriptives of these indicators are displayed in Panel B of Table D1.

[7] This is not always exactly feasible since we rely on day-county level averages, whereas some EU thresholds rely on hourly averages.

[8] Note that CO and $SO_2$ are omitted. The maximum county-level CO concentration that we observe is 2.8 ppm, which is significantly below the EU threshold and the WHO 8-hour threshold of 8.7 ppm. All $SO_2$ values also lie significantly below the maximum daily EU threshold of 125 $\mu g/m^3$. Interestingly we observe the same pattern for the US (Environmental Protection Agency (EPA), 2013).

- ***$O_3$* non-compliance day** = 1 if the max. $O_3$ level >120 $\mu g/m^3$, 0 else.
- ***$NO_2$* non-compliance day** = 1 if the average $NO_2$ level >40 $\mu g/m^3$, 0 else.
- ***$PM_{10}$* non-compliance day** = 1 if the average $PM_{10}$ level >50 $\mu g/m^3$, 0 else.

According to these definitions, 12% of all 1.5 million county-day observations are *$NO_2$ and $PM_{10}$ non-compliance days*. This translates into 44 days per year. Thirty-four days per year are *$O_3$ non-compliance days*.

**Comparison to US Thresholds.** In principle, the pollution regulation in the US is similar to the one in the EU: the US Environmental Protection Agency (EPA) implements pollution concentration thresholds and requires all US states to comply. However, the EPA thresholds are significantly less strict than the EU ones: The $PM_{10}$ threshold is a 24 hour average concentration of 150 $\mu g/m^3$. The $O_3$ threshold is an 8 hour average concentration of 159 $\mu g/m^3$. And the $NO_2$ threshold is an annual average concentration of 107 $\mu g/m^3$ or a maximum hourly concentration of 203 $\mu g/m^3$ (Environmental Protection Agency (EPA), 2013).[9] Thus, the threshold levels for $NO_2$ and $PM_{10}$ are 2 to 3 times larger in the US, which should be kept in mind when comparing the results of this study to related US studies. In Germany, from 1999 to 2008, the US regulatory thresholds for $PM_{10}$, $O_3$ and $NO_2$ were *never* exceeded (see Table D1). The *measured* pollution concentrations in the US are likewise 2 to 3 times higher than in Germany—at least for $PM_{10}$ and $O_3$, while $NO_2$ concentrations are very similar (Environmental Protection Agency (EPA), 2013). However, note that the variation in concentrations across US regions is very large and the distribution overlaps with the distribution of concentration levels in Germany.

**Identifying Variation.** Figure 4 illustrates several crucial empirical facts about the identifying pollution variation using the example of ozone ($O_3$): *(a)* Figure 4a in the top left corner illustrates that the annual number of *$O_3$ non-compliance days* is highly correlated with the annual maximum ozone concentration. This shows that the binary non-compliance indicator represents and captures high ozone concentrations well.

**[Insert Figure 4 about here]**

*(b)* Figure 4b in the top right corner shows the pollution variation with respect to the daily maximum ozone concentration per county and year (black), as well as the number of annual non-compliance days per county and year (red). Both distributions are roughly normal and have large

[9] The original scales for $NO_2$ and $O_3$ are expressed in "parts per million (ppm)" and have to be converted to "micrograms per cubic meter of air ($\mu g/m^3$)". The annual threshold for $NO_2$ is 0.053 ppm and the hourly maximum 0.1 ppm. For $O_3$, the "annual fourth-highest daily maximum 8 hours concentration, averaged over 3 years," must not exceed 0.075 ppm.

supports. This shows that the identifying variation is based on a broad set of counties and not just a small subset of non-representative high ozone-level counties with permanently high levels of ozone. It is worthwhile to note that *every* single German county had non-compliance days between 1999 and 2008. In fact, the number of non-compliance days ranges between 10 and 553 across the counties and over the ten years.

*(c)* Figure 4c in the bottom left corner plots the annual number of non-compliance days along with the GDP growth rate, while Figure 4d plots the annual number of non-compliance days along with the annual maximum temperature. The graphs illustrate that there is not much correlation between economic activity and high ozone concentrations, but that there is a strong correlation between high temperatures and high ozone levels at the county level. This relationship is also illustrated in Figure 15 in Appendix D. The relationship derives from the chemical process that leads to high ozone levels. High temperatures and sunshine are important input factors for the photochemical oxidation process between CO and $NO_x$ and thus for the development of ozone (cf. Arya, 1998; European Environment Agency, 2013). It that sense one can think of extreme temperatures causally triggering extreme ozone levels. Since high ambient temperatures are exogenous to individuals, so are high ozone levels. The empirical analysis exploits both.

The equivalent graphs to Figure 4 for nitrogen dioxide ($NO_2$) and particular matter ($PM_{10}$) can be found in Appendix D (Figures 13 and 16). The conclusions drawn for $O_3$ also hold for $NO_2$ and $PM_{10}$. For example, from 1999 to 2008, even the German county with the least $NO_2$ or $PM_{10}$ pollution experienced 4 ($NO_2$) and 8 ($PM_{10}$) *non-compliance days*, respectively. High $NO_2$ and $PM_{10}$ concentrations are also triggered by high temperatures.

### 2.5 Interpolation of Weather and Pollution Measures

To obtain the working datasets, we *(i)* had to interpolate the point measures of the weather and pollution monitors into the county space, *(ii)* aggregate and normalize all information at the (daily) county level, and *(iii)* merge the register datasets with the pollution, weather, and the socioeconomic dataset (see Appendix F) at the day-county level. Assuming that the number of counties is time-invariant and 400, we should obtain $400 \times 365 \times 10 = 1,460,000$ rows, each representing one county on a given day.

**Interpolation of Weather and Pollution Measures**

Hanigan et al. (2006) discuss and compare different approaches of how to calculate population exposure estimates of daily weather and pollution conditions from monitors. The approach chosen here makes use of the geographical centroid of each county: one calculates the weather and pollution conditions for every county and day as the inverse distance weighted average of all ambient monitors within a radius of 60 km (37.5 miles) of the county centroid. Thus, denoting by $\delta_{ij}$ the distance between a location $i$ (a county centroid) and a monitor $j$, one can define the weighting scheme as:

$$w_{ijd} = \begin{cases} \frac{1}{\delta_{ij}} & \text{if } i \neq j \text{ and } \delta_{ij} < 60 \\ 1 & \text{if } \delta_{iM_{id}} > 60 \text{ and } j = M_{id} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $M_{id}$ denotes the nearest station outside location $i$. Thus, whenever there are no stations within a radius of 60 kilometers, the measure from the nearest station outside this radius is used. Henceforth, we call this interpolation approach simply Inverse Distance Weighting (IDW).

# 3 Empirical Approach and Identification

## 3.1 Econometric Approach

As a first step, we investigate the temperature and pollution-health relationship nonparametrically by plotting scatterplots along with Kernel-weighted local polynomial smooth plots. Next, we run the following model by OLS:

$$\begin{aligned} Y_{cd} = \alpha + \sum_{h=20}^{36} \beta_h \, MaxTemp_{cd}^h \qquad (2) \\ + \sum_{j=2}^{468} \nu_j \, county_j + \sum_{k=2}^{52} \zeta_k \, week_k + \sum_{m=\text{Feb1999}}^{\text{Dec 2008}} \sigma_m \, month_m + \theta \, X_{ct} + \epsilon_{cd} \end{aligned}$$

where, depending on the specification, $Y_{cd}$ either denotes the hospital admission rate or the mortality rate per 100,000 population in county $c$ on day $d$. $MaxTemp_{cd}^h$ are a series of temperature regressors of interest that equal 1 if the maximum daily temperature in the county *exceeds* $h$ degrees and is zero otherwise. We plot the 17 temperature coefficient estimates to semiparametrically evaluate the temperature-health relationship, net of seasonal influences. This means that we net out county fixed effects, $\sum_{j=2}^{468} \nu_j \, county_j$ , calendar week fixed effects, $\sum_{k=2}^{52} \zeta_k \, week_k$ , year-month fixed effects, $\sum_{m=\text{Feb 1999}}^{\text{Dec 2008}} \sigma_m \, month_m$, as well as yearly county-level covariates, $X_{ct}$. The latter vector contains demographics, the share of private hospitals, the bed density, or the county-level GDP per capita (see Appendix F).

Finally, our basic fully parametric approach is based on the following model:

$$
\begin{aligned}
Y_{cd} = {} & \alpha + \beta \, ET_{cd} + \gamma \, PNC_{cd} + \phi \, W_{cd} + \rho \, P_{cd} + \psi \, W_{cd} \times P_{cd} \\
& + \sum_{j=2}^{468} \nu_j \, county_j + \sum_{k=2}^{52} \zeta_k \, week_k + \sum_{m=\text{Feb 1999}}^{\text{Dec 2008}} \sigma_m \, month_m + \theta \, X_{ct} + \epsilon_{cd}
\end{aligned}
\tag{3}
$$

As mentioned earlier, the parametric analysis is implemented using two main approaches: the "*Unconditional*" and the "*Conditional*" Weather and Pollution *Approach*. Both approaches routinely control for all covariates that appear in the second row in equation (3), i.e., county fixed effects, calendar week fixed effects, year-month fixed effects, as well as yearly county-level covariates.

Approach I is the *Unconditional Approach* that does not net out contemporaneous weather and pollution conditions. In addition to the sets of covariates listed above, only one weather or pollution variable of interest is added to the model. Depending on the exact specification, for example, this could be either the dummy *Hot Day*, the dummy *Cold Wave Day*, or a pollution non-compliance dummy like *$NO_2$ non-compliance day* (see Section 3.3 and 2.4). One can think of this approach as a reduced form "intention-to-treat" approach where the main regressor of interest absorbs all weather and pollution conditions that are correlated with the exogenous weather or pollution indicator.

Approach II is the *Conditional Approach*, net of contemporaneous weather and pollution conditions. A saturated model is estimated that includes—in addition to the weather or pollution regressor of interest—all covariates listed in the first row of equation (3), i.e., *E*xtreme *T*emperature dummies in the vector $ET_{cd}$, a set of *P*ollution *N*on-*C*ompliance dummies $PNC_{cd}$, the 7 "raw" continuous weather measures, as well as 15 own interactions of these measures at the day-county level, represented by the vector $W_{cd}$ (see Tables C1 and D1).[10] The *Conditional Approach* also makes use of $P_{cd}$ which contains 5 continuous pollution measures—$O_3$, $NO_2$, $SO_2$, $NO_2$, $PM_{10}$—their quadratic and cubic terms, as well as the 10 most relevant interactions between these 5

[10] Obviously, when estimating the effect of a *Hot Day* this vector does *not* additionally include the temperature, but such measures as precipitation, sunshine, or air pressure (see Table C1 in Appendix C).

pollutants (cf. Figure 17). Finally, Approach II also considers cross-interaction effects between $W_{cd}$ and $P_{cd}$. In total, in addition to the binary variable of interest, the *Conditional*,saturated, *Approach* includes more that 80 additional weather and pollution control variables—in levels, quadratic, cubic and interacted terms. As such, the *Conditional Approach* disentangles the single pollutant or extreme temperature effect from simultaneous weather and pollution conditions.

## 3.2 Identification of Population Health Effects Using Hospitalization and Mortality Censuses

First of all, one needs to consider that we "only" observe the universe of deaths and inpatient treatments, i.e., hospital admissions that require the patient to stay over night. This excludes mild conditions that are treated in outpatient settings. Since this paper intends to assess the population health effects of weather and pollution, the underlying assumption here is that adverse health effects not requiring an overnight stay in a hospital are negligible relative to inpatient treatments and mortality effects. This assumption essentially means that we obtain a lower bound total population health effect triggered by weather and pollution conditions.

Second, note that we interpret the estimates strictly as contemporaneous short-run effects on population health. While this approach has several methodological advantages—one of them is the immediate and obvious dose-response relationship that substantially ameliorates concerns about confounding impacts of third unobserved factors—one has to keep in mind that this approach abstracts away from long-term effects on health. The main neglected long-term effect is certainly the adverse effect of in-utero and early childhood exposure to adverse environmental conditions (cf. van den Berg, 2006; Currie and Almond, 2011; Zivin and Neidell, 2013). For example, Currie et al. (2013) estimate that the overall discounted long-term societal costs of being born with low-birth weight are at least \$100,000. While the paper abstains from comment on any long-term effects of climatic conditions on health, please note that the approach considers the immediate effect on hospitalizations and deaths of newborns and children.

Third, while we are able to observe every single hospital admission, data protection laws prohibit us from analyzing panel data. This means that we are unable to observe hospital readmissions. According to representative Socio-Economic Panel Study (SOEP) data, about 13% of all Germans were admitted to a hospital in 2010. About 2% (15% conditional on an admission), had more than one hospital stay in 2010 (Wagner et al., 2007). Not being able to identify readmissions would be particularly worrisome if we were interested in treatments of chronic diseases such as diabetes where patients are obliged to return to the hospital in regular intervals.[11]

Fourth, we implicitly assume that all severe health effects triggered by weather and pollution eventually lead to a hospital admission or death. We believe that this is a reasonable assumption. German geography, combined with the institutional setting of the German health care system, supports the assumption. First of all, the German population density is relatively high. Germany has 82 million residents living in an area roughly the size of the US state Montana. The average German population density is about seven times as high as the US population density (231 vs. 32 people per $km^2$) (U.S. Census Bureau, 2012; German Federal Statistical Office, 2012). The hospital bed density is also much higher. Germany has a total of 2,045 hospitals while Montana has only 70 hospitals (German Federal Statistical Office, 2013b). Per 100,000 population, Germany's health care infrastructure offers 824 hospital beds, while the US has only 304 (OECD, 2012). This illustrates that geographic hospital access barriers, such as travel distances, are low in Germany and significantly lower than in the US.

To date, Germany has 402 counties or county-equivalents ("urban municipalities"). The average population is about 190,000 but varies from 35,000 to 3.5 million for Berlin (see Table F1 in Appendix F). As compared to the US, the area size of German counties is smaller and the population density is higher. The US has 3,144 counties but is 27 times the size of Germany. The average US county population is about 100,000 but variation is much larger than in Germany and ranges from 82 inhabitants in the smallest to 10 million inhabitants in the largest US county, Los Angeles County (United States Census Bureau, 2013). Hence, the German counties are more homogeneous. Still it is fair to say that US and German counties are comparable both in terms of their administrative function in the two federalist states, as well as their overall structure.

Lastly, the uninsurance rate in Germany is below 0.5%. The public health care system covers 90% of the population and copayment rates in the public scheme are uniform and low.[12] The overwhelming majority of hospitals can be accessed independently of insurance status and provider networks are almost unknown in Germany. Thus, insurance barriers to hospital access are also low in Germany, and certainly lower than in the US.

Given these very low geographic and institutional access barriers, it is reasonable to assume that severe health conditions ultimately lead to hospitalizations or death

[11] Note that, using the age, gender and county-level information, we could apply propensity score matching methods to probabilistically identify readmissions.

[12] If total out-of-pocket expenditures do not exceed 2% of the individual's income (1% for people with chronic conditions), the daily copayment for inpatient stays is €10 in the public system.

### 3.3 County-Level Aggregation and Measurement Errors in Pollution and Weather Measures

Every empirical study has to deal with the issue of measurement errors. It is known that classical measurement error attenuates parameter estimates. In case of non-classical measurement error, the direction of the bias is unclear. Moreoever, measurement error in the dependent variables inflates standard errors (Chen et al., 2011). Since this study makes use of several rich high-frequency administrative register datasets it *(i)* certainly does not have a power issue, and *(ii)* the overall data quality is very high.

However, particularly when it comes to the mapping of monitor point measures into space, it is obvious that one has to deal with measurement errors.[13] To assess the measurement error that is introduced via the IDW method, we perform the following (indirect) test: For each weather and pollution *monitor* (*not* county centroid), one calculates the IDW value using the weighting scheme in equation (1). The crucial point is that the weighting scheme attaches weight 0 to the own station.[14] Thus, for each ambient monitor and all weather and pollution measures from that monitor $Z_d$, we calculate a cross-validated $\tilde{Z}_d = Z_d \Omega_d$; where $\Omega_d$ is the symmetric matrix of weights for day $d$ with elements $\omega_{ijd} = w_{ijd}/\sum_k w_{ikd}$. In other words, we predict the values of each monitor using all surrounding monitors and the IDW interpolation method. Then, we assess the accuracy of the IDW interpolation by calculating Pearson's correlation coefficient for the variables $Z$ and $\tilde{Z}$. The results of this exercise are in column (1) of Table G1 in Appendix G. Column (2) compares the IDW method to a simple nearest neighbor (NN) interpolation that just correlates the values of each station with its nearest neighbor.

Table G1 illustrates that *(a)* the IDW method dominates the simpler NN weighting scheme: only for air pressure does the NN method deliver better accuracy. Besides, it becomes clear that *(b)* our IDW interpolation algorithm delivers a very acceptable accuracy with correlation coefficients ranging from 0.413 for the max. CO concentration to 0.981 for the mean temperatures. Overall, the correlation values for pollution range between 0.4 and 0.8 while the weather measures mostly deliver even better results. Note that this paper particularly relies on minimum, mean, and maximum temperature measurements, all of which deliver excellent accuracy results with correlation values ranging above 0.95 (column (1) of Table G1). This means that we are able to

[13] Currie et al. (2013) show that if one uses variation in the toxic exposure of plants as source of exogenous pollution variation, at least in the US, the measurable impact of the emitted pollutants mostly lies within the radius of one mile.

[14] When using the county centroid in the IDW interpolation of point measures into county space, the closest monitor obviously gets the largest weight.

predict 95% of the variation of the temperature measured by monitor X using our IDW method and all surrounding monitors.[15] Note that the results are robust to considering individual years instead of the entire pooled sample (results available upon request).

Table G2 in Appendix G show results of a similar test for the generated extreme weather indicators and confirms the results of Table G1. Basically, one finds that the overall share of correctly predicted heat and cold indicator values is above 99%, as is the share of correctly predicted zeros. Since there is only a small percentage of extreme temperature events, "false positives" have a larger impact on estimates than "false negatives". Thus, it is reassuring to see that *(i)* IDW clearly outperforms NN in that respect, and *(ii)* the share of false positives is low and less than 20% in the case of heat.

Finally, we calculate the Reliability Ratio (RR) $\alpha$ that indicates the magnitude of measurement errors and thus the attenuation bias (Hyslop and Imbens, 2001):

$$\alpha = \frac{\text{Cov}\left(Z, \tilde{Z}\right)}{\text{Var}\left(\tilde{Z}\right)} \tag{4}$$

The RR is relatively high and lies around 0.8 for the most important indicators.

As a last conceptual point, please note that the issue of introducing measurement error when extrapolating point measures into space is methodologically not fundamentally different from the issue of unknown individual exposure to weather and pollution conditions. We approximate the individual level exposure to weather and pollution on a given day by taking inverse distance weighted averages of the daily measures of the next monitors. Even if we knew the exact ambient weather and pollution conditions at the exact locations of residence of all German residents, we would still *(i)* have to take daily averages in ambient conditions, *(ii)* lack knowledge about the exact length, place, and time of the day spent outdoors by the individuals, and thus *(iii)* deal with exposure-related measurement error of unknown form.

As time, empirical methods, and data collection advance, researchers will have access to increasingly more and better data that reduce measurement errors. Considering our extremely precise coefficient estimates stemming from various high-quality register datasets and comparing

[15] One concern with this interpolation test is that a seemingly high degree of accuracy might be driven by time trends and seasonal variation in the variables. Thus, we calculate alternative accuracy correlation measures that are based on transformed versions of $Z$ which had first been nonparametrically adjusted for individual day effects. As seen, the correlation coefficients in columns (3) and (4) of Table G1 drop somewhat, but still show that there is a considerable correlation between imputed and actual values. For the temperature measures, the time trend and season-adjusted correlation values all lie around 0.7. It should also be noted that this method of controlling for day time effects is very conservative in the sense that it is likely to remove "too much" variation from the data since one cannot disentangle the "true" correlation between monitors and climatic measures from day effects. By removing the daily mean one obviously also removes part the non time-trend correlation.

this data quality to the data quality of self-reported survey data often used in empirical studies, we believe and argue that the issue of measurement error is of secondary importance for the general findings of this paper.

### 3.4 Identification of Exogenous Pollution and Weather Effects Using Daily Spikes in High Pollution and Extreme Temperatures

From an identification point of view, the appealing aspect of using weather and pollution variation to estimate their impact on health is that weather and pollution is very likely to be orthogonal to the error term in equation (3) above. More precisely, it is very plausible that pollution and weather variation at the day-county level is exogenous to the outcomes of any one individual. Remember that the parametric models net out a rich array of seasonal and time effects and solely rely on high-frequency, daily within county variation. Positive and negative pollution and temperature shocks are then linked to contemporaneous health effects at the day-county level.

One could still identify three identification concerns: *(i)* based on (un)observables, people may self-select into living in specific regions, *(ii)* pollution *levels* may be correlated with economic activities which, in turn, may affect health outcomes, *(iii)* individual-level exposure to weather and pollution conditions is unknown and adaption behavior may bias the "true" causal effect downward.

A few recent papers address some of these concerns by using variation in traffic as an instrument for CO, $PM_{10}$, and $O_3$ exposure (Knittel et al., 2011; Moretti and Neidell, 2011; Schlenker and Walker, 2011). While these approaches as stimulating and worthwhile to pursue, this paper abstains from instrumenting pollution levels with traffic activity.

Instead, this paper addresses the potential concerns in the following ways. First of all, with respect to *(i)*: It is of course true that people with specific characteristics may self-select into specific regions. This is of particular concern for studies that rely on small geographic regions— one may question the external validity of the findings. One particular strength of the approach used in this paper is that it relies on the universe all hospital admissions and deaths from the fourth largest industrial nation in the world over one decade. To the extent that one is interested in the real-world effects of weather and pollution on population health in a given geographic area, one should consider and *include* sorting into regions; the identified parameters then represent the effects on population health once geographic preferences are accounted for. In the case of Germany, it should be added that (intergenerational) geographic mobility is historically very low.

Using the SOEP we find that, in a given year, only about 1% of all SOEP respondents move, which also includes within-county moving (Wagner et al., 2007; SOEP, 2012).

Second, as far as *(ii)* is concerned: it is obvious that the *level* of regional economic activity and the regional pollution *level* may be correlated. This is particularly worrisome when pollution and health outcome data are linked on a highly aggregated level, e.g., when the unit of observation is the year or month and studies do not or cannot account for year and region fixed effects.

However, as a first argument, recall that we rely on high-frequency data, recorded on a *daily* county-level basis. We do not only consider county fixed effects, but also week fixed effects and month-year fixed effects. Moreover, we make use of binary indicators that indicate *changes* in high pollution concentrations in these fixed effects models. Econometrically, this means that we exclusively focus on daily county-level increases or decreases in high pollution concentrations, i.e., EU non-compliance days and alternative high pollution thresholds. Economic activity, in contrast, does not fluctuate strongly at the day-county level.

Moreover, a robustness check relates changes in high pollution concentrations to hospitalizations that may stem from an increased economic activity: treatments due to physical injuries caused by accidents. We do not find any evidence that there is a meaningful relationship between these two factors. For example, in the *Unconditional Model*, an ozone non-compliance day is associated with a 0.024 ppt. (0.5%) lower standardized accident rate at the daily county-level. However, this association is clearly not statistically significant with a p-value of 0.52 (detailed results available upon request).

In addition, Figures 4c, 13c, and 16c do not suggest a significant relationship between economic activity and high levels of pollution. Looking at annual county-level variation in GDP growth and pollution concentrations, it becomes clear that changes in extreme pollution are not primarily driven by economic activity but rather by high temperatures. On the annual county level, only the maximum $NO_2$ concentration is positively correlated with the GDP growth per capita. However, even this correlation is rather small. An increase in the growth rate by 1 ppt. increases the average maximum $NO_2$ level by 0.25 $\mu g/m^3$ or 0.5%. High $O_3$ and $PM_{10}$ concentrations are in fact *negatively* correlated with GDP growth. Thus, assuming a positive association between economic activity and adverse health effects would cause us to *under*estimate the adverse effect of pollution on health. Thus economic activity is very unlikely to significantly confound our identified pollution-health relationship.

As an important last point, note that concern *(ii)*—economic activity may affect both, pollution and health—is clearly irrelevant when it comes to extreme temperature events which are definitely exogenous to any one human's behavior. Research in atmospheric science strongly suggests that extreme pollution levels are triggered by high temperatures. The discussion in Appendix D strongly supports this. Particularly $O_3$ and $PM_{10}$ are secondary pollutants and oxidants; the oxidation process requires sunshine and heat. For example, the relationship between ozone and the temperature is almost linear (see Figure 15a) and the daily county-level correlation between the maximum temperature and the maximum ozone concentration is 0.7(!). The equivalent non-parametric graphs between $PM_{10}$, $NO_2$ and the mean temperature are U-shaped (e.g., see Figure 12a). When daily temperatures exceed $20^\circ C$ ($68^\circ F$), $PM_{10}$ and $NO_2$ concentrations increase strongly with correlations of about 0.25. Overall, one can think of extreme temperatures and weather conditions playing a significant causal role in *producing* high pollution concentrations. Since weather is exogenous to individuals, so are changes in high pollution concentrations that are triggered by high temperatures.

Third, with respect to the third potential identification concern *(iii)* and adaptation behavior: We argue that we intentionally want to estimate an effect that would equal an "intention-to-treat (ITT)" estimate in other settings, including avoidance behavior and human adaptation to extreme temperatures and pollution. This parameter is the relevant parameter for policymakers. Any policy action should be based on this parameter. We do not deny that people engage in avoidance behavior and spend less time outdoor when pollution levels and temperatures are high. It is also clear that it is a challenging and relevant task to study avoidance behavior. However, we believe that in this setting, a parameter measuring the health effects of a theoretical 24 hours exposure to high pollution levels, heat or cold events would not be policy-relevant. This exercise has been and can be better conducted by medical scientists in laboratory settings (cf. Stewart et al., 1970; Anderson et al., 1973; Hackney et al., 1975; Kerr et al., 1979; Horstman et al., 1988; Lippmann, 1989; Jäppinen et al., 1990; Dye et al., 2001). We think that real-world data offers great advantages over such experimental studies. The relationship that this paper intends to expose is: given that people adjust their behavior to climatic conditions, how would a decrease in the number of annual days with heavy ambient air pollution affect population health? Or: Given that humans have the capacity to adjust to extreme temperatures, based on real-world behavioral data from today, how would climate change in the form of more heat events most likely affect population health? However, without question, this ITT estimate represents a lower bound estimate as compared to a "full exposure" estimate.

Also please note that it is beyond the scope of this paper to make projections about human behavioral adaptation and/or technological progress that could facilitate adaption behavior in the future. Such projections are inherently uncertain and notoriously difficult to make. However, recent state-of-the-art empirical evidence shows that humans adapt to adverse climatic conditions and that adaptation has increased over time (cf. Deschênes, 2012; Zivin and Neidell, 2013; Barreca et al., 2013). Given this recent empirical evidence, an approach that assumes no further increases in adaptation behavior produces conservative estimates of the potential adverse health effects of climate change.

Finally it should be re-iterated what has been discussed at various places throughout the manuscript: thanks to its climatic conditions and four seasons, Germany is particularly well-suited for this type of study that links daily increases as well as decreases in extreme temperatures and pollution to immediate severe health effects. As Figures 3, 4, 13, and 16 demonstrate, the identification of parameters is based on a broad set of counties and largely avoids out-of-sample predictions. *All* German counties experienced variation in the extreme temperature variables as well as in the non-compliance pollution indicators of interest. Identification is not based on a small non-representative subset of high pollution, extremely hot or extremely cold counties, but has broad support.

# 4 Results

## 4.1 Nonparametric Relationship Between Temperature, Pollution and Health

Figure 5a-d shows scatterplots of hospital admission rates and the daily county-level *(a)* maximum temperature, *(b)* minimum temperature, *(c)* mean $NO_2$, and *(d)* mean $PM_{10}$ concentration—along with local polynomial smooth plots and confidence bands. Interestingly, the equivalent mortality rate graphs follow very similar patterns, but the patterns are even less pronounced (available upon request).

Overall, at first sight, it is difficult to detect an unambiguous positive relationship between hospital admissions and extreme temperature or pollution conditions. The data seem to be very noisy: hospital admission rates vary widely across the whole range of temperature and pollution values on the x-axis. The smoothed polynomial plots appear surprisingly flat.

**[Insert Figure 5 about here]**

However, having a closer look, a few peculiarities are of note: *(a)* The higher the temperature, the wider the confidence bands. Moreover, the admission rate seems to smoothly—but only very slightly and linearly—increase between 20 ˚C and 36 ˚C (68 ˚F and 97 ˚F). Then, admissions dip slightly, before they significantly increase for temperatures above 38 ˚C (100 ˚F).

*(b)* For daily minimum temperatures below -10 ˚C (14 ˚F), the admission rate seems to increase slowly and smoothly down to -18 ˚C (-2 ˚F). Then, again, admissions seem to dip, before they increase significantly for temperatures below -22 ˚C (-8 ˚F).

*(c)* The relationship between mean $NO_2$ concentrations and admissions increases strongly for $NO_2$ concentrations between 10 and 20 $\mu g/m^3$ and subsequently slightly for concentrations up to $63\mu g/m^3$. Between $63\mu g/m^3$ and $74\mu g/m^3$ one observes a clear increase in hospitalizations, then a drop, and for concentrations above $78\mu g/m^3$, a strong increase.

*(d)* Finally, the relationship between mean $PM_{10}$ concentrations and admissions looks pretty flat but is also slightly bumpy, up to ambient concentrations of around $56\mu g/m^3$. Then, one observes a relatively clear and strong increase of admissions up to the recorded maximum $PM_{10}$ value. This strong increase has broad support in actual observed $PM_{10}$ values—more that 200,000 county-day observations or about 12% of all observations carry $PM_{10}$ concentrations above the EU alert threshold of $50\mu g/m^3$.

### 4.2 Semiparametric Relationship Between Temperature, Pollution and Health

Next we investigate the temperature-health relationship semiparametrically using flexible temperature and pollution cut-off variables. More specifically, we run the model in equation (2) by OLS. We always employ *Approach I*, i.e, the model nets out county fixed effects, week fixed effects, and month-year fixed effects, but does not consider contemporaneous weather and pollution conditions other than the variable of interest. The variables of interest are a series of dummy variables which equal 1 if daily temperature or pollution conditions exceed a certain threshold that we allow to vary in this particular specification. In contrast, the main models below make use solely of the official *Hot Day* and *Cold Day* definitions as well as the official EU pollution alert thresholds (see Section 3.3).

However, this section intends to illustrate the marginal health impact of one additional temperature degree or a pollution concentration increase of 10 $\mu g/m^3$. Thus, Figure 6—illustrating heat effects—plots the coefficient estimates, $\sum_{h=20}^{36} \beta_h \, MaxTemp_{cd}^h$, of the regression in equation 2. Since the model controls for all 17 maximum temperature dummies simultaneously, the plotted graph in Figure 6a shows the marginal temperature impact of one degree, relative to the baseline category of less than 20 ˚C (68 ˚F).

Figure 6a illustrates that the temperature range from 26 to 28 $^\circ$C (79 to 82 $^\circ$F) induces additional admissions, but at a low rate. For temperatures above 30 $^\circ$C (86 $^\circ$F), one observes more pronounced adverse health effects. Thus, it is reasonable to follow the convention and define a *HotDay* as a day with temperature above 30 $^\circ$C (86 $^\circ$F), and use that cut-off to estimate the impact of one additional *HotDay* on health. We employ this convention henceforth.

Figure 6b applies the same approach for the daily minimum temperature. As seen, the coefficient estimates are very flat around the x-axis. One observes, if any, a slightly and partially negative relationship between cold and hospital admissions. We discuss potential explanations for this negative effect of cold on admissions below.

**[Insert Figure 6 about here]**

Figures 6c and d show the results for $NO_2$ and $PM_{10}$. The x-axis of Figure 6c shows the ambient $NO_2$ concentration in $\mu g/m^3$, where 40 equals the official EU alert threshold. One observes significantly positive effects on admissions when concentrations exceed this threshold. Recall that the US threshold is significantly higher, with an annual average concentration of 107 $\mu g/m^3$. However, surprisingly, the marginal impact of concentrations even below the EU threshold of 40 $\mu g/m^3$ seems to be strong and significant. This suggests that even the EU alert threshold is too high to avoid adverse health effects, even in case of full compliance by the EU member states. Another explanation could be that other contemporaneous weather and pollution factors confound the "pure" relationship between $NO_2$ and health, e.g., Figure 17 shows that $O_3$ and $NO_2$ are *negatively* correlated which means that $O_3$ levels are higher when $NO_2$ levels are lower. Recall that the underlying *Unconditional Approach* used here does *not* consider contemporaneous weather and pollution conditions.

The findings for Figures 6d and $PM_{10}$ are very clear and confirm what we observe in the nonparametric Figure 5d above. One finds that adverse health effects significantly increase for $PM_{10}$ values above 50 $\mu g/m^3$. This is also the official EU alert threshold and thus seems to be well-targeted, while the US threshold of 150 $\mu g/m^3$ is clearly too high to avoid adverse population health effects, assuming similar behavioral responses in the US and Germany. In general, one can say that all graphs in Figure 6 are in line with the nonparametric scatterplots in Figure 5. However, the dose-response relationship between temperature and single pollutants on the one hand, and health on the other is carved out in a much clearer way in Figure 6. The equivalent nonparametric and semiparametric graphs for the mortality rates follow very similar patterns, albeit the adverse health effects are less pronounced. They are available upon request.

The following section focuses on models that solely employ the official *Hot* and *Cold Day* definitions as well as the official EU alert thresholds for pollutants. Instead of varying the threshold parameters, the next section analyzes the differences between *Approach I* and *II* as discussed in Section 3.1, i.e., the differences in estimates when considering vs. not considering contemporaneous weather and pollution conditions in addition to the parameter of interest.

### 4.3 Health Effects of Extreme Temperatures: Unconditional and Conditional on Contemporaneous Weather and Pollution Conditions

Table 1 illustrates the impact of extreme heat and cold on health. Panel A shows the effects on hospitalizations and Panel B the effects on mortality. The dependent variable always measures the all-cause hospitalization or mortality rate, i.e., does not distinguish by diagnoses. Each column in the panels represents one model estimated according to equation (3). Columns (1) to (4) run *Approach I*, the *Unconditional Approach*. This means that these models do *not* control for any contemporaneous weather or pollution conditions and solely focus on the extreme heat or cold measures as indicated in the rows. For example, the result in column (1) of Panel A measures the overall *Hot-Day-Effect*, including correlated factors such as higher ozone levels, more sunshine, or less rainfall.

Column (1) shows that a *Hot Day*, i.e., a day with a maximum temperature of more than 30 $^\circ$C (86 $^\circ$F) leads to a 5.4% increase in hospitalizations and to a 9.8% increase in deaths. For the whole of Germany, this translates into 2,500 additional hospital admissions and 240 additional deaths. Again, please note that this represents the overall Hot-Day-Effect, including all other weather and pollution conditions that prevail on a *Hot Day*.

Column (2) shows the effect of the fourth consecutive Hot Day. A *Heat Wave Day* increases hospital admissions by about 6% and mortality by about 20%, i.e., the mortality effect doubles after 3 consecutive *Hot Days*. Appendix C1 shows that 2% of all county-day observations during the time period between 1999 and 2008 were *Hot Days*. Only 0.3% of all observations were *Heat Wave Days* in Germany. This translates into an annual average of 7.2 *Hot Days*, of which 1.2 are *Heat Wave Days*.

Column (3) shows the overall effect of a *Cold Day*, i.e., a day with a minimum temperature of less than -10 $^\circ$C (14 $^\circ$F). We find that, not netting out other weather and pollution conditions, a *Cold Day* leads to 2% *fewer* hospital admissions and 1% more deaths, i.e., the effects are *(i)* significantly smaller as compared to the health effects of heat events. The *(ii) de*crease in hospitalizations is very likely an artifact of higher hospital admission costs and less outdoor activities, e.g., through snowfall and bad traffic and weather conditions on *Cold Days* (Schwartz et al., 2004). This is reinforced by column (4) which shows that hospital admissions significantly decrease by almost 8% on the fourth consecutive *Cold Day* while mortality is not affected by *Cold Waves*. However, again, please keep in mind that these effects do not capture only the effect of extreme cold on health, but also the effect of correlated conditions such as lower $O_3$ and higher $NO_2$ pollution (see Figures 12 and 15 in the Appendix).

[**Insert Table 1 about here**]

Column (5) show results from estimation using the fully saturated *Conditional Approach II*. Here, we add an extensive set of 7 continuous weather measures, such as sunshine and precipitation as well as 15 interaction terms between those weather measures (cf. Table C1).[16] Moreover, we simultaneously control for high pollution non-compliance days. In addition, the continuous mean values of the 5 pollutants—CO, $O_3$, $NO_2$, $SO_2$, $PM_{10}$—their 5 quadratic and 5 cubic terms as well as 10 of their cross interactions are added to the model (cf. Table D1). The covariates model the rich non-linear interactions between pollutants, as shown in Figure 17 in the Appendix. Lastly, the model includes 25 interaction terms between the 7 continuous weather and the 5 continuous pollution measures modelling the nonparametric associations displayed in Figures 12 and 15. In total, *Approach II* adds 77 continuous pollution and weather measures as well as their interactions to the model, plus up to 4 extreme temperature and 3 high pollution non-compliance indicators.

When one considers this rich set of contemporaneous weather and pollution conditions, in the *Conditional Approach II*, extreme cold alone does not affect mortality at all. The mortality effects of about +1% in the *Unconditional Approach I* are further reduced and become insignificant. The *Cold Day* effect on hospitalizations is also small in size (0.72***; +1.2%), but highly significant. Interestingly, the negative *Cold Wave Day* hospitalization coefficient remains large in size (-9.6%) and significant. The finding that the first (presumably unexpected) *Cold Day* slightly increases mortality but *Cold Wave Days* strongly decrease hospitalizations yield strong evidence in support for the "higher transportation cost hypothesis" outlined above. Bad weather and transportation conditions associated with several subsequent extremely *Cold Days* are the most likely reason for the drop in hospital admissions. This is in line with existing research from epidemiology (Schwartz et al., 2004). The drop in admissions does not seem to trigger an (immediate) increase in mortality.

[16] The model does not consider the three plain temperature indicators since the main variables of interest are the four extreme temperature indicators.

Netting out all climatic factors that prevail on *Hot Days*, in the *Conditional Approach II*, the hospitalization effect of a *Hot Day* is reduced by the factor 2 and the effect of a *Heat Wave Day* is even reduced by a factor of 6. However, both coefficient estimates are still highly significant and of meaningful size ($\sim$+2.5%). As compared to the *Unconditional Approach I*, the *Hot Day* mortality effect is reduced by the factor 6 to +1.5%. The *Heat Wave Day* mortality effect shrinks by a factor of 2 to (a still large) +10% increase in the death rate on the fourth consecutive *Hot Day*. When adding the sets of covariates step-wise, it becomes clear that the main reduction in coefficient sizes is due to the inclusion of the *(i)* seven continuous weather measures such as precipitation, the *(ii)* five continuous mean pollution measures as well as *(iii)* the five quadratic and five cubic terms of the pollutants (see Panels A of Tables C1 and D1). Overall, these sharp declines in coefficient sizes illustrate the importance of considering other health-damaging weather and pollution conditions associated with extreme heat.

### 4.4 Health Effects of High Pollution Concentrations: Unconditional and Conditional on Contemporaneous Weather and Pollution Conditions

The setup of Table 2 is identical to that of Table 1. Considering the *Unconditional Approach I*, it is easy to see that for all three pollutants—$NO_2$, $O_3$ and $PM_{10}$—the following holds: When ambient pollution levels cross EU alert thresholds, hospital admissions increase significantly. However, the effects for $O_3$ and $PM_{10}$ are small and have magnitudes of about 1%. On the other hand, the $NO_2$ effect is relatively large and associated with 8.7% more hospital admissions.

Mortality rates increase by between 1.5% and 4.5% when $NO_2$, $O_3$ and $PM_{10}$ levels increase above EU alert thresholds. For example, during a *$PM_{10}$ non-compliance day* when concentrations increase above 50 $\mu g/m^3$, the death rate increases by 2.5% or about 61 deaths for the whole of Germany. These effects are of significant relevance since 12.8% of all county-day observations in the sample carry $PM_{10}$ concentrations above EU alert thresholds. No county entirely avoids violation of the EU norms over the ten years of observation. Similarly high pollution levels are reached for $O_3$ (9.3% of all obs.; +4.5% deaths) and $NO_2$ (11.9% of all obs.; +1.3% deaths).

**[Insert Table 2 about here]**

However, interestingly and maybe surprisingly, when considering the extremely rich set of concurrent weather and pollution conditions—Figures 12 and 15 illustrate the nonlinear relationships between pollution levels and weather conditions—all formerly significant associations between pollutants and mortality dramatically shrink in size and become insignificant. The effects on hospital admissions remain partly significant, but the coefficients are very small in size; the effects in percentage terms tend towards zero. This finding is absolutely in line with research in medical science and epidemiology, where high concentrations of a single pollutant are seen rather as an indicator for overall general adverse environmental conditions that put strain on human bodies. Laboratory experiments show that concentrations of single pollutants need to be extremely high—higher than they typically occur in outdoor environments, at least in Germany—before adverse physical health effects such as lung, pulmonary, or respiratory function responses can be detected (cf. Stewart et al., 1970; Anderson et al., 1973; Hackney et al., 1975; Kerr et al., 1979; Horstman et al., 1988; Lippmann, 1989; Jäppinen et al., 1990; Dye et al., 2001).

However, for the regulator, the relevant pollution parameter of interest should be the one that the *Unconditional Approach I* identifies. To date, regulatory thresholds always apply to unconditional pollution levels and do not consider simultaneous weather and pollution conditions. In the EU and the US, measures from official ambient monitors are taken on a daily basis. If they exceed official thresholds, action must be taken. It that sense, it may be of interest for the researcher that one only detects small health-damaging effects of single pollutants when comprehensively considering all other health-damaging weather and pollution conditions, but the policy implications of this empirical exercise are questionable.

Both in Germany and the US, $SO_2$ and CO concentration levels rarely exceed alert thresholds. However, in Germany, in more than ten percent of all county level observations, $PM_{10}$, $O_3$, and $NO_2$ levels were significantly elevated above EU thresholds and associated with more deaths and hospitalizations. $NO_2$ levels even below the EU threshold are associated with adverse health effects. This suggests significant public health benefits from stricter regulation and/or stricter enforcement of the existing regulation. It should also be recalled that the US regulatory thresholds for $PM_{10}$ and $NO_2$ are 2 to 3 times higher as compared to the EU regulatory thresholds (Environmental Protection Agency (EPA), 2013); actual average $PM_{10}$ and $O_3$ levels in the US are also 2 to 3 times higher than in Germany (Environmental Protection Agency (EPA), 2011). Obviously, the public health benefits from lowering the US thresholds and actual pollution concentrations could be tremendous.

### 4.5 Cause-Specific Health Effects of Extreme Heat Conditional on Contemporaneous Weather and Pollution Conditions

Table 3 now disentangles the extreme temperature effects by diagnoses applying the *Conditional Approach II.*[17] We can summarize the following:

First, *Hot Days* significantly increase cardiovascular, respiratory, metabolic and neoplasmic hospital admissions. Infections are unaffected. The latter is confirmed when looking at mortality effects in Panel B. Plausibly, *Hot Days* do not trigger infectious and metabolic deaths.

Second, ongoing heat—i.e., the fourth consecutive *Hot Day*—increases all of the above listed diagnosis-specific deaths. The same holds true for hospital admissions; however, neoplasmic and cardiovascular health shocks requiring inpatient stays seem to be triggered by a single *Hot Day*, and are not exacerbated by the onset of a heat wave.

**[Insert Table 3 about here]**

Third, in line with expectations, metabolic and neoplasmic health shocks are totally unrelated to extreme cold while cardiocascular, respiratory and infectious diseases are triggered by extreme cold. As mentioned above, longer periods of extreme cold are most likely associated with bad traveling conditions and therefore *de*crease admissions significantly.

Fourth, *Cold Days* and *Cold Wave Days* are not associated with higher mortality rates for any of the cause-specific deaths. All coefficient estimates are very small in size.

Figures 7 graphically illustrates the effect of extreme heat on the different disease types and also considers the overall relevance of the different disease groups. Hospitalization effects are evenly distributed across diagnostic categories. The effects of one *Hot Day* is relatively moderate but always significant. They range between 2% and 6% across disease categories. With increasing duration, heat particularly triggers infections as well as metabolic health issues, whereas cardiovascular and neoplasmic admissions occur at the onset of heat events (Figure 7b).

**[Insert Figure 7 about here]**

The latter may be due to the fact that people with cancer die at a significantly higher rate during heat events (Figure 7c and d). Figures 7c and d also illustrate that cardiovascular (50%) and neoplasmic (25%) deaths make up 75% of all heat-related deaths. One *Hot Day* slightly elevates the rate of these two types of deaths, and also respiratory deaths, but only by between

[17] The results are similar for *Approach I* but more pronounced and available upon request.

2% and 4%. Ongoing heat boosts all type of deaths, independent of diagnoses—in particular respiratory deaths (+35%) and deaths dues to infections (+60%), but also cardiovascular (+9%) and metabolic (+17%) deaths.

### 4.6 Robustness Checks

Table 4 presents a series of robustness checks. The reference specification is always the effect of one *Hot Day* on the hospital admission rate in the *Unconditional Approach I* (column (1) in Table 1). All findings also hold when using the mortality rate as the dependent variable. These results are available upon request.

Column (1) in Panel A of Table 4 reports results with standard errors clustered at the state instead of the county level (Cameron and Miller, 2011). Column (2) applies two-way clustering by county and date (Cameron et al., 2011). As compared to the standard specification, standard errors roughly double, but the coefficient estimates remain highly significant at the one percent level.

The next three columns add nation-level (column (3)), state-level (column (4)), and county-level (column (5)) time trends to the model. The latter two specifications reduce the magnitude of the estimated *Hot Day* coefficients somewhat. However, they remain highly significant and of an economically meaningful size (4.7% and 3.4%).[18]

**[Insert Table 4 about here]**

The first column in Panel B show results from another way of modelling the heat-health relationship illustrated in Figures 5 and 6. Here, the model includes the maximum daily temperature as a continuous variable, along with the *Hot Day* dummy and an interaction between *Hot Day* and a continuous variable that captures the difference between the average maximum temperature that prevails on *Hot Days*, 31.9 $^\circ$C (89.4 $^\circ$F), and the county-specific maximum temperature on a given *Hot Day*. In other words: The interaction term indicates the degree to which hospitalizations additionally increase with every temperature increase above 32 $^\circ$C (89 $^\circ$F). An average *Hot Day* increases admissions by about 3%. For every degree Celsius above 32 $^\circ$C (90 $^\circ$F), the admission rate rises by another 1%—a factor four times as large as the general impact factor of a one degree increase in temperature. This tells us that increasing temperatures are especially harmful once the *Hot Day* threshold has been surpassed.

[18] The slightly larger decrease found with the addition of county-level trends is partly due to the fact that we have to restrict this specification to the years 2006 to 2008 due to computer memory constraints.

Column (2) of Panel B interacts *Hot Day* with a dummy for weekends. This specification indirectly tests behavioral adaptations to heat—under the assumption that individuals have more and better options to engage in adaptation behavior on weekends as compared to weekdays. Although the coefficient estimate of the interaction term carries a negative sign, it is small in size and not statistically significant from zero. This suggests that adaptation behavior may exist but is unlikely to play an economically significant role here. However, other mechanisms may be at work on weekends. For example, stress related to work may exacerbate the adverse health effects of heat on weekdays. This is in line with the sharp decrease in hospitalizations on weekends—by a staggering 50%. To the extent that this hypothesis is true, it would reinforce a potential positive adaptation effect on health on weekends. Thus, finding no significant reduction in heat-related admissions on weekends strengthens the notion that behavioral adaptation behavior may exist but, at least in Germany, there is suggestive evidence that it is of secondary importance to the general heat-health relationship. Other explanations for the strong decrease in admissions on weekends could result from institutional factors related to hospital management.

Columns (3) and (4) indirectly test another form of adaptation behavior: Namely, whether the human body adapts to heat and warmer temperatures when humans live in warmer versus colder regions. Econometrically, we define a "warm region" as a region where the mean annual county-level temperature falls into the highest temperature quartile for Germany ($>10.2^\circ$C ($50^\circ$F)). Analogously defined is a "cold region" which is a county with a mean annual temperature below the lowest temperature quartile ($<9.0^\circ$C ($48^\circ$F)). Accordingly, we define two dummy variables, *Cold* and *Warm Region* and add them to the models in levels and in interactions with the *Hot Day* indicator. There is clear evidence in line with the human body adaptation hypothesis since, in warm regions, the effect of a *Hot Day* is 0.8 ppt. (1.4%) smaller than the average *Hot Day* effect. Likewise, in cold regions, the effect of a *Hot Day* is 0.8 ppt. larger, although it is imprecisely estimated. Note that this finding would also be in line with heat-(in)sensitive individuals sorting into colder (warmer) regions. In any case, although human body adaption or sorting seems to exist, it is also clear that it amounts to a relatively small overall effect and does not alter our basic findings: In warm regions, *Hot Days* still lead to 4.2% more hospital admissions and in cold regions, to 6.7% more admissions.

# 5 Monetized Health Costs of One Additional *Hot Day*: Implications of Climate Change

## 5.1 Increase in Deaths Due to Heat: Who Dies and Is There "Harvesting"?

Next, we explore in more detail: Who dies during weather events? Obviously, the conclusion from this analysis has important implications in assessing the economic relevance of additional heat events due to climate change. The literature discusses a phenomenon called the "harvesting hypothesis" (cf. Rabl, 2005; Fung et al., 2005). According to the harvesting hypothesis heat events temporarily lead to a higher mortality rate, particularly among the elderly who are already in bad health. The hypothesis suggests that people who die during heat events would have died a few days later, even in the absence of the heat event. If this were true, then the overall effect of heat events on population health would be dramatically reduced since heat would only reduce the life expectancy of the old and sick by a few days. Empirically, a decline in mortality rates in the days following a heat event is often cited as evidence strongly in line with the harvesting hypothesis.

This paper makes several contributions to the harvesting debate. First of all, we do not only focus on mortality but also on hospital admissions. In the *Hospital Admission Census*, we see the age (group) of the admitted patients, how long they stayed in the hospital, and whether they died subsequent their admission. This allows us to pinpoint the population health loss of heat in a much more precise way. For people who do not die after an admission, one knows precisely how many days they had to stay in the hospital. This allows us to calculate the overall number of hospital days triggered by a heat event. Thus, our research design allows us to evaluate the harvesting hypothesis with reference to several different endpoints in order to gain a much deeper understanding of its practical relevance.

Second, this paper tests whether mortality rates actually decline in the days after a heat event. If the results were completely driven by harvesting, one would expect the mortality effect observed on *Hot Days* to be completely reversed during the next few days. Figure 8 provides a test. It is based on the *Unconditional Approach I* not considering contemporaneous weather and pollution conditions and plots the development of the mortality rate during the days preceding and following a *Hot Day*. As expected, the adverse health effects peak on the *Hot Day* itself. Within three days following a *Hot Day*, the effect decreases strongly to below +2%, but is still significantly greater than zero. As expected, during the days prior to a *Hot Day*, as well as after day three following a *Hot Day*, there are virtually zero remaining effects of the heat event.[19] This can be interpreted as evidence against the harvesting hypothesis because one would expect that harvesting would lead to a *de*crease in mortality rates in the days immediatetly after a heat event.

**[Insert Figure 8 about here]**

However, the results presented in Figure 8 may potentially suffer from omitted variable bias: the temporal correlation of temperature and also pollution is significant. The plain correlation coefficient between the maximum daily temperatures on a *Hot Day* and the following day is 0.6. For ozone it is even higher, 0.7. More specifically, the average maximum temperature on a day preceding a *Hot Day* is 28.9 $^\circ$C (84 $^\circ$F) and on a day following a *Hot Day* is 28.4 $^\circ$C (83 $^\circ$F). Consequently, maximum ozone levels on the days prior and subsequent to a *Hot Day* are also highly elevated, between 129 and $127\mu$g$/m^3$. The underlying model displayed in Figure 8 does *not* disentangle the health effects due to high temperatures and adverse pollution conditions during the days surrounding a *Hot Day*. Nor does it take into account that the leads and lags are likely to be *Hot Days* as well. Since we know that environmental conditions on the days preceding and following a *Hot Day* are likely to lead to adverse health (and thus increases in hospitalizations and mortality), it becomes clear that this impact factor works in opposition to potential harvesting effects and may actually obscure it. Disentangling the two opposing forces is challenging.

Next, Figure 9 plots the coefficient estimates from interaction terms between *Hot Day* and 5-year age group dummies that have been added to the model.[20] It is easy and nice to see that the hospitalizations on *Hot Days* are driven by people above 55, and in particular by the elderly between 71 and 80 years of age.[21] This finding yields strong evidence in favor of the harvesting hypothesis. It may be reconciled with the apparent refutation of a harvesting effect in Figure 8 with reference to the time frame: the four-day window in Figure 8 might have been too short to identify a reversal of mortality rates. Thus we now turn to a more long-term perspective.

**[Insert Figure 9 about here]**

Finally, a sound test for the empirical relevance of the harvesting hypothesis is to aggregate data at the *year*-county level. Using this method, one can test whether the occurrence of one additional *Hot Day* has a significant impact on the annual mortality and hospital admission

[19] The according graph for hospitalizations looks very similar and is available upon request.
[20] The results here are for the *Unconditional Approach I* but are almost identical for the *Conditional Approach II*.
[21] The plain *Hot Day* coefficient estimate is of magnitude 0.16 (i.e. 0.3% of the mean) and not statistically different from zero.

rate. In other words: If it were true that heat events triggered persistent adverse health effects that would not have occurred in the counterfactual state, then an additional *Hot Day* should also significantly elevate the annual mortality and hospitalization rate, not only the daily one. However, due to data limitations and power issues, researchers often cannot implement this test since one obviously needs enough years of observation with enough variation in the annual number of *Hot Days*. In addition, the number of regional units of observations—in this case counties—should be sufficiently large. Our data and setting fulfills all of these conditions. Column (5) in Panel B of Table 4 reports results from the test and uses data aggregated at the county-year level, resulting in 4,356 observations. The results show a coefficient of size 0.03—i.e. reduced by a factor of 100(!) as compared to the standard estimate in column (1) of Table 1. However, this coefficient is statistically significant at the five percent level. It translates into 25 additional hospital admissions due to one additional *Hot Day* per year. This means that we can assume that these 25 people would not have been hospitalized during that year in the absence of one *Hot Day*. Obviously, this finding delivers strong support for the harvesting hypothesis. The finding is reinforced by the mortality data which yields a highly significant coefficient of 0.0024. Again, the yearly coefficient is more than 100 times smaller than the daily one, and translates into only minor annual mortality increases of 0.08% or 2 people per additional *Hot Day* in Germany.

## 5.2 Monetizing the Health Loss of One Additional *Hot Day*

As a last step, this paper seeks to assess and monetize the total health effects triggered by extreme weather conditions and to derive implications from climate change. Although this exercise requires many assumptions, we believe that it is an important first step to conduct an evidence-based cost-benefit analyses of climate change regulation.

Given the complex nature of climate change, it is not surprising that projections are relatively vague. According to the Stern (2006) report, over the past 100 years, the world's temperature increased by 0.74 $^\circ$ C (1.33 $^\circ$ F) and will further increase by between 1.8 and 4 $^\circ$ C in the next 100 years. However, concrete statements are hard to find in this report. The Intergovernmental Panel on Climate Change (IPCC) states that it is very likely that hot extremes, heat waves and heavy precipitation events will continue to become more frequent (Intergovernmental Panel on Climate Change (IPCC) (2007), p. 46, 53). The underlying state-of-the art global climate model of the Intergovernmental Panel on Climate Change (IPCC) is the third version of the so-called *Hadley Centre Coupled Model (HadCM3)* (Pope et al., 2000). These climate models are extremely complex and require many underlying assumptions and scenarios. Deschênes and

Greenstone (2011) make use of the *HadCM3 model* and the "business-as-usual" scenario to predict the change in the number of *Hot Days* for 2070 to 2099 relative to 1968 to 2002 and different US regions.[22] For the region whose climate comes closest to Germany's, New England, Deschênes and Greenstone (2011) estimate a 20% increase in the number of *Hot Days*. Hübler et al. (2008) make use of the Regional Climate Model *REMO* and predict "two to five times as many hot days [for Germany from 2071 to 2100 relative to 1971 to 2000]" (p. 383).

Given the difficulty and inherent uncertainty of making such long-term predictions (Heal and Millner, 2013), for the following reasons, the remainder of the paper focuses on the monetized health effects of *one additional* Hot Day: *(i)* Additional *Hot Days* are extremely plausible climate change predictions and are always referred to in predictions of climate change models. *(ii)* One additional *Hot Day* is an intuitively plausible concept. The monetized health effects can be easily adapted to varying climate change predictions. *(iii)* One additional *Hot Day* represents an increase of about 14% in the total number of*Hot Days*, which is very much in line with the Deschênes and Greenstone (2011) prediction for New England using *HadCM3*. *(iv)* Finally, we abstain from estimating the impact of fewer *Cold Days* for two reasons: First, the empirical models do not yield strong evidence for the notion that extreme cold significantly affects population health in Germany. Second, climate change projections concerning *Cold Days* are not unambiguous. On the one hand, the Intergovernmental Panel on Climate Change (IPCC) (2007) projects that snow cover will contract (globally) in the future. On the other hand, loss of arctic sea ice has been linked to the recent extreme cold weather in North America and Europe (Liu et al., 2012). The latter finding suggests that climate change would lead to both more heat *and* cold events in the mid-latitudes.

Table 5 summarizes the results from the empirical models and calculates the total health costs of one *Hot Day* under competing assumptions. The basis for these calculations is Table 1 and equivalent tables using the dependent variables *Hospital Days* and *Hospital Death* reported in Appendix A1. The health effects that we monetize consist of *(i)* hospital days due to a *Hot Day*, *(ii)* death after a hospital stay due to a *Hot Day*, *(iii)* immediate death due to a *Hot Day*—in accordance with the dependent variables (see Table A1 and B1). Table 5 presents the results using estimates obtained from three main models: *Approach I*, *Approach II* on the day-county level, and the approach outlined in column (5) of Table 4, which aggregates at the *year* county level, completely internalizing the harvesting effect.

The first three columns of Table 5 show results that evaluate the economic value of hospital

[22] *Hot Days* are in this case defined as days with a mean daily temperature above 90 °F.

days that are triggered by one additional *Hot Day*. Column (1) shows results where we simply multiply the number of triggered hospital days by the average health care costs of one hospital day in Germany, which is €500 (German Federal Statistical Office, 2013b). Column (2) shows the value of the *Hot Day*-induced loss in labor productivity by multiplying the rough share of the working population, 50%, by the number of hospital days and the average daily gross wage in 2012, including employer-mandated benefits: €150 (German Federal Statistical Office, 2013a). Columns (3) and (4) show results where we convert the number of hospital days into Quality-Adjusted Life Years (QALYs) by assuming that 365 hospital days equal a loss of one QALY (column (3)) and half a QALY (column (4)). We evaluate one QALY with €100,000 ($130,000) (Shiroiwa et al., 2010; Kniesner et al., 2010; Robinson et al., 2013). Note that alternative assumptions about the value of a QALY do not significantly alter the main findings below.

[**Insert Table 5 about here**]

Column (5) shows results that assess the value of the total number of deaths, which is the sum of deaths after a hospital stay as well as immediate deaths. Again, one QALY is evaluated at a value of €100,000. Results in the first two rows are derived from the two approaches that use the data at the day-county level and that ignore harvesting, we assume that people who died would have lived another calendar year absent the heat event. For the third approach—results displayed in the third row—which aggregates at the year-county level and accounts for harvesting, we assume that people who died would have lived another 30 years.[23]

As can be seen in the final two columns: *(i)* the upper and lower bound QALY assumptions barely affect the estimates (and neither do varying assumptions about their value). *(ii)* The *Unconditional Approach I* yields the largest monetized health loss estimates and the approach that aggregates at the year-county level, and thus accounts for harvesting, yields the lowest monetized health loss. *(iii)* All estimates are relatively close and relatively small in size. The estimated monetized losses range from €6m to €43m per *Hot Day* for an entire nation with a GDP of €2.5 trillion and 82 million residents. The according values for the US would lie between $30m and $212m. The values equal between €0.07 ($0.10) and €0.52 ($0.68) per resident.[24] *(iv)* Assuming that climate change leads to a permanent increase of one additional *Hot Day* per year and taking the largest annual loss estimate of €43m, the nominal health-related welfare loss over one life cycle, i.e., 80 years, would accumulate to €3.4bn for Germany. Applying a discount rate

[23] 30 years is roughly the difference between the average current age of Germans and their life expectancy. Alternative assumptions do not alter the main findings.

[24] Assuming an exchange rate of 1.3 and that the US has 311/82=3.8 times as many residents.

of 2.5% reduces this sum to €470m or about €6 ($8) per resident. The according values for the US would be $16.8bn and $2.3bn, respectively.

Finally, it should be stressed that these back-of-the-envelope calculations solely consider the *health*-related costs of one additional *Hot Day*. They also ignore any health effects that do not manifest in immediate hospital admissions or death. We also abstract away from costs associated with health-related avoidance behavior, as well as from any health effects stemming from potential climate change-related increases in flooding, hurricanes, and tornados. As a comparison, the second costliest hurricane in US history—Hurricane Sandy—is estimated to have cost 72 human lives in the US, most of whom would not have died absent the hurricane (Blake et al., 2012). Assuming that these humans would have lived another 30 years, the monetized mortality-loss of Sandy would be $280m or about $1 per US resident. The total loss is an estimated $50bn, the large majority of which is attributed to property damage.

## 6 Conclusion

Around the world, climate change—with its potentially adverse effects on mankind—and the question of appropriate regulatory measures are heavily debated. The issue will continue to be at the top of policy agendas. To put policymakers into a position to be able to seriously evaluate and balance costs and benefits of climate change and according regulatory efforts, scientists have to provide state-of-the art empirical analyses and cost projections.

This paper assesses more comprehensively than any previous paper the adverse population health effects of extreme temperatures and pollution. Weather and pollution are inherently linked. Thus it is necessary to consider many high quality measures of both. At the day-county level, we link an extensive set of administrative weather and pollution measures from more than 2,000 ambient monitors obtained over a time period of 10 years to two register datasets: *(i)* a mortality census comprising all deaths on German territory from 1999 to 2008, and*(ii)* a hospital census of all admissions from 1999 to 2008. All databases together allow us to comprehensively analyze the short-term, immediate, health effects of weather and pollution and to draw conclusions of the implications of climate change for population health.

This study makes the following important contributions: First, in line with the existing literature, it finds that extreme heat triggers significant increases in adverse health events that lead to hospital stays or deaths. The length of a heat wave determines which types of diseases are primarily triggered. For example, infectious and metabolic hospital admissions strongly increase with the length of a heat event. However, the admission rate for other disease categories, e.g., cardiovascular or neoplasmic diseases, are front-loaded and tend to occur at the beginning of heat events. Thus the average impact of a *Hot Day*—a day with maximum temperatures above 30 $^\circ$C (86 $^\circ$F)—on hospital admissions is about 5% and remains relatively stable over time. A *Hot Day* also leads to a 10% increase in the overall mortality rate. Again, as for hospitalizations, the heat-related causes of death vary with the length of the heat event, e.g., metabolic and infectious deaths are not affected at the beginning of heat events; ongoing extreme heat particularly boosts respiratory and infectious deaths.

Second, we do not find empirical evidence that extreme cold significantly affects population health. All estimated effects are very small or even negative. The latter applies to the impact of cold waves on hospital admissions and is very likely an artifact of the associated bad outdoor and driving conditions that may prevent some hospital admissions.

Third, pollution levels above EU alert thresholds are significantly associated with both increased hospitalizations and deaths. However, this only holds when *ignoring* simultaneous weather conditions and other pollutants that may also drive adverse health effects. Climate remains a poorly understood, complex system, but it is well known that certain climatic conditions serve as input factors for the formation of others. For example, ozone is an oxidant and the chemical product of CO and $NO_x$ under the influence of high temperatures and sunshine. This explains why these climatic factors are correlated. When one disentangles and controls for these simultaneous climatic factors, the impact of a *single* pollutant via elevated ambient concentration levels converges to zero. This is in line with findings from medical scientists and epidemiologists who showed in laboratory experiments that surprisingly high pollution levels of single pollutants are required before significant adverse health effects could be detected in humans. These and the findings from this study suggest that the real-life adverse health effects of pollutants mostly stem from a *combination* of several adverse climatic factors and elevated pollutants that generally prevail on days with high ambient air pollution. Thus, high pollution levels of single pollutants can also be interpreted as general indicators of adverse outdoor environmental conditions. However, since—to date—regulators around the world only regulate "unconditional" pollution levels and do not consider contemporaneous climatic conditions or implement action plans accordingly, we see the unconditional effects as the relevant policy parameters of interest. When ambient pollution levels of $O_3$, $NO_2$ and $PM_{10}$ increase above the current EU thresholds, which happens on 10% of all days in Germany, hospitalizations and deaths clearly and significantly increase by between 1 and 9%. This strongly suggests that lower EU pollution alert thresholds would be beneficial for population health and would save lives—at least to the extent that lower thresholds actually lead to lower pollution. For example, one day with ozone levels above 120 $\mu g/m^3$—a value very similar and regularly exceeded for both the US and the EU—would lead to 1.3 fewer deaths per 1 million residents. In the US, $NO_2$ and $PM_{10}$ regulatory thresholds as well as actual concentration levels are two to three times higher than in the EU (Environmental Protection Agency (EPA), 2011, 2013). To the extend that our findings are transferable to the US, tightening the regulation to EU standards and enforcing it strictly could translate into 38 ($NO_2$) and 74 ($PM_{10}$) fewer deaths per 100 million residents and avoided high pollution day. Obviously, the same applies to a stricter enforcement of the EU regulation, which has been violated in more than 10% of all German day observations.

Fourth, in general, all findings differ significantly in size depending on whether one considers a rich set of simultaneous weather and pollution conditions *in addition to* the temperature or pollution variable of interest. For example, the adverse health effects of a *Hot Day* decrease by factors of two or more when comprehensively controlling for all contemporaneous climatic conditions like high ozone or particular matter concentrations that typically prevail on *Hot Days*.

Fifth, we apply several methods to assess the validity of the harvesting hypothesis according to which mostly older people die or are admitted to hospitals during heat events—and that those people would have died or would have been admitted absent the heat event. We provide strong support in favor of the harvesting hypothesis by *(i)* looking at the evolution of mortality and admission rates before and after heat events, *(ii)* looking at the age structure of those admitted on *Hot Days* and finally, and most convincingly, *(iii)* exploiting the richness of the data by aggregating it at the year-county level and using annual variation in the number of *Hot Days* as the identifying variation. If one applies the latter test that accounts for harvesting comprehensively at the annual level, the basic heat-health relationship remains robust and highly significant, but the strength of the dose-response function is reduced by the factor 100. This finding strongly supports the view that the adverse health effects of heat we study are overwhelmingly temporary phenomena with little long-lasting impact on population health.

Finally, we monetize the health effects of one additional *Hot Day*—a very concrete and plausible prediction of climate change. We provide the results for three main approaches: *(a)* estimates obtained at the day-county level, which ignore contemporaneous climatic conditions and harvesting, *(b)* estimates at the day-county level that consider contemporaneous climatic conditions, but ignore harvesting, and *(b)* the year-county level estimates, which ignore contemporaneous climatic conditions, but account for harvesting. The total estimated health loss of one *Hot Day* represents a monetary welfare loss of between €6m and €43m for the whole of Germany ($30m to $212m for the US)—or up to €0.52 ($0.68) per resident. The size of the overall loss decreases as we move from approach *(a)* to *(c)*. Whether one considers contemporaneous environmental conditions or not affects the results by a factor of two. Even more important is whether one considers harvesting: consideration of the harvesting phenomenon affects results by a factor of six.

As a last point, we would like to stress the limitations of this study. First of all, this paper solely studies the *health* effects of extreme temperature and pollution. Second, it does not consider health effects that lead to ambulatory doctor visits or no treatments at all. However, our calculations demonstrate that mild health effects do not seem to matter a lot when it comes to the overall monetized health effects. Moreover, a very large share of the serious health effects should be captured by this study. One important exception may be fetal health effects which may have long-lasting and expensive impacts (Currie et al., 2013). We acknowledge that we do not consider adverse health effects of avoidance behavior in the estimates but, at the same time, this omission should not significantly impact the central findings of this study, which incorporates avoidance behavior in its estimates.[25] More importantly, to the extend that climate change leads to an increase in floods, tornadoes, and hurricanes, we underestimate the total health effects. However, in the time period from 1993 to 2006, the average *total* number of deaths from all these natural disasters combined was 273 in the US (Goklany, 2009). Even if 20% of these incidences were triggered by climate change, the overall impact on the total cost estimates would be relatively moderate. Lastly, this study solely focusses on short-term adverse health effects of extreme climatic conditions. It entirely disregards any long-term effects that extreme temperatures or high pollution concentrations may have on health (cf. van den Berg, 2006; Currie and Almond, 2011; Zivin and Neidell, 2013; Currie et al., 2013).

We see this study as a first step to better assess climate change-related social costs. More studies on other regions and other outcome measures are instrumental for a better understanding of how weather, pollution and human health interact.

## References

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

Almond, D., L. Edlund, and M. Palme (2009). Chernobyl's subclinical legacy: prenatal exposure to

[25] Obviously, for predictions about the future, assumptions about avoidance behavioral matter and the adverse health effects would be mitigated if avoidance behavior further increased (cf. Barreca et al., 2013) and reinforced if people engaged in less avoidance behavior in the future.

radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics 124*(4), 1729–1772.

Anderson, E. W., R. J. Andelman, J. M. Strauch, N. J. Fortuin, and J. H. Knelson (1973). Effect of low-level carbon monoxide exposure on onset and duration of angina pectorisa study in ten patients with ischemic heart disease. *Annals of Internal Medicine 79*(1), 46–50.

Arya, S. P. (1998). *Air Pollution Meteorology and Dispersion* (1st ed.). Oxford University Press.

Barck, C., T. Sandström, L. J., G. Halldén, M. Svartengren, V. Strand, S. Rak, and G. Bylin (2002). Ambient level of NO2 augments the inflammatory response to inhaled allergen in asthmatics. *Respiratory Medicine 96*(11), 907 – 917.

Barreca, A., K. Clay, O. Deschênes, M. Greenstone, and J. S. Shapiro (2013). Adapting to climate change: the remarkable decline in the U.S. temperature-mortality relationship over the 20th century. NBER Working Papers 18692.

Basu, R. and J. M. Samet (2002). Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic Reviews 24*(2), 190–202.

Blake, E. S., T. B. Kimberlain, R. J. Berg, J. P. Cangialosi, and J. L. Beven II (2012). *Tropical Cyclone Report.* National Hurricane Center. `www.nhc.noaa.gov/data/tcr/AL182012_Sandy.pdf`, last accessed on November 24, 2013.

Blomberg, A., M. T. Krishna, R. Helleday, M. Söderberg, M.-C. Ledin, F. J. Kelly, A. J. Frew, S. T. Holgate, and T. Sandström (1999). Persistent Airway Inflammation but Accommodated Antioxidant and Lung Function Responses after Repeated Daily Exposure to Nitrogen Dioxide. *American Journal of Respiratory and Critical Care Medicine 159*(2), 536–543.

Broeckaert, F., K. Arsalane, C. Hermans, E. Bergamaschi, A. Brustolin, A. Mutti, and A. Bernard (2000). Serum clara cell protein: A sensitive biomarker of increased lung epithelium permeability caused by ambient ozone. *Respiratory Medicine 108*(6), 533 – 537.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Cameron, C. and D. Miller (2011). Robust Inference with Clustered Data. In A. Ullah and D. E. Giles (Eds.), *Handbook of Empirical Economics and Finance*, pp. 128. CRC Press.

Centers for Disease Control and Prevention (2012). *Carbon Monoxide Poisoning.* `http://ephtracking.cdc.gov/showCarbonMonoxideLanding.action`, last accessed on November 24, 2012.

Chay, K. Y. and M. Greenstone (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics 118*(3), 1121–1167.

Chen, X., H. Hong, and D. Nekipelov (2011). Nonlinear models of measurement errors. *Journal of Economic Literature 49*(4), 901–37.

Clay, K., W. Troesken, and M. R. Haines (2013). Lead and mortality. *Review of Economics and Statistics*. forthcoming.

Currie, J. and D. Almond (2011). Chapter 15: Human capital development before age five. Volume 4, Part B of *Handbook of Labor Economics*, pp. 1315 – 1486.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2013). Do housing prices reflect environmental health risks? Evidence from more than 1600 toxic plant openings and closings. NBER Working Papers 18700.

Currie, J., E. A. Hanushek, E. M. Kahn, M. Neidell, and S. G. Rivkin (2009). Does pollution increase school absences? *The Review of Economics and Statistics 91*(4), 682–694.

Currie, J., M. Neidell, and J. F. Schmieder (2009). Air pollution and infant health: lessons from New Jersey. *Journal of Health Economics 28*(3), 688–703.

Currie, J. and J. F. Schmieder (2009). Fetal exposures to toxic releases and infant health. *American Economic Review 99*(2), 177–83.

Currie, J. and R. Walker (2011). Traffic congestion and infant health: evidence from E-ZPass. *American Economic Journal: Applied Economics 3*(1), 65–90.

Curriero, F. C., K. S. Heiner, J. M. Samet, S. L. Zeger, L. Strug, and J. A. Patz (2002). Temperature and mortality in 11 cities of the Eastern United States. *American Journal of Epidemiology 155*(1), 80–87.

Deschênes, O. (2012). Temperature, human health, and adaptation: A review of the empirical literature. NBER Working Papers 18345.

Deschênes, O. and M. Greenstone (2011). Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the U.S. *American Economic Journal: Applied Economics 3*(4), 152–85.

Deschênes, O., M. Greenstone, and J. Guryan (2009). Climate change and birth weight. *American Economic Review 99*(2), 211–17.

Deschênes, O. and E. Moretti (2009). Extreme weather events, mortality and migration. *The Review of Economics and Statistics 91*(4), 659–681.

Devlin, R. B., L. J. Folinsbee, F. Biscardi, G. Hatch, S. Becker, M. C. Madden, M. Robbins, and H. S. Koren (1997). Inflammation and cell damage induced by repeated exposure of humans to ozone. *Inhalation Toxicology 9*(3), 211–235.

Dye, J. A., J. R. Lehmann, J. K. McGee, D. W. Winsett, A. D. Ledbetter, J. I. Everitt, A. J. Ghio, and D. L. Costa (2001). Acute pulmonary toxicity of particulate matter filter extracts in rats: coherence with epidemiologic studies in utah valley residents. *Environmental Health Perspectives 109*(S3), 395–403.

Edwards, J. H. Y. and C. Langpap (2012). Fuel choice, indoor air pollution and children's health. *Environment and Development Economics 17*(4), 379–406.

Ehrlich, R., J. Findlay, J. Fenters, and D. Gardner (1977). Health effects of short-term inhalation of nitrogen dioxide and ozone mixtures. *Environmental Research 14*(2), 223 – 231.

Environmental Protection Agency (EPA) (2011). *Our Nations Air: Status and Trends Through 2010.* `www.epa.gov/airtrends/2011/`, last accessed on November 14, 2013.

Environmental Protection Agency (EPA) (2013). *National Ambient Air Quality Standards (NAAQS).* `http://www.epa.gov/air/criteria.html`, last accessed on September 14, 2013.

European Environment Agency (2012). *Air Quality in Europe 2011 Report.* available at `http://www.eea.europa.eu/publications/air-quality-in-europe-2011`, last accessed on November 24, 2012.

European Environment Agency (2013). *Air pollution by ozone across Europe during summer 2012.* available at `http://www.eea.europa.eu/publications/air-pollution-by-ozone-across-EU-2012`, last accessed on February 1, 2013.

Evans, M. F. and V. K. Smith (2005). Do new health conditions support mortality-air pollution effects? *Journal of Environmental Economics and Management 50*(3), 496–518.

Federal Institute for Research on Building, Urban Affairs and Spatial Development (2012). *INKAR 2011. Indicators and Maps on Spatial Development.* CD-ROM.

Fung, K., D. Krewski, R. Burnett, and F. Dominici (2005). Testing the harvesting hypothesis by time-domain regression analysis. i: Baseline analysis. *Journal of Toxicology and Environmental Health, Part A 68*(13-14), 1137–1154.

German Federal Statistical Office (2012). *Statistical Yearbook 2012 for the Federal Republic of Germany.* Metzler-Poeschel.

German Federal Statistical Office (2013a). www.destatis.de, last accessed on April 20, 2013.

German Federal Statistical Office (2013b). *Federal Health Monitoring.* www.gbe-bund.de, last accessed on April 20, 2013.

Goklany, I. M. (2009). Deaths and death rates from extreme weather events: 1900-2008. *Journal of American Physicians and Surgeons 14*(4), 102–109.

Greenstone, M. and R. Hanna (2011). Environmental regulations, air and water pollution, and infant mortality in India. NBER Working Papers 17210.

Greenstone, M. and B. K. Jack (2013). Envirodevonomics: A research agenda for a young field. NBER Working Papers 19426.

Hackney, J. D., W. S. Linn, J. G. Mohler, F. E. Pedersen, P. Breisacher, and A. Russo (1975). Human health effects of air pollutants. *Archives of Environmental Health: An International Journal 30*(8), 379–384.

Hammitt, J. and Y. Zhou (2006). The economic value of air-pollution-related health risks in China: a contingent valuation study. *Environmental & Resource Economics 33*(3), 399–423.

Hanigan, I., G. Hall, and K. B. Dear (2006). A comparison of methods for calculating population exposure estimates of daily weather for health research. *International Journal of Health Geographics 5*(38), 1–6 (online version).

Hanna, R., E. Duflo, and M. Greenstone (2012). Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves. NBER Working Papers 18033.

Heal, G. and A. Millner (2013). Uncertainty and decision in climate change economics. Working Paper 18929, National Bureau of Economic Research.

Horstman, D., E. J. Seal, L. Folinsbee, P. Ives, and L. Roger (1988). The relationship between exposure duration and sulfur dioxide-induced bronchoconstriction in asthmatic subjects. *International Journal of Health Geographics 49*(1), 38–47.

Hübler, M., G. Klepper, and S. Peterson (2008). Costs of climate change; the effects of rising temperatures on health and productivity in Germany. *Journal of Ecological Economics 68*, 381–393.

Hyslop, D. R. and G. W. Imbens (2001). Bias from classical and other forms of measurement error. *Journal of Business & Economic Statistics 19*(4), 475–481.

Intergovernmental Panel on Climate Change (IPCC) (2007). *Climate Change 2007: Synthesis Report.* `http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_synthesis_report.htm`, last accessed on January 25, 2013.

Jäppinen, P., V. Vilkka, O. Marttila, and T. Haahtela (1990). Exposure to hydrogen sulphide and respiratory function. *47*(12), 824–828.

Karlsson, M. and M. Schmitt (2011). Only in the heat of the moment? A study of the relation between weather and mortality in Germany. Health, Econometrics and Data Group (HEDG) Working Papers 11/27, Department of Economics, University of York.

Kerr, H., T. Kulle, M. McIlhany, and P. Swidersky (1979). Effects of nitrogen dioxide on pulmonary function in human subjects: An environmental chamber study. *Environmental Research 19*(2), 392 – 404.

Kniesner, T., W. Viscusi, and J. Ziliak (2010). Policy relevant heterogeneity in the value of statistical life: New evidence from panel data quantile regressions. *Journal of Risk and Uncertainty 40*(1), 15–31.

Knittel, C. R., D. L. Miller, and N. J. Sanders (2011). Caution, drivers! Children present: Traffic, pollution, and infant health. NBER Working Papers 17222.

Lagravinese, R., L. Habin, F. Moscone, and E. Tosetti (2013). The impact of air pollution on hospital admissions: evidence from Italy. Departmental Working Papers of Economics 170, University Roma Tre, Department of Economics.

Lawther, P., A. Macfarlane, R. Waller, and A. Brooks (1975). Pulmonary function and sulphur dioxide, some preliminary findings. *Environmental Research 10*(3), 355 – 367.

Li, Z., F. Niu, J. Fan, Y. Liu, D. Rosenfeld, and Y. Ding (2011). Meta-analysis of association between particulate matter and stroke attack. *Nature Geoscience 4*, 888–894.

Lippmann, M. (1989). Health effects of ozone a critical review. *JAPCA 39*(5), 672–695.

Liu, J., J. A. Curry, H. Wang, M. Song, and R. M. Horton (2012). Impact of declining Arctic sea ice on winter snowfall. *Proceedings of the National Academy of Sciences of the United States of America (PNAS) 109*(11), 4074–4079.

Mayr, F. B., A. Spiel, J. Leitner, C. Marsik, P. Germann, R. Ullrich, O. Wagner, and B. Jilma (2005). Effects of carbon monoxide inhalation during experimental endotoxemia in humans. *American Journal of Respiratory and Critical Care Medicine 171*(4), 354–360.

Moretti, E. and M. Neidell (2011). Pollution, health, and avoidance behavior: evidence from the ports of Los Angeles. *Journal of Human Resources 46*(1), 154–175.

Neidell, M. (2009). Information, avoidance behavior, and health: the effect of ozone on asthma hospitalizations. *Journal of Human Resources 44*(2), 450–478.

Nilsson, P. (2009). The long-term effects of early childhood lead exposure: evidence from the phase-out of leaded gasoline. `https://sites.google.com/site/nilssonjanpeter/`, last accessed on January 25, 2013.

OECD (2012). *OCED Health StatExtracts 2012*. `http://stats.oecd.org`, last accessed on November 20, 2012.

OECD (2013). OECD Health Data: Health status. `http://dx.doi.org/10.1787/data-00540-en`, last accessed on August 16, 2013.

Omaye, S. T. (2002). Metabolic modulation of carbon monoxide toxicity. *Toxicology 180*(62), 139–150.

Penney, D. G. (2000). *Carbon Monoxide Toxicity.* CRC Press.

Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton (2000). The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics 16*(2-3), 123–146.

Pope III, C. A., R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA 287*(9), 1132–1141.

Potter, Thomas D.; Colman, B. R. (2002). *Handbook of Weather, Climate and Water: Atmospheric Chemistry, Hydrology and Societal Impacts* (1st ed.). Wiley-Interscience.

Quah, E. and T. L. Boon (2003). The economic cost of particulate air pollution on health in Singapore. *Journal of Asian Economics 14*(1), 73–90.

Rabl, A. (2005). Air pollution mortality: Harvesting and loss of life expectancy. *Journal of Toxicology and Environmental Health, Part A 68*(13-14), 1175–1180.

Robinson, A., D. Gyrd-Hansen, P. Bacon, R. Baker, M. Pennington, and C. Donaldson (2013). Estimating a WTP-based value of a QALY: The 'chained' approach. *Social Science & Medicine 92*(C), 92–104.

Sandstrom, T., N. Stjernberg, A. Eklund, M. Ledin, L. Bjermer, B. Kolmodin-Hedman, K. Lindstrom, L. Rosenhall, and T. Angstrom (1991). Inflammatory cell response in bronchoalveolar lavage fluid after nitrogen dioxide exposure of healthy subjects: a dose-response study. *European Respiratory Journal 4*(3), 332–339.

Schlenker, W. and W. R. Walker (2011). Airports, air pollution, and contemporaneous health. NBER Working Papers 17684.

Schwartz, J., J. M. Samet, and J. A. Patz (2004). Hospital admissions for heart disease: The effects of temperature and humidity. *Epidemiology 15*(6), 755–761.

Seinfeld, John H.; Pandis, S. N. (2006). *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (2nd ed.). Wiley-Interscience.

Shah, P. S. and T. Balkhair (2011). Air pollution and birth outcomes: a systematic review. *Environment International 37*(2), 498 – 516.

Shiroiwa, T., Y.-K. Sung, T. Fukuda, H.-C. Lang, S.-C. Bae, and K. Tsutani (2010). International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness? *Health Economics 19*(4), 422–437.

SOEP (2012). Data for the years: 1984-2011. doi:10.5684/soep.v28, version 29.

Stern, N. (2006). Stern review on the economics of climate change. http://webarchive.nationalarchives.gov.uk/+/http://www.hm-treasury.gov.uk/sternreview_index.htm, last accessed on March 1, 2013.

Stewart, R. D., J. E. Peterson, E. D. Baretta, R. T. Bachand, M. J. Hosko, and A. A. Herrmann (1970). Experimental human exposure to carbon monoxide. *Archives of Environmental Health: An International Journal 21*(2), 154–164.

United States Census Bureau (2013). *County Totals Datasets: Population, Population Change and Estimated Components of Population Change: April 1, 2010 to July 1, 2012.* http://www.census.gov/popest/data/counties/totals/2012/CO-EST2012-alldata.html, last accessed on August 27, 2013.

United States Global Change Research Program (2009). Global Climate Change Impacts in the United States. In J. M. M. Thomas R. Karl and T. C. Peterson (Eds.), *United States Global Change Research Program.* Cambridge University Press.

U.S. Census Bureau (2012). *US Census 2010.* http://2010.census.gov/2010census/data/, last accessed on November 20, 2012.

van den Berg, Gerard J.; Lindeboom, M. P. F. (2006). Economic conditions early in life and individual mortality. *American Economic Review 96*(1), 290–302.

Villena, M., E. Monroy, and M. Villena (2008). Is relative risk due to air pollution linked to income? Infant and elderly mortality in Santiago, Chile. Working Papers 6, Departamento de Economía, Universidad de Concepción.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP)--Scope, Evolution and Enhancements. *Journal of Applied Social Science Studies (Schmollers Jahrbuch) 127*(1), 139–169.

World Health Organization (2000). *Air Quality Guidelines for Europe* (second edition ed.). available at www.euro.who.int/document/e71922.pdf, last accessed on February 1, 2013.

World Health Organization (2003). *Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide.* available at www.euro.who.int/document/e79097.pdf, last accessed on February 1, 2013.

World Health Organization (2011). *Air quality and health: Fact sheet No 313 Updated September 2011.* available at http://www.who.int/mediacentre/factsheets/fs313/en/, last accessed on February 1, 2013.

Wright, E. S., D. Dziedzic, and C. S. Wheeler (1990). Cellular, biochemical and functional effects of ozone: new research and perspectives on ozone health effects. *Toxicology Letters 51*(2), 125 – 145.

Zivin, J. S. G. and M. J. Neidell (2013). Environment, health, and human capital. *Journal of Economic Literature 51*(3), 689–730.

**Figure 1:** Distribution of Official German Ambient Weather and Pollution Monitors

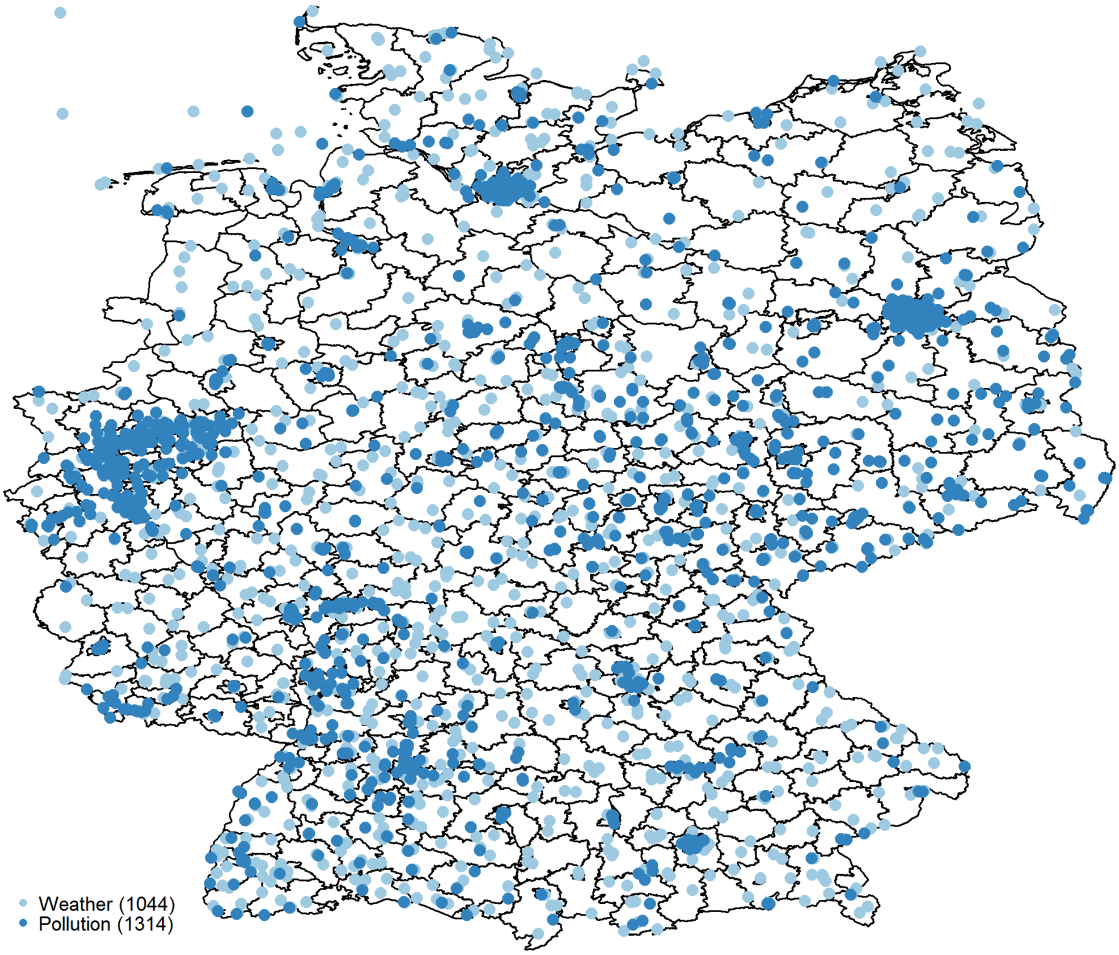
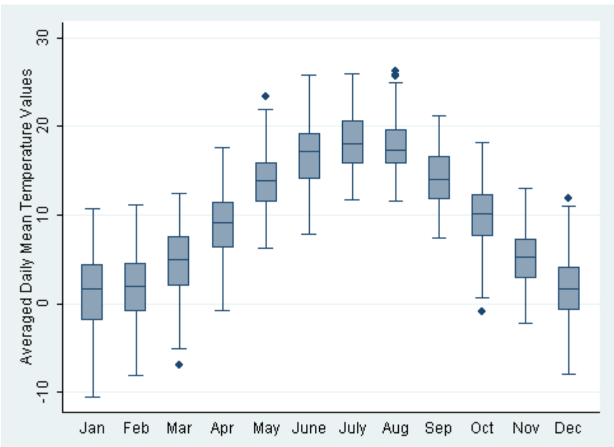



**Figure 2:** (a) Boxplot of Mean Temperature Over Month and (b) Temperature Variation Over 10 Years



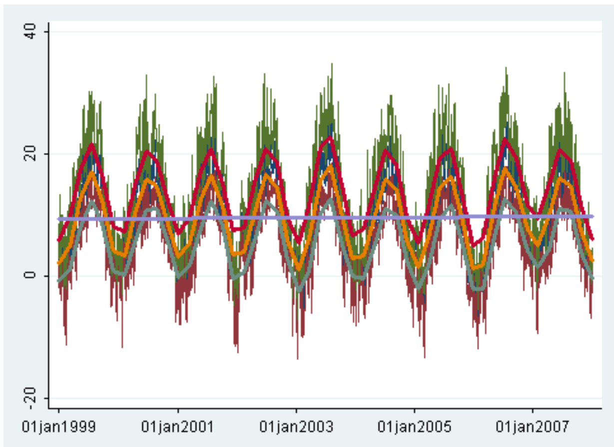



**Figure 3:** Distributions of Max. and Min. Temperatures and Number of Hot and Cold Days

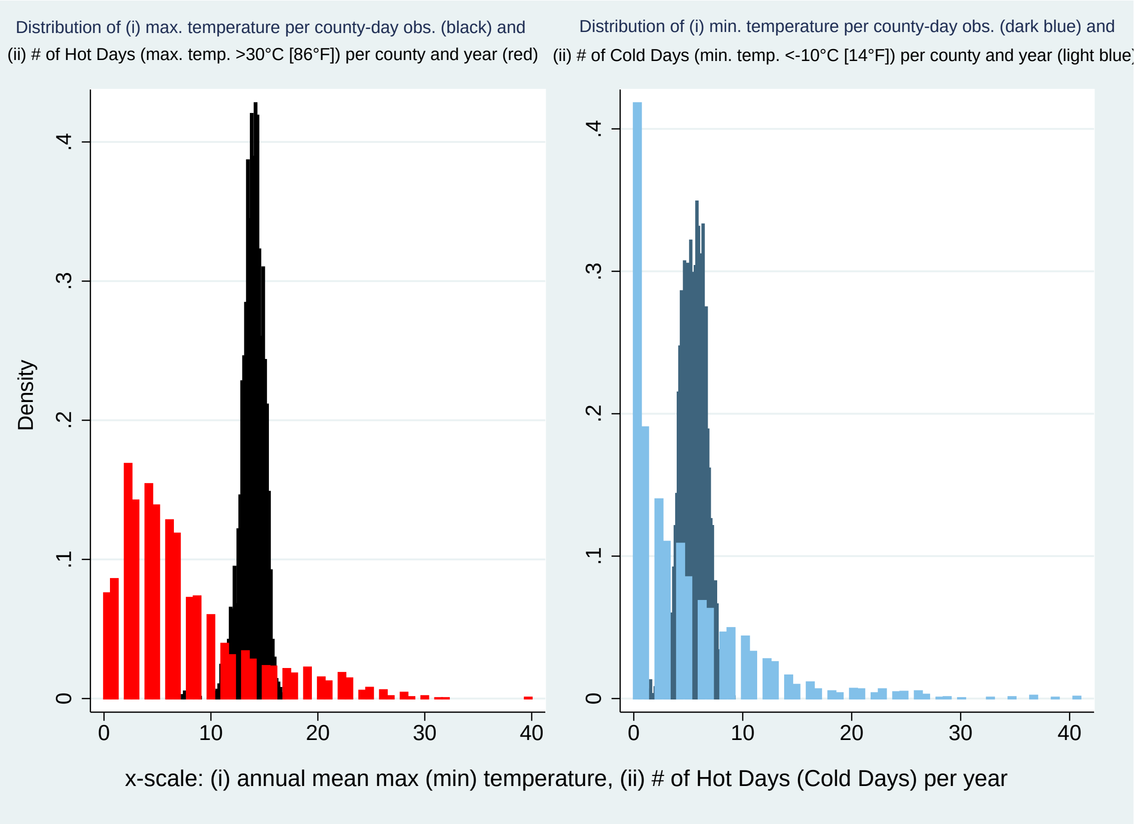



**Figure 4:** Distribution of Ozone ($O_3$) Concentration and Non-Compliance Days: Identifying Variation

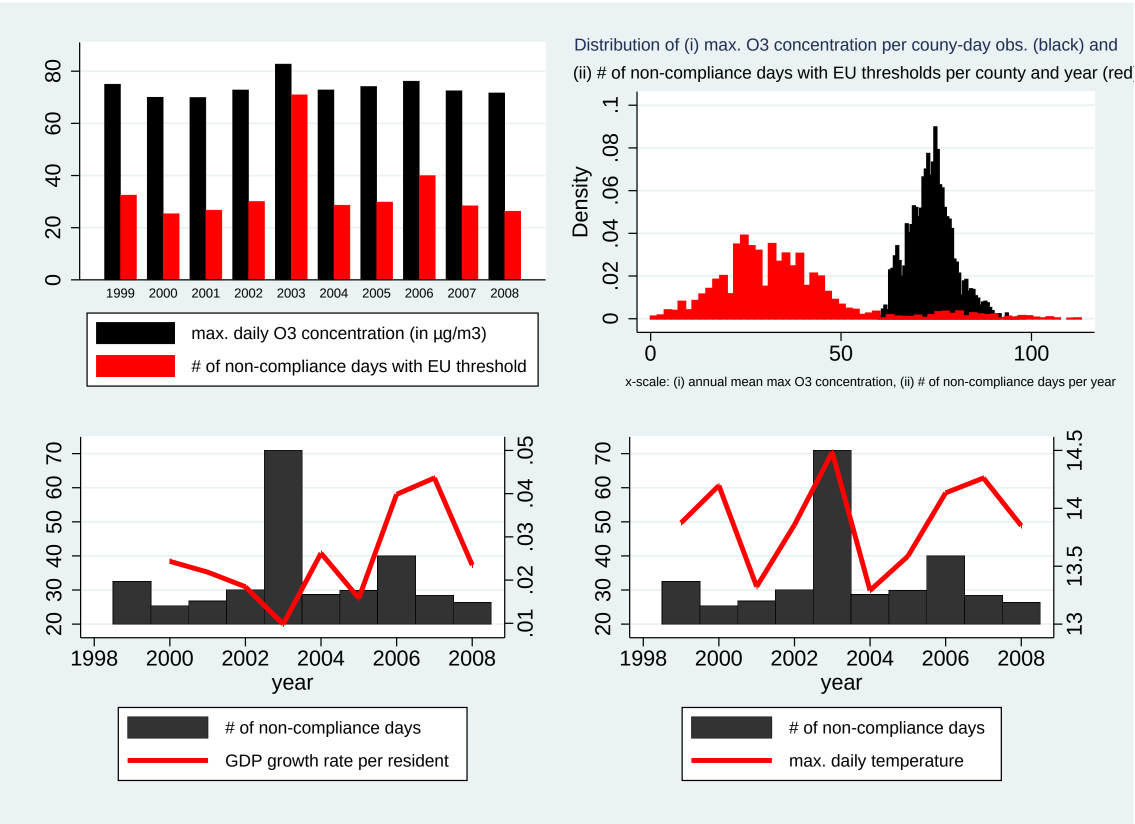


**Figure 5:** Nonparametric Relationship Between Temperature, Pollution and Hospitalizations

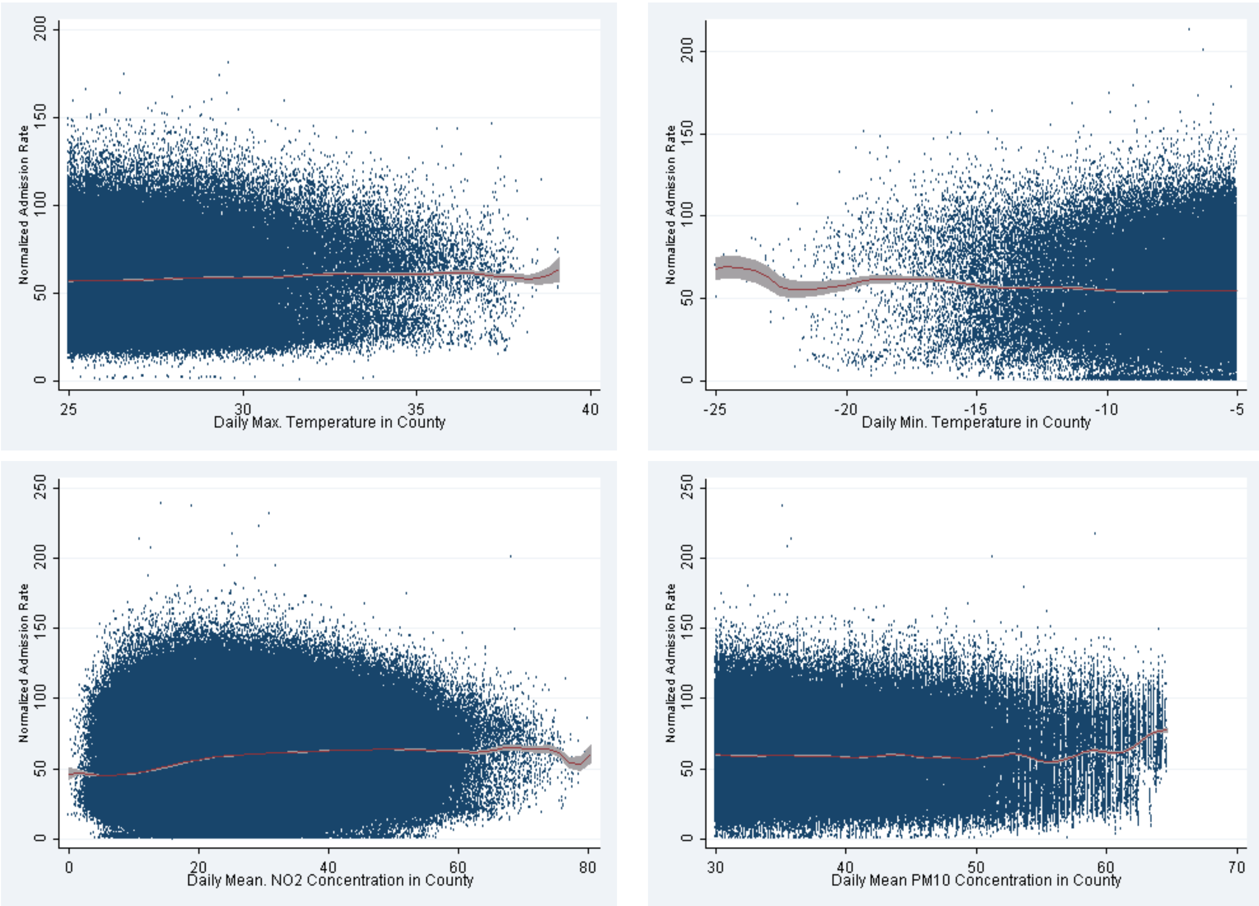


**Figure 6:** Semiparametric Relationship Between Temperature, Pollution and Hospital Amissions Net of Seasonal and County Effects (*Unconditional Approach I*, equation (3))

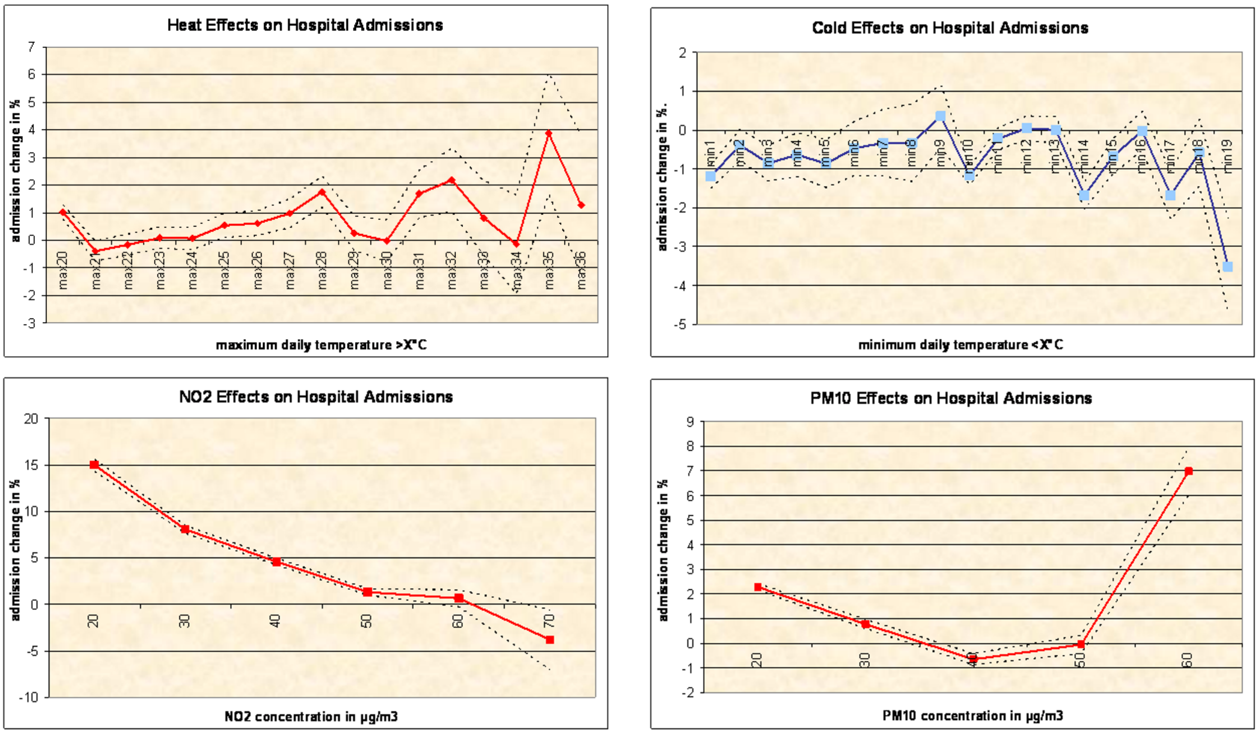



**Figure 7:** Effect of Heat on Cause-Specific Hospitalizations and Mortality (*Conditional Approach II*, equation (3))

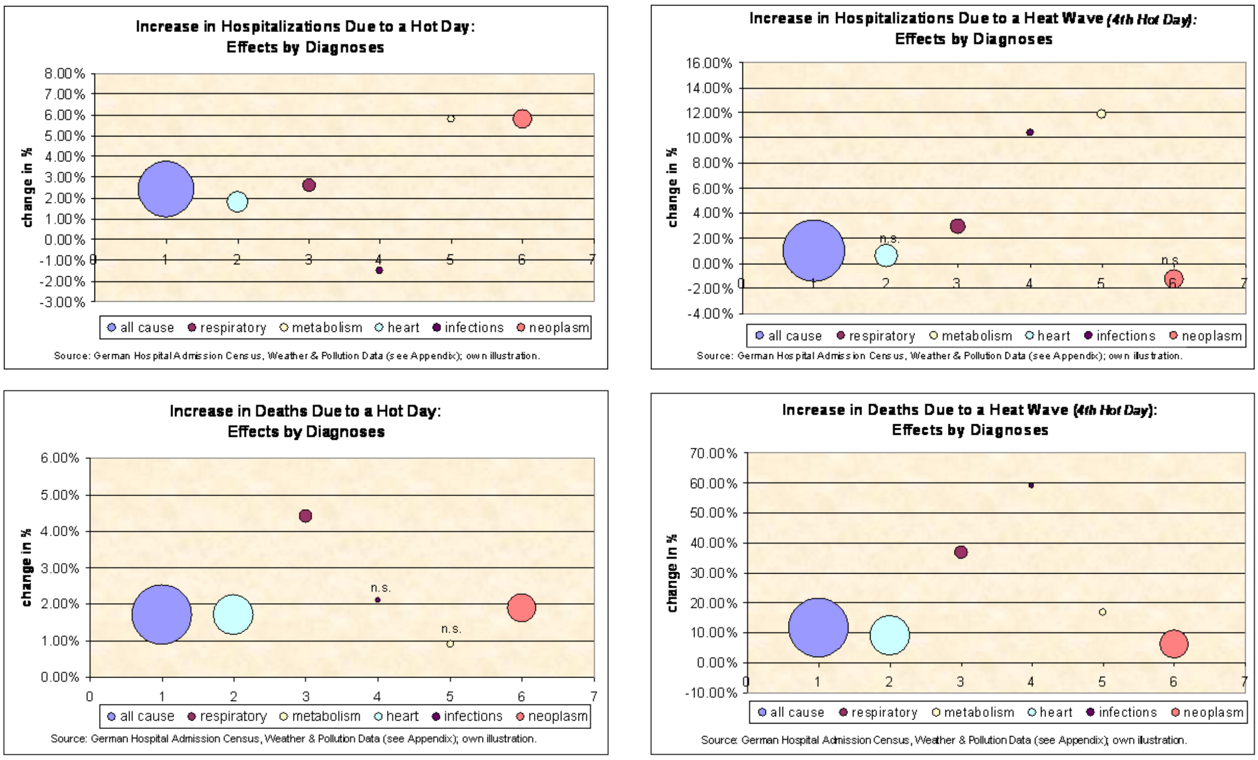


**Figure 8:** Mortality Rates 4 Days Before and After a *Hot Day* Net of Seasonal and County Effects (*Unconditional Approach I*, equation (3))

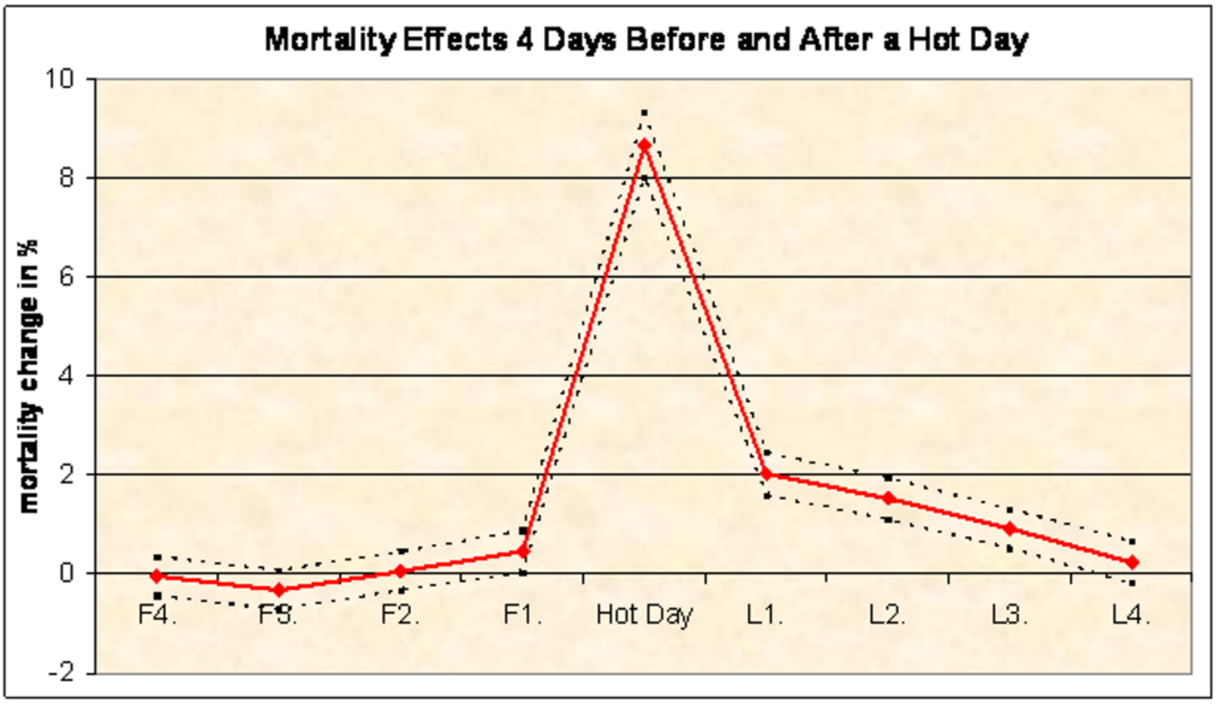



**Figure 9:** Age Structure of Hospital Admissions on *Hot Days*: Plotted Interaction Terms Between *Hot Day* and Age Groups (*Unconditional Approach I*, equation (3))

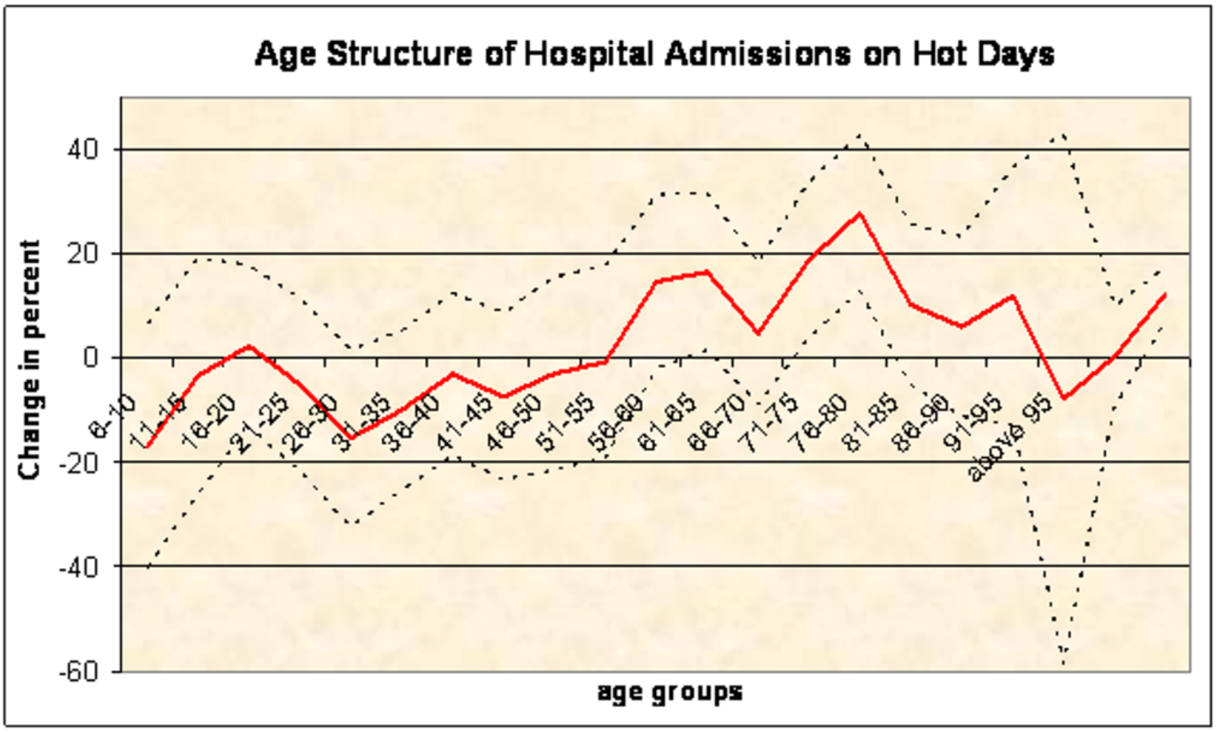


**Table 1:** The Impact of Extreme Temperature on Health: Conditional and Unconditional on Contemporaneous Pollution and Weather Conditions

| **Panel A: Hospitalizations** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
|---|---|---|---|---|---|
| *Hot Day* | 3.1063*** | | | | 1.4075*** |
| | (0.1037) | | | | (0.1389) |
| *Heat Wave Day* | | 3.5085*** | | | 0.5781** |
| | | (0.2261) | | | (0.2559) |
| *Cold Day* | | | -1.2093*** | | 0.7235*** |
| | | | (0.1271) | | (0.1899) |
| *Cold Wave Day* | | | | -4.4559*** | -5.5656*** |
| | | | | (0.3533) | (0.5137) |
| change in % | +5.4% | +6.1% | -2.1% | -7.7% | |
| N | 1,590,454 | 1,590,454 | 1,590,454 | 1,590,454 | 1,429,928 |
| $R^2$ | 0.4030 | 0.4028 | 0.4028 | 0.4030 | 0.4030 |
| **Panel B: Mortality** | | | | | |
| *Hot Day* | 0.2939*** | | | | 0.05069*** |
| | (0.0098) | | | | (0.0114) |
| *Heat Wave Day (>3 Hot Days)* | | 0.5935*** | | | 0.3476*** |
| | | (0.0308) | | | (0.0311) |
| *Cold Day* | | | 0.0246* | | -0.0118 |
| | | | (0.0130) | | (0.0167) |
| *Cold Wave Day (>3 Cold Days)* | | | | 0.0455 | 0.0223 |
| | | | | (0.0330) | (0.0343) |
| change in % | +9.8% | +19.8% | +0.8% | +1.5% | |
| N | 1,518,000 | 1,518,000 | 1,518,000 | 1,518,000 | 1,364,921 |
| $R^2$ | 0.0172 | 0.0170 | 0.0169 | 0.0166 | 0.0177 |
| County, week, & month-year fixed effects | yes | yes | yes | yes | yes |
| Age, gender, county & hospital controls | yes | yes | yes | yes | yes |
| 7 continuous weather measures + 15 interactions | no | no | no | no | yes |
| 5 continuous pollution measures + 5 quadratic + 5 cubic + 10 interaction terms | no | no | no | no | yes |
| 3 pollution EU Non-Compliance Indicators ($O_3$, $NO_2$, $PM_{10}$) | no | no | no | no | yes |
| 25 interaction terms weather & pollution | no | no | no | no | yes |

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses are clustered at the county level. Data sources are discussed in Section 2. All specifications estimate the model in equation (3) by OLS. Each column in each panel represents one model. Models only differ by the sets of covariates included as indicated. In Panel A, the dependent variable is the daily incidence of hospital admissions per 100,000 population at the county level (mean: 57.99, see Appendix A). In Panel B, the dependent variable is the daily mortality rate per 100,000 population at the county level (mean: 2.99, see Appendix B). For example, according to column (1) in Panel A, a *Hot Day*—defined as the max. temperature exceeding 30 ˚C (86 ˚F)—triggers 3.1 additional hospital admissions per 100,000 pop. This represents an increase by 5.4% and translates into 2,542 additional admissions for the whole of Germany with its 82 million inhabitants, or roughly 1 additional daily admission per hospital. As shown in Table C1, about 2% of all days are *Hot Days* in Germany, between 7 and 8 per year. Columns (1) to (4) show the *Unconditional Approach*, where the regressor of interest absorbs the effects of all contemporaneous weather and pollution conditions, whereas column (5) shows the *Conditional Approach* that nets out all contemporaneous weather and pollution conditions. Weather conditions are specified and defined as explained in Section 2 and Appendix C. Column (5) has fewer observations since $PM_{10}$ data for 2000 is not available.

**Table 2:** The Impact of High Pollution Levels ("EU Non-Compliance Days") on Health: Conditional and Unconditional on Contemporaneous Weather Conditions and Other Pollutants

| **Panel A: Hospitalizations** | **(1)** | **(2)** | **(3)** | **(4)** |
|---|---|---|---|---|
| *$NO_2$ EU Non-Compliance Day* | 5.0272*** | | | 0.5595*** |
| | (0.1230) | | | (0.1599) |
| *$O_3$ EU Non-Compliance Day* | | 0.4064*** | | -1.6547*** |
| | | (0.0649) | | (0.1241) |
| *$PM_{10}$ EU Non-Compliance Day* | | | 0.6842*** | 0.2725 |
| | | | (0.0838) | (0.1723) |
| change in % | +8.7% | +0.7% | +1.2% | |
| N | 1,590,454 | 1,590,454 | 1,590,454 | 1,429,928 |
| $R^2$ | 0.4030 | 0.4028 | 0.4028 | 0.4664 |
| **Panel B: Mortality** | | | | |
| *$NO_2$ EU Non-Compliance Day* | 0.0375*** | | | 0.0083 |
| | (0.0041) | | | (0.0068) |
| *$O_3$ EU Non-Compliance Day* | | 0.1348*** | | -0.006 |
| | | (0.0051) | | (0.0069) |
| *$PM_{10}$ EU Non-Compliance Day* | | | 0.0741*** | 0.0039 |
| | | | (0.0087) | (0.0148) |
| change in % | +1.3% | +4.5% | +2.5% | |
| N | 1,518,000 | 1,518,000 | 1,518,000 | 1,364,921 |
| $R^2$ | 0.0166 | 0.0171 | 0.0166 | 0.0486 |
| County, week, & month-year fixed effects | yes | yes | yes | yes |
| Age, gender, county & hospital controls | yes | yes | yes | yes |
| 7 continuous weather measures + 15 interactions | no | no | no | yes |
| 5 continuous pollution measures + 5 quadratic + 5 cubic + 10 interaction terms | no | no | no | yes |
| 4 extreme temp. indicators (*Hot Day*, *Heat Wave*, *Cold Day*, *Cold Wave*) | no | no | no | yes |
| 25 interaction terms weather & pollution | no | no | no | yes |

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses are clustered at the county level. Data sources are discussed in Section 2. All specifications estimate the model in equation (3) by OLS. Each column in each panel represents one model. Models only differ by the sets of covariates included as indicated. In Panel A, the dependent variable is the daily incidence of hospital admissions per 100,000 population at the county level (mean: 57.99, see Appendix A). In Panel B, the dependent variable is the daily mortality rate per 100,000 population at the county level (mean: 2.99, see Appendix B). For example, according to column (1) in Panel B, a $NO_2$ EU Non-Compliance Day—defined as a day with the average $NO_2$ level exceeding the EU alert threshold of 40 $\mu g/m^3$—triggers 0.0375 additional deaths per 100,000 pop. This represents an increase by 1.3% and translates into 31 additional deaths for the whole of Germany. As shown in Table D1, about 12% of all days are $NO_2$ EU Non-Compliance Days in Germany, 44 per year. Columns (1) to (3) show the *Unconditional Approach*, where the regressor of interest absorbs the effects of all contemporaneous weather and pollution conditions, whereas column (4) shows the *Conditional Approach* that nets out all contemporaneous weather and pollution conditions. Pollution conditions are specified and defined as explained in Section 2 and Appendix D. Column (4) has fewer observations since $PM_{10}$ data for 2000 is not available.

**Table 3:** The Impact of Extreme Temperature on Health by Diagnoses: Conditional on Weather Conditions and Other Pollutants

| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
|---|---|---|---|---|---|---|
| **Panel A: Hospitalizations** | all causes | heart | respiratory | infections | metabolism | neoplasm |
| *Hot Day* | 1.4075*** | 0.1680*** | 0.0946*** | -0.0195** | 0.0954*** | 0.3805*** |
| | (0.1389) | (0.0275) | (0.0164) | (0.0093) | (0.0110) | (0.0276) |
| *Heat Wave Day (>3 Hot Days)* | 0.5781** | 0.0548 | 0.1061*** | 0.1403*** | 0.1960*** | -0.0825 |
| | (0.2559) | (0.0527) | (0.0342) | (0.0218) | (0.2610) | (0.0553) |
| *Cold Day* | 0.7235*** | 0.1823*** | 0.0588*** | 0.0244** | -0.0027 | 0.0435 |
| | (0.1899) | (0.0369) | (0.0224) | (0.0117) | (0.0134) | (0.0394) |
| *Cold Wave Day (>3 Cold Days)* | -5.5656*** | -0.8324*** | -0.2908*** | -0.0338 | -0.2035*** | -1.0375*** |
| | (0.5137) | (0.0998) | (0.0589) | (0.0253) | (0.0304) | (0.1066) |
| share of all causes | 100% | 15.7% | 6.2% | 2.3% | 2.8% | 11.3% |
| N | 1,429,928 | 1,429,928 | 1,429,928 | 1,429,928 | 1,429,928 | 1,429,928 |
| $R^2$ | 0.4030 | 0.3869 | 0.1967 | 0.0691 | 0.1733 | 0.3958 |
| **Panel B: Mortality** | all causes | heart | respiratory | infections | metabolism | neoplasm |
| *Hot Day* | 0.0507*** | 0.0233*** | 0.0085*** | 0.0008 | 0.0009 | 0.0143** |
| | (0.0114) | (0.0079) | (0.0036) | (0.0014) | (0.0023) | (0.0062) |
| *Heat Wave Day (>3 Hot Days)* | 0.3476*** | 0.1248*** | 0.0708*** | 0.0221*** | 0.0165*** | 0.0467*** |
| | (0.0311) | (0.0193) | (0.0091) | (0.0039) | (0.0053) | (0.0138) |
| *Cold Day* | -0.0118 | -0.0106 | -0.0042 | 0.0011 | 0.0041 | 0.0017 |
| | (0.0167) | (0.0124) | (0.0044) | (0.0021) | (0.0028) | (0.0085) |
| *Cold Wave Day (>3 Cold Days)* | 0.0223 | 0.0033 | -0.0007 | 0.0023 | 0.0097 | -0.0152 |
| | (0.0343) | (0.0264) | (0.0106) | (0.0043) | (0.0074) | (0.0177) |
| share of all causes | 100% | 46.3% | 6.4% | 1.3% | 2.4% | 25.7% |
| N | 1,364,921 | 1,364,921 | 1,364,921 | 1,364,921 | 1,364,921 | 1,364,921 |
| $R^2$ | 0.0177 | 0.0486 | 0.0139 | 0.0023 | 0.0028 | 0.0426 |
| County, week, & month-year fixed effects | yes | yes | yes | yes | yes | yes |
| Age, gender, county & hospital ind. | yes | yes | yes | yes | yes | yes |
| 7 cont. weather ind. + 15 interactions | yes | yes | yes | yes | yes | yes |
| 5 cont. pollution ind. + 5 quadratic + 5 cubic + 10 interaction terms | yes | yes | yes | yes | yes | yes |
| 3 pollution EU Non-Compliance Indicators | yes | yes | yes | yes | yes | yes |
| 25 interactions weather & pollution | yes | yes | yes | yes | yes | yes |

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses are clustered at the county level. Data sources are discussed in Section 2. All specifications estimate the model in equation (3) by OLS. Each column in each panel represents the *Conditional Approach* that nets out all contemporaneous weather and pollution conditions. In Panel A, the dependent variable is the daily incidence of hospital admissions per 100,000 population at the county level (mean: 57.99, see Appendix A). In Panel B, the dependent variable is the daily mortality rate per 100,000 population at the county level (mean: 2.99, see Appendix B). For example, according to column (1) in Panel A, a Hot Day—defined as the max. temperature exceeding 30 °C (86 °F)—triggers 1.4 additional hospital admissions per 100,000 pop. This represents an increase by 2.4% and translates into 1,148 additional admissions for the whole of Germany. Weather conditions are specified and defined as explained in Section 2 and Appendix C.

**Table 4:** The Impact of Extreme Heat on Normalized Hospitalizations: Robustness Checks

| **Panel A** | **cluster at state level (1)** | **2-way cluster (2)** | **linear & quadratic time trend(3)** | **linear & quadratic state time trends (4)** | **linear & quadratic county time trends [2006-2008](5)** |
|---|---|---|---|---|---|
| *Hot Day* | 3.1063*** | 3.1063*** | 3.1063** | 2.7266*** | 1.9695*** |
| | (0.1918) | (0.1918) | (0.1037) | (0.0954) | (0.1942) |
| change in % | +5.4% | +5.4% | +5.4% | +4.7% | +3.4% |
| N | 1,590,454 | 1,590,454 | 1,590,454 | 1,590,454 | 1,590,454 |

| **Panel B** | **×(temp.>32 ˚C) (1)** | **×weekend (2)** | **×warm region (3)** | **×cold region (4)** | **aggregated at annual level (5)** |
|---|---|---|---|---|---|
| *Hot Day*×[column header] | 0.6848*** | -0.1238 | -0.9710*** | 0.5438** | |
| | (0.0519) | (0.1671) | (0.1642) | (0.2429) | |
| *Hot Day* | 1.8235*** | 2.3651*** | 3.4272*** | 3.0156*** | 0.0298** |
| | (0.1026) | (0.0944) | (0.1305) | (0.1103) | (0.0133) |
| max. daily temp. | 0.1737*** | | | | |
| | (0.0043) | | | | |
| Weekend | | -31.5713*** | | | |
| | | (0.2838) | | | |
| N | 1,590,454 | 1,590,454 | 1,590,454 | 1,590,454 | 4,356 |

* p<0.1, ** p<0.05, *** p<0.01; standard errors in parentheses are clustered at the county level except for columns (1) of Panel A which clusters at the state and column (2) of Panel A which clusters at the county and day level (2-way cluster). Data sources are discussed in Section 2. Each column in each panel represents one *Unconditional Approach* model, i.e., the model does not control for other contemporaneous weather and pollution conditions. The dependent variable is always the hospitalization rate (mean: 57.99, see Table A1); the reference estimate is the one in Column (1) of Table 1. All specifications estimate a model similar to equation (3) by OLS. More precisely, in Panel A, column (3) adds a nation-wide linear and quadratic time trend. Column (4) adds state-level time trends and column (5) adds county-level time trends (for 2006-2008 only because of computer memory constraints). The first column in Panel B adds a continuous measure for the maximum daily temperature as well as an interaction term between the maximum daily temperature and the average maximum *Hot Day* temperature (31.9 ˚C (89.4 ˚F)). Thus, the interaction term estimates the marginal effect of one temperature degree above 31.9 ˚C. Column (2) of Panel B adds a weekend dummy and interacts it with *Hot Day*. Column (3) and (4) add a dummy for *warm region* (mean annual county-level temperature falls into the highest temperature quartile for Germany (>10.2 ˚C (50 ˚F))) and *cold region* (mean annual temperature below the lowest temperature quartile (<9.0 ˚C (48 ˚F))) as well as their interactions with *Hot Day*. Column (5) in Panel B aggregates the data at the annual county level and estimates the impact of one additional *Hot Day* per year.

**Table 5:** The Monetized Health Effects of One Additional Hot Day

| | ***Hospitalizations*** | | | | ***Mortality*** | ***Total*** | |
|---|---|---|---|---|---|---|---|
| | **Health Care Expenditures (1)** | **Lost Labor (2)** | **Lost QALYs (upper bound) (3)** | **Lost QALYs (lower bound) (4)** | **remaining life years (5)** | **(1)-(3) +(5)** | **(1)+(2) +(4)+(5)** |
| Unconditional Approach, daily county level, (Approach I) | 19,000×€ 500 =€ 9.5m | 0.5×19,000×€ 150 =€ 1.4m | (19,000/365)×€ 100,000×1.0 =€ 5.2m | (19,000/365)×€ 100,000×0.5 =€ 2.6m | 270×1×€ 100,000 =€ 27m | ∼€ 43.1m | ∼€ 40.5m |
| Conditional Approach, daily county level, (Approach II) | 8,000×€ 500 =€ 4.0m | 0.5×8,000×€ 150 =€ 0.6m | (8,000/365)×€ 100,000×1.0 =€ 2.2m | (8,000/365)×€ 100,000×0.5 =€ 1.1m | 78×1×€ 100,000 =€ 7.8m | ∼€ 14.6m | ∼€ 13.5m |
| Unconditional Approach, annual county level | 180×€ 500 =€ 90,000 | 0.5×180×€ 150 =€ 13,000 | (180/365)×€ 100,000×1.0 =€ 50,000 | (180/365)×€ 100,000×0.5 = € 25,000 | 2×30×€ 100,000 =€ 6m | ∼€ 6.2m | ∼€ 6.1m |

The table shows the health-related costs associated with one *Hot Day*. The first row is based on the *Unconditional Approach I* that does not consider additional weather or pollution control variables other than *Hot Day*. The underlying models that estimate how many hospital days are triggered by a *Hot Day* are similar to equation (3) but use *Hospital Days* as dependent variable (see Appendix A1). The second row is based on the *Conditional Approach II* and a saturated model that simultaneously considers a rich set of weather and pollution controls. These first two approaches are based on daily county-level observations and include potential harvesting effects. The third row considers harvesting and is based on aggregated annual county-level data (see column (5) in Panel B of Table 4). Column (1) makes use of the fact that an average hospital day in Germany is reimbursed with € 500. Column (2) considers that the average daily wage in Germany is € 150. Columns (3) and (4) assume that 365 hospital days equal a loss of 1 and 0.5 QALYs, respectively. One QALY is evaluated with € 100,000. Column (5) assumes that the remaining life expectancy for those who die during heat events is 1 year for rows one and two (excluding harvesting) and 30 years for row three (including harvesting). We do not discount the monetized health-related loss in welfare. Under *Approach I* in the first row, a discount rate of 2.5% would reduce the costs over 80 years from € 3.2bn to € 1.4bn or € 17 per resident. The table does not consider health issues that lead to outpatient treatments. Neither does it consider health-related avoidance behavior costs or adverse health effects due to tornados, hurricanes, or floods.

# Appendix A: *Hospital Admission Census*

The first register dataset is the *Hospital Admission Census*. It contains the universe of hospital admissions from 1999 to 2008. This is a restricted access dataset provided by the German Federal Statistical Office (*Statistische Ämter des Bundes und der Länder*). We observe every single of the more than 17 million annual hospital admissions. The data contain the following information on the individual admission level:

- age in 18 age groups
  *(0-2 yrs., 3-5 yrs., 6-9 yrs., 10-14 yrs.,..., 60-64 yrs., 65-75 yrs., >75 yrs.)*
- gender *(binary indicator)*
- county of residence [*between 442 (1999) and 413 (2008) counties*]
- day of admission
- length of stay *(censored at 85 days)*
- died in hospital *(binary indicator)*
- primary diagnosis *(ICD-10, 3 digit)*
- surgery needed *(binary indicator)*
- primary hospital department *(43 categories)*
- #hospital beds *(12 categories)*
- hospital location *(federal state level; 16 states)*
- private hospital *(binary indicator)*
- hospital identifier

As described in Section 2.5, we normalize, aggregate, and merge this dataset with the other datasets at the day-county level. As such, we obtain the following descriptive statistics for the hospital admission data:

**Table A1:** Hospital Admission Census: Dependent Variables per 100,000 pop. (Daily County-Level, 1999-2008)

| **Variable** | **Mean** | **Std. Dev.** | **N** |
|---|---|---|---|
| All-cause hospitalization rate | 57.99 | 25.71 | 1,590,454 |
| Hospital days | 488.87 | 267.21 | 1,590,454 |
| | | | |
| Cardiovascular hospitalization rate | 9.1116 | 4.9216 | 1,590,454 |
| Cardiovascular hospital days | 83.69 | 55.96 | 1,590,454 |
| Cardiovascular deaths | 0.4532 | 0.6423 | 1,590,454 |
| | | | |
| Respiratory hospitalization rate | 3.6013 | 2.5195 | 1,590,454 |
| Respiratory hospital days | 27.93 | 23.39 | 1,590,454 |
| Respiratory deaths | 0.1557 | 0.3685 | 1,590,454 |
| | | | |
| Infectious hospitalization rate | 1.3442 | 1.1759 | 1,590,454 |
| Infectious hospital days | 10.45 | 13.36 | 1,590,454 |
| Infectious deaths | 0.0509 | 0.2072 | 1,590,454 |
| | | | |
| Neoplastic hospitalization rate | 6.54 | 5.1076 | 1,590,454 |
| Neoplastic hospital days | 56.92 | 49.24 | 1,590,454 |
| Neoplastic deaths | 0.2812 | 0.5022 | 1,590,454 |
| | | | |
| Metabolic hospitalization rate | 1.6476 | 1.5454 | 1,590,454 |
| Metabolic hospital days | 15.48 | 18.39 | 1,590,454 |
| Metabolic deaths | 0.02534 | 0.1489 | 1,590,454 |

*Source:* German Federal Statistical Office (*Statistische Ämter des Bundes und der Länder*). The *German Hospital Admission Census* includes the county of residence and the day when the patient was hospitalized. The *hospitalization rate* counts the daily incidence of hospitalizations per 100,000 pop. on the county level. *Hospital days* is the sum of all hospital days that were triggered on a given day, i.e., it is the product of the hospitalization rate and the length of stay. *Deaths* counts the number of hospital deaths per 100,000 pop. on the county level. Reference point is always the day when the patient was hospitalized. The patient died sometime after being admitted, but not necessarily on the day of admission. German data protection laws prohibit us from reporting min. and max. values.

## Appendix B: *Mortality Census*

The second register dataset is the *Mortality Census*. It contains the universe of deaths on German territory from 1999 to 2008. This is a restricted access dataset provided by the German Federal Statistical Office (*Statistische Ämter des Bundes und der Länder*). We observe all of the 0.8 million annual deaths. The data contain the following information at the individual admission level:

- age in years
- gender *(binary indicator)*
- county of residence [*between 442 (1999) and 413 (2008) counties*]
- day of death
- primary cause of death *(ICD-10, 3 digit)*

As described in Section 2.5, we normalize, aggregate, and merge this dataset with the other datasets at the day-county level. As such, we obtain the following descriptive statistics.

**Table B1:** Mortality Census: Dependent Variables per 100,000 pop. (Daily County-Level, 1999-2008)

| Variable | Mean | Std. Dev. | N |
|---|---|---|---|
| Mortality rate | 2.9897 | 1.5229 | 1,518,000 |
| Cardiovascular mortality rate | 1.3839 | 1.0788 | 1,518,000 |
| Respiratory mortality rate | 0.1918 | 0.4039 | 1,518,000 |
| Infectious mortality rate | 0.0374 | 0.1749 | 1,518,000 |
| Metabolic mortality rate | 0.0973 | 0.2889 | 1,518,000 |
| Neoplasmic mortality rate | 0.7676 | 0.2889 | 1,518,000 |

*Source:* German Federal Statistical Office (*Statistische Ämter des Bundes und der Länder*). The mortality statistic includes the county of residence and the day of death. The *mortality rate* counts the daily mortality rate per 100,000 pop. at the county level. German data protection laws prohibit us from reporting min. and max. values.

# Appendix C: Official Weather Data

The third register dataset contains daily weather measures from up to 1,044 ambient weather stations. The data are provided by the German Meteorological Service (*Deutscher Wetterdienst (DWD)*). It covers the years from 1999 to 2008. The following weather measures were collected on a daily basis:

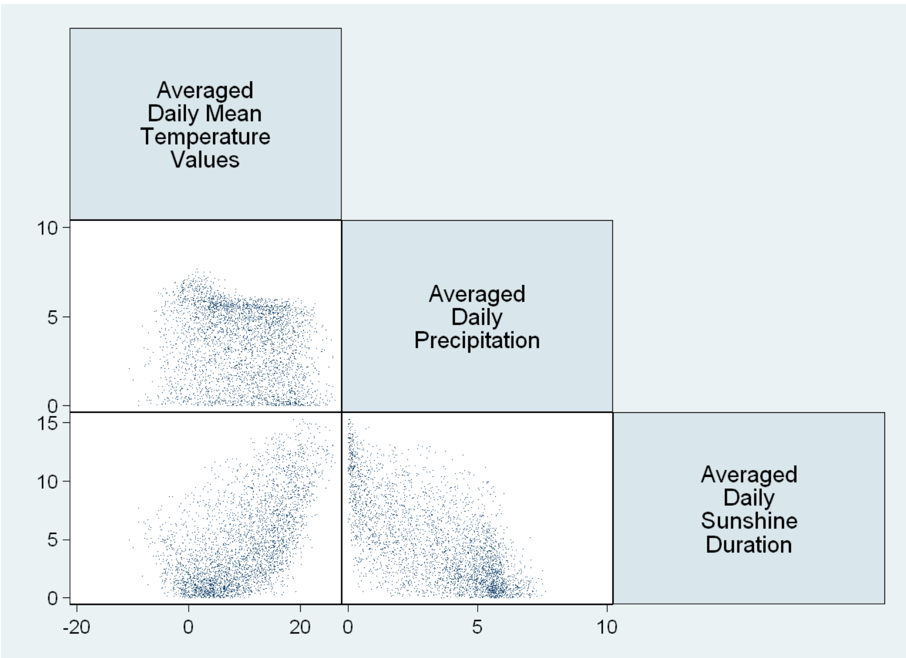
- average temperature in $^\circ$ C [measured *2 m (6'7") above ground*]
- minimum temperature in $^\circ$ C [measured *2 m (6'7") above ground*]
- maximum temperature in $^\circ$ C [measured *2 m (6'7") above ground*]

- total hours of sunshine
- precipitation level in mm per day

- average humidity in percent
- average storm force
- max. wind speed in km per hour (Beauford scale)
- average cloud coverage in percent
- vapor pressure in hectopascal (hPa)
- min. air pressure in hectopascal (hPa) measured [*5 cm (2 inches) above ground*]

As described in Section 2.5, in a first step, we interpolate the point measure into the county space. Then we merge the weather dataset with the other datasets at the day-county level.

**Figure 10:** Scatter Matrix Illustrating Associations Between Temperature, Sunshine, and Precipitation



Panel A of Table C1 shows the descriptive statistics for the raw measures as collected by the DWD. Figure 10 illustrates the associations between the temperature, the hours of sunshine and the precipitation levels. Panel B contains the generated weather condition indicators, i.e., our main variables of interest in the regression models.

**Table C1:** Weather Data (Daily County-Level, 1999-2008)

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| ***A. Raw Measures*** | | | | | |
| Average temperature in $^\circ$C *(2 m (6'7") above ground)* | 9.5573 | 7.3047 | -19 | 30.6 | 1,590,454 |
| Minimum temperature in $^\circ$C *(2 m (6'7") above ground)* | 5.4671 | 6.4965 | -25.01 | 23.8 | 1,590,454 |
| Maximum temperature in $^\circ$C *(2 m (6'7") above ground)* | 13.8912 | 8.5608 | -14.1 | 39.07 | 1,590,454 |
| Total hours of sunshine | 4.6252 | 4.2373 | 0 | 16.7 | 1,590,454 |
| Precipitation level | 2.2246 | 4.2154 | 0 | 144.98 | 1,590,454 |
| Average humidity | 78.3161 | 11.4307 | 10 | 100 | 1,590,454 |
| Average cloud coverage | 5.3128 | 2.1534 | 0 | 8.23 | 1,590,454 |
| Average storm force | 3.6065 | 2.0856 | 0 | 26.3 | 1,590,454 |
| Max. wind speed | 10.4964 | 4.4462 | 0 | 54 | 1,590,454 |
| Vapor pressure | 9.8876 | 3.9981 | 0.5 | 25.9 | 1,590,454 |
| Min. air pressure *(5 cm (2 inches) above ground)* | 3.8456 | 6.5299 | -29.01 | 22 | 1,590,454 |
| ***B. Extreme Temperature Indicators*** | | | | | |
| *Hot Day* (max temp. >30$^\circ$C (86$^\circ$F)) | 0.0197 | 0.1389 | 0 | 1 | 1,590,454 |
| *Heat Wave Day* (4$^{\text{the}}$ consecutive *Hot Day*) | 0.0032 | 0.0568 | 0 | 1 | 1,590,454 |
| *Cold Day* (min temp. <-10$^\circ$C (86$^\circ$F)) | 0.0124 | 0.1106 | 0 | 1 | 1,590,454 |
| *Cold Wave Day* (4$^{\text{the}}$ consecutive *Cold Days*) | 0.0018 | 0.0421 | 0 | 1 | 1,590,454 |

*Source:* German Meteorological Service (*Deutscher Wetterdienst (DWD)*). The information was recorded on a daily basis by up to 1,044 ambient weather monitors that are distributed across the German counties (see Figure 1). The number of weather stations varies from year to year. The weather indicators displayed cover the years 1999 to 2008. As described in Section 2.5, all point measures from the stations are interpolated into the county space by means of deterministic inverse distance weighting (IDW). Level of analysis is the day×county level. Hence, with exactly 400 counties in each year, we would obtain $400 \times 365 \times 10 = 1,460,000$ observations. However, the number of counties varies across years from 442 (1999) to 413 (2008).

## Appendix D: Official Pollution Data

The fourth register dataset contains daily pollution measures from up to 1,314 ambient monitors. The data are provided by the German Federal Environmental Office (*Umweltbundesamt (UBA)*). It covers the years from 1999 to 2008. Measures of the following pollutants have been recorded on a daily basis:

- average concentration of carbon monoxide (CO) in parts per million (ppm)
- minimum concentration of carbon monoxide (CO) in ppm
- maximum concentration of carbon monoxide (CO) in ppm

- average concentration of ozone ($O_3$) in micrograms per cubic meter of air ($\mu g/m^3$)
- minimum concentration of ozone ($O_3$) in $\mu g/m^3$
- maximum concentration of ozone ($O_3$) in $\mu g/m^3$

- average concentration of nitrogen dioxide ($NO_2$) in $\mu g/m^3$
- minimum concentration of nitrogen dioxide ($NO_2$) in $\mu g/m^3$
- maximum concentration of nitrogen dioxide ($NO_2$) in $\mu g/m^3$

- average concentration of sulphur dioxide ($SO_2$) in $\mu g/m^3$
- average concentration of particular matter ($PM_{10}$) in $\mu g/m^3$; since 2000

As described in Section 2.5, in a first step, we interpolate the point measure into the county space via IDW. Then we merge the pollution dataset with the other datasets at the day-county level. Panel A of Table D1 shows the descriptive statistics for the raw measures. The next section describes the chemical composition of the five pollutants, their health hazards, and discusses their tempo-spatial variation. Panel B of Table D1 contains the generated high pollution concentration indicators. The thresholds are modelled after the alert thresholds of the European Union (see Section 2.4 and European Environment Agency (2012)).

**Table D1:** Pollution Data (Daily County-Level, 1999-2008)

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| ***A. Raw Measures*** | | | | | |
| Average CO in ppm | 0.4342 | 0.1794 | 0.0023 | 1.3083 | 1,594,154 |
| Min. CO in ppm | 0.2326 | 0.0911 | 0 | 0.6 | 1,594,154 |
| Max. CO in ppm | 0.8145 | 0.38 | 0.025 | 2.8 | 1,594,154 |
| Average $O_3$ in $\mu$g/$m^3$ | 45.9786 | 22.0423 | 0.8612 | 135.79 | 1,594,154 |
| Min. $O_3$ in $\mu$g/$m^3$ | 17.9888 | 13.8282 | 0 | 79.6 | 1,594,154 |
| Max. $O_3$ in $\mu$g/$m^3$ | 73.7943 | 31.5263 | 1.1673 | 192.15 | 1,594,154 |
| Average $NO_2$ in $\mu$g/$m^3$ | 26.8907 | 10.6284 | 0.0278 | 80.3095 | 1,594,154 |
| Min. $NO_2$ in $\mu$g/$m^3$ | 12.6384 | 5.9959 | 0 | 39.5 | 1,594,154 |
| Max. $NO_2$ in $\mu$g/$m^3$ | 46.4607 | 16.3252 | 0.5 | 132.1 | 1,594,154 |
| Average $SO_2$ in $\mu$g/$m^3$ | 3.7256 | 1.6115 | 0.0654 | 12.5435 | 1,594,154 |
| Average $PM_{10}$ in $\mu$g/$m^3$ | 24.3097 | 11.4625 | 2.0625 | 64.625 | 1,432,822 |
| ***B. Pollution Non-Compliance Indicators*** | | | | | |
| $O_3$ *non-compliance day* (max level >120 $\mu$g/$m^3$) | 0.0929 | 0.2903 | 0 | 1 | 1,594,154 |
| $NO_2$ *non-compliance day* (av. level >40 $\mu$g/$m^3$) | 0.1194 | 0.3243 | 0 | 1 | 1,594,154 |
| $PM_{10}$ *non-compliance day* (av. level >50 $\mu$g/$m^3$) | 0.1278 | 0.3339 | 0 | 1 | 1,594,154 |

*Source:* German Federal Environmental Office (*Umweltbundesamt (UBA)*). The information was recorded on a daily basis by up to 1,317 ambient pollution monitors that are distributed across the German counties (see Figure 1). The number of counties and weather stations vary from year to year. The pollution measures displayed cover the years 1999 to 2008. As described in Section 2.5, all point measures from the stations are interpolated into the county space by means of deterministic inverse distance weighting (IDW). Level of analysis is the day×county level. Hence, with exactly 400 counties in each year, we would obtain $400 \times 365 \times 10 = 1,460,000$ observations. However, as explained in Section 2.5, the number of counties varies across years from 442 (1999) to 413 (2008). CO stands for "carbon monoxide" and ppm for "parts per million." $NO_2$ stands for "nitrogen dioxide," $O_3$ stands for "ozone," $SO_2$ stands for "sulphur dioxide," and $PM_{10}$ stands for "particular matter." $\mu g/m^3$ stands for micrograms per cubic meter of air. The high pollution concentration "non-compliance" days are modelled after the alert thresholds of the European Union (European Environment Agency, 2012) and Section 2.4.

# NO$_2$, O$_3$, CO, SO$_2$, PM$_{10}$: Occurence, Health Hazards, and Variation across Space and Time

## D1.2 Nitrogen Dioxide (NO$_2$)

Nitrogen dioxide is a red-brown toxic gas that is formed by oxidation of nitrogen monoxide (NO). $NO_x$—describing the sum of NO and NO$_2$—is a product of combustion processes under high temperature that happen in automobile engines or fossil fuel power plants; it is also an important intermediate in the chemical industry.

Since $NO_x$ is one main ingredient in the formation of $O_3$ (see below) and highly correlated with the other pollutants, isolating its single impact on human health is challenging. One purpose of this study is to disentangle the health effects of the single pollutants from one another and the weather conditions. Experts by the WHO and the EU warn that "epidemiological studies of $NO_2$ exposure from outdoor air are limited in being able to separate these effects" (World Health

Organization (2003), p. 46; European Environment Agency (2012), p. 39). Evidence for negative health effects mainly comes from indoor toxicological studies showing that $NO_x$ has a negative effect on respiratory functions (cf. Ehrlich et al., 1977; Kerr et al., 1979; Sandstrom et al., 1991; Blomberg et al., 1999; Barck et al., 2002).

**Figure 11:** Nitrogen Dioxide ($NO_2$) Variation Across Counties and Over Time

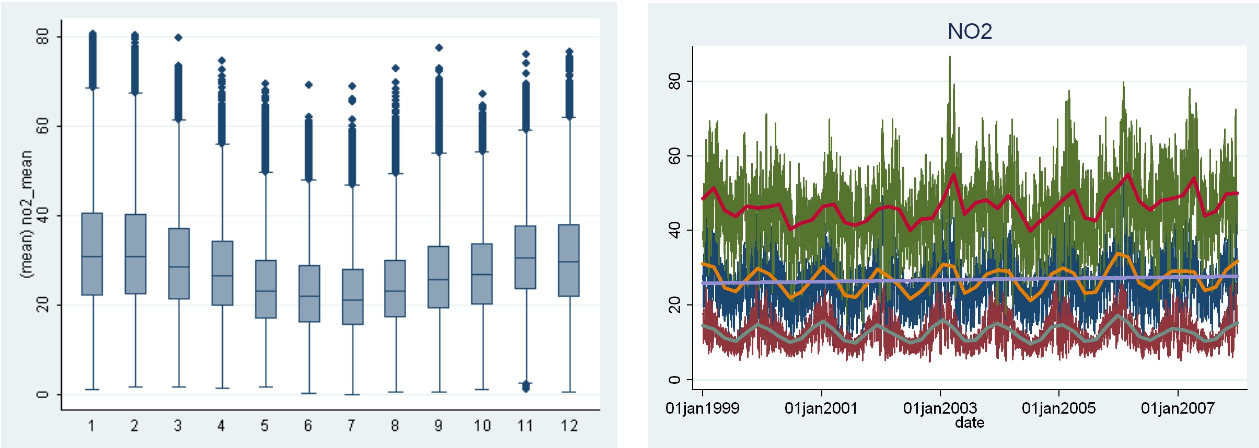
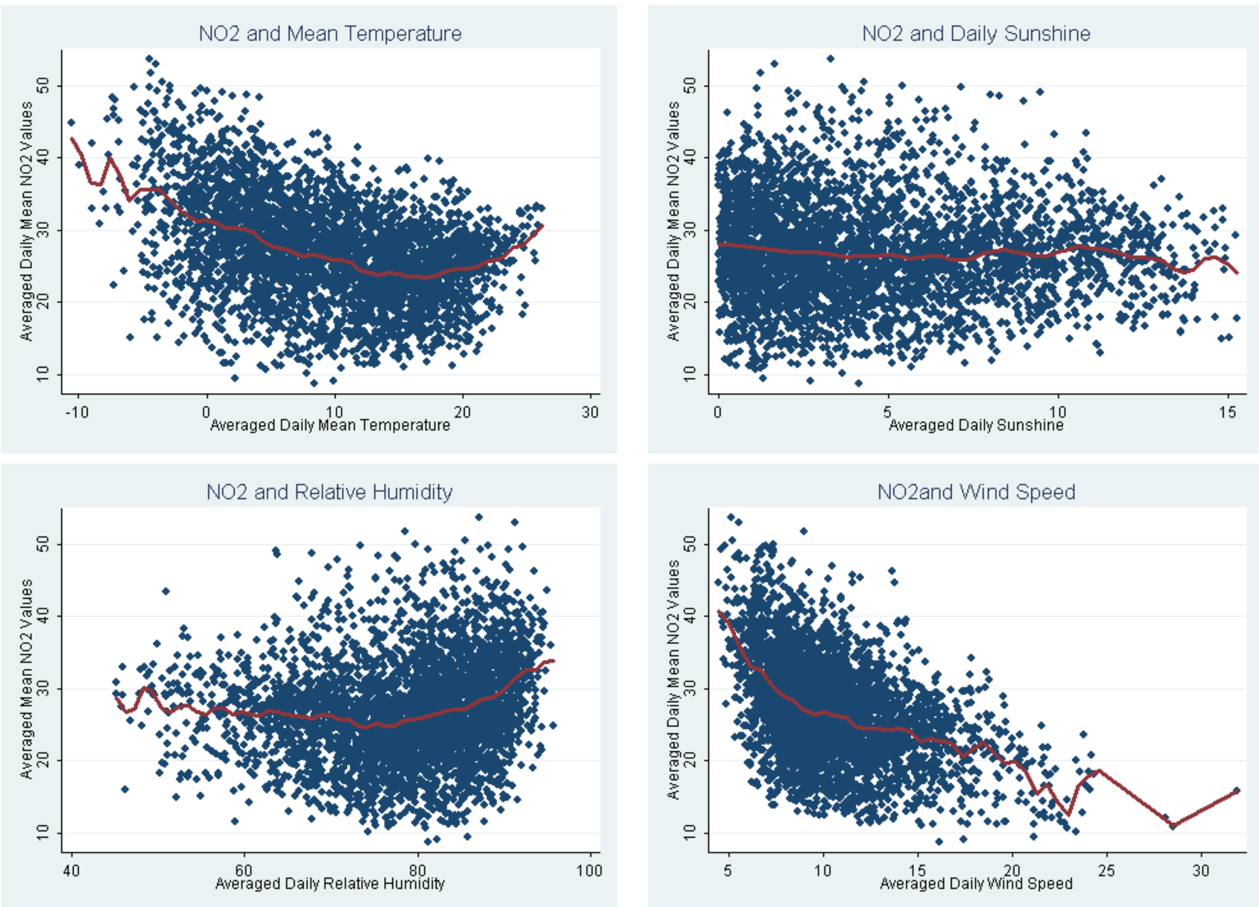



**Figure 12:** Association Between Nitrogen Dioxide ($NO_2$) and Weather




The $NO_2$ concentration is measured in $\mu g/m^3$. The European Union (EU) applies a long-term threshold of 40 $\mu g/m^3$ and an hourly alert threshold of 400 $\mu g/m^3$. If exceeded for more than three hours, authorities are required to implement short-term action plans (European Environment Agency, 2012). The thresholds in the US are much larger—an annual average $NO_2$ concentration of 107 $\mu g/m^3$ or a maximum daily hourly concentration of 203 $\mu g/m^3$ (Environmental Protection Agency (EPA), 2013).

Figure 11a shows a boxplot of the mean daily $NO_2$ levels across German counties and over the twelve months of a year (averaged over 10 years). There is some seasonal variation with lower $NO_2$ levels during the summer month, but most striking is the huge variation within months across counties. The average value over all years and counties is $27\mu g/m^3$ and very similar to the actually measured values in the US, despite the more generous regulatory thresholds (Environmental Protection Agency (EPA), 2011).

Figure 11b shows the mean, minimum, and maximum daily $NO_2$ levels over the time period from 1999 to 2008. First, we observe a significant difference between minimum and maximum daily values throughout the years. Second, there seems to exist a slightly increasing trend in $NO_2$ levels over the 10-year period.

**Figure 13:** Distribution of Nitrogen Dioxide ($NO_2$) Concentration and Non-Compliance Days: Identifying Variation

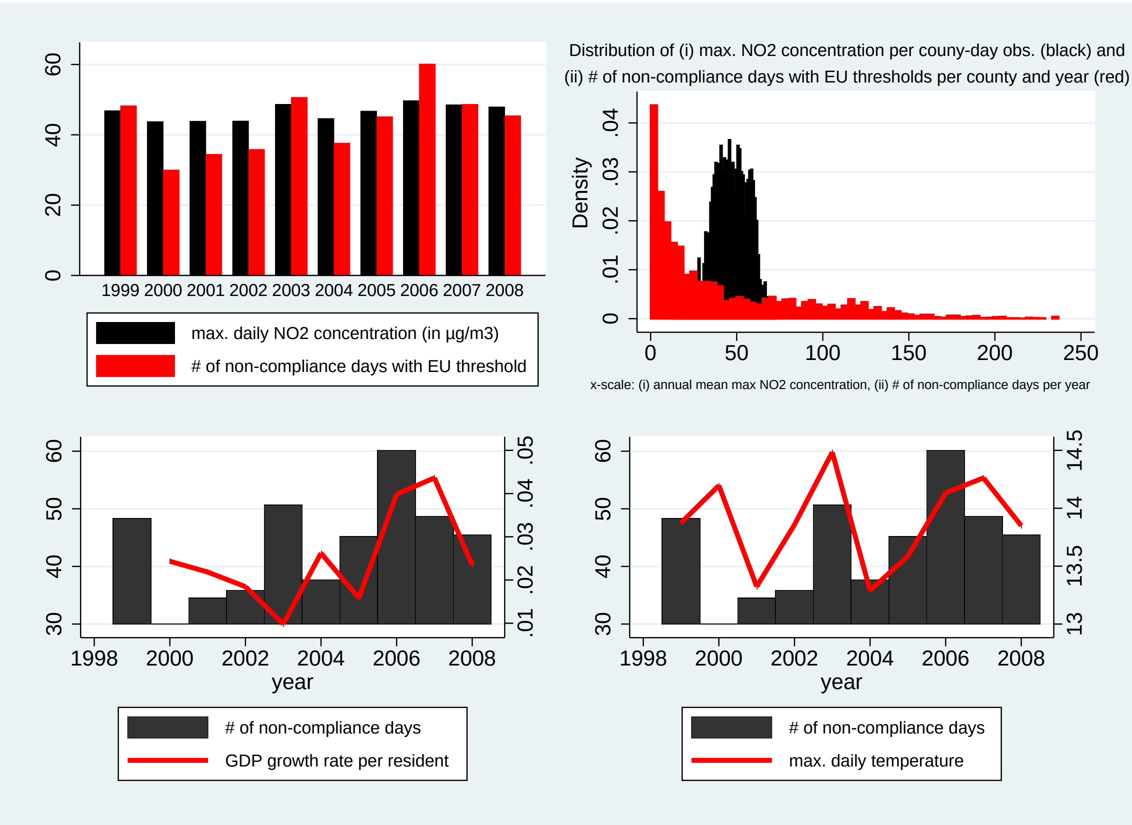


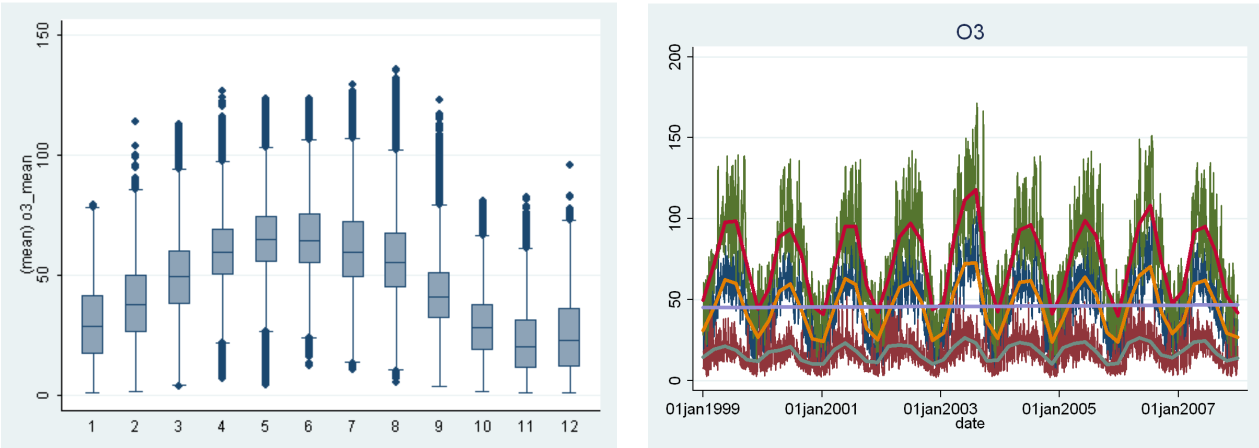

Figure 12 reveals the relationship between $NO_2$ and some of the weather indicators in Table D1. One observes a slightly negative correlation between $NO_2$, the mean temperature and the wind speed. On the other hand, humidity levels of more than 80% seem to be positively correlated with $NO_2$. The is no correlation with hours of sunshine.

Figure 13 is the equivalent to Figure 4 for $O_3$, which was discussed in Secion 2.4. Figure 13 shows that the variation in high concentrations of $NO_2$ has wide support across the German counties—every single county exceeded the EU thresholds several times during the ten years under consideration. The right upper corner of Figure 13 (13b) shows the distributions of both the continuous $NO_2$ measure as well as the binary non-compliance indicator. The left lower corner (Figure 13c) shows that high $NO_2$ concentrations do not seem to be correlated with economic activity at the annual county level, but rather with the maximum temperature (Figure 13d).

### D1.3 Ground-Level Ozone ($O_3$)

Ozone is an oxidant and may lead to respiratory hazards. It is called a "secondary pollutant" since it is formed by various photochemical reactions between carbon monoxide (CO), nitrogen oxides ($NO_x$) and free oxygen molecules (O) (European Environment Agency, 2013). The ground-level ozone concentration is measured in $\mu g/m^3$. According to the European Union (EU), values below 100 $\mu g/m^3$ do not pose a threat to human health. Very high ozone concentrations of more than 240 $\mu g/m^3$ may lead to asthma, bronchitis, chest pain, coughing, throat irritation, or congestion, but also to more severe conditions such as heart attacks or other cardiopulmonary problems (cf. Hackney et al., 1975; Lippmann, 1989; Wright et al., 1990; Devlin et al., 1997; Broeckaert et al., 2000).

**Figure 14:** Ozone ($O_3$) Variation Across Counties and Over Time




In the EU, an hourly concentration of more than 180 $\mu g/m^3$ requires that the population is officially informed by the national authorities. The health alert threshold requires the hourly concentration to not exceed 240 $\mu g/m^3$. The EU Air Quality Directive specifies that a daily maximum 8-hour average of 120 $\mu g/m^3$ should not be exceeded by the member states to avoid health hazards (European Environment Agency, 2012). In the US, the according threshold is an 8 hour average concentration of 160 $\mu g/m^3$ (Environmental Protection Agency (EPA), 2013).

As shown in Table D1 above, in Germany, the average ozone level is 45.98, but average daily values vary from 0.86 to 135.79. Minimum daily values vary from 0 to 79.6, whereas maximum daily county averages range between 1.17 and 192.15 $\mu g/m^3$. In comparison, in the US in 2010, the average ozone concentration was about 150 $\mu g/m^3$ and thus only slightly below the regulatory threshold. A quarter of all sites measured above-threshold concentrations on at least four days of the year (Environmental Protection Agency (EPA), 2011).

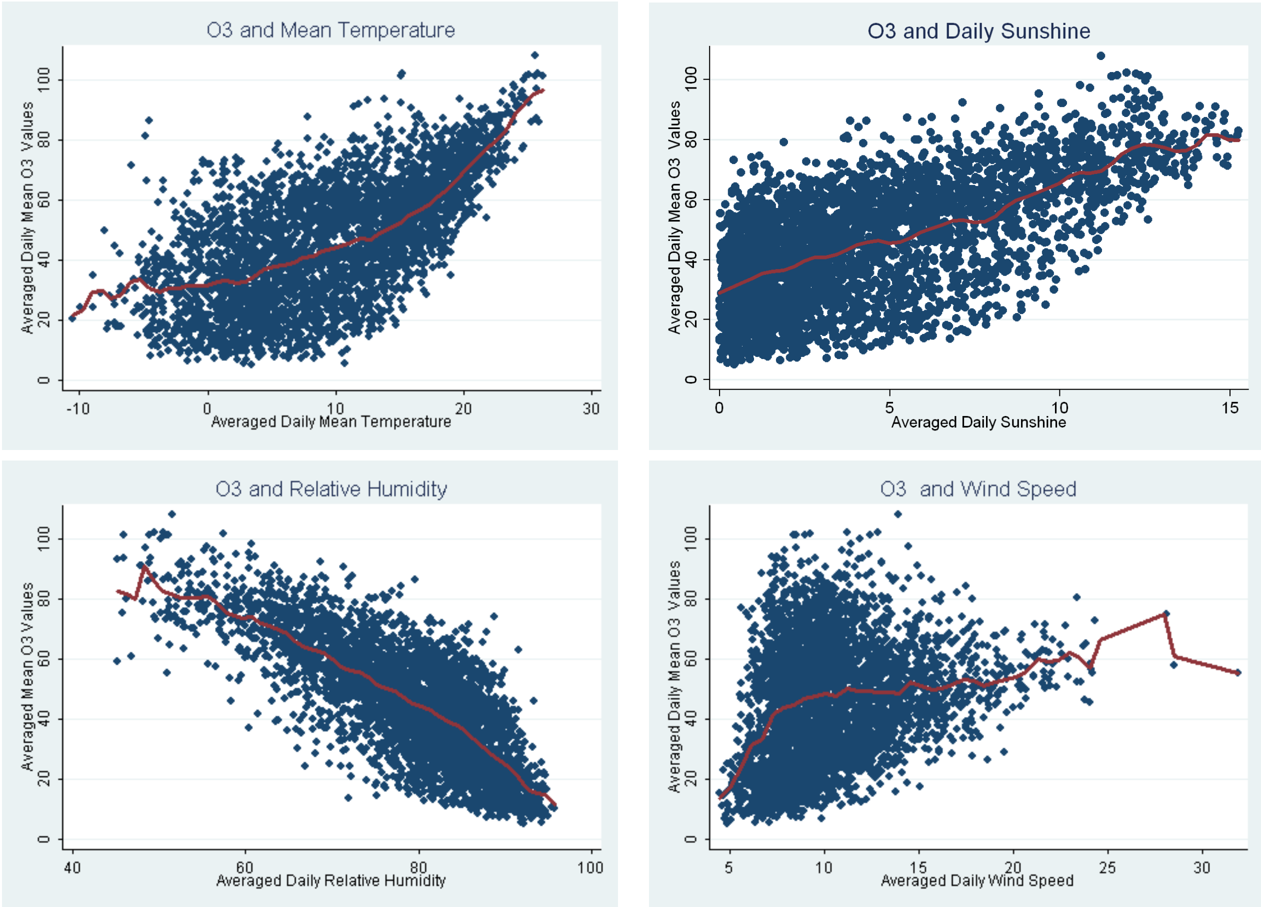
Figure 14a shows the $O_3$ variation across counties and over calendar months. First, there is enormous variation in ozone levels across counties within months. Second, ozone levels increase significantly over the summer months. This is due to the fact that ground-level ozone is highly and positively correlated with both the temperature and the hours of sunshine and thus negatively correlated with humidity (Figure 15). Over the time period from 1999 to 2008, both the variation and the levels of ozone seem to have been stable (Figure 14b).

The equivalent to Figure 13 for ozone is Figure 4 which has been discussed in the main text in Section 2.4.

### D1.4 Particular Matter ($PM_{10}$)

Particular matter (PM) is a generic term and describes aerosol particles—or athmospheric aerosol—which can be of different size and chemical composition. $PM_{10}$ refers to particles with a diameter of at most 10 micrometres. PM may have a "natural" origin and stem from sea salt,

**Figure 15:** Association Between Ozone ($O_3$) and Weather



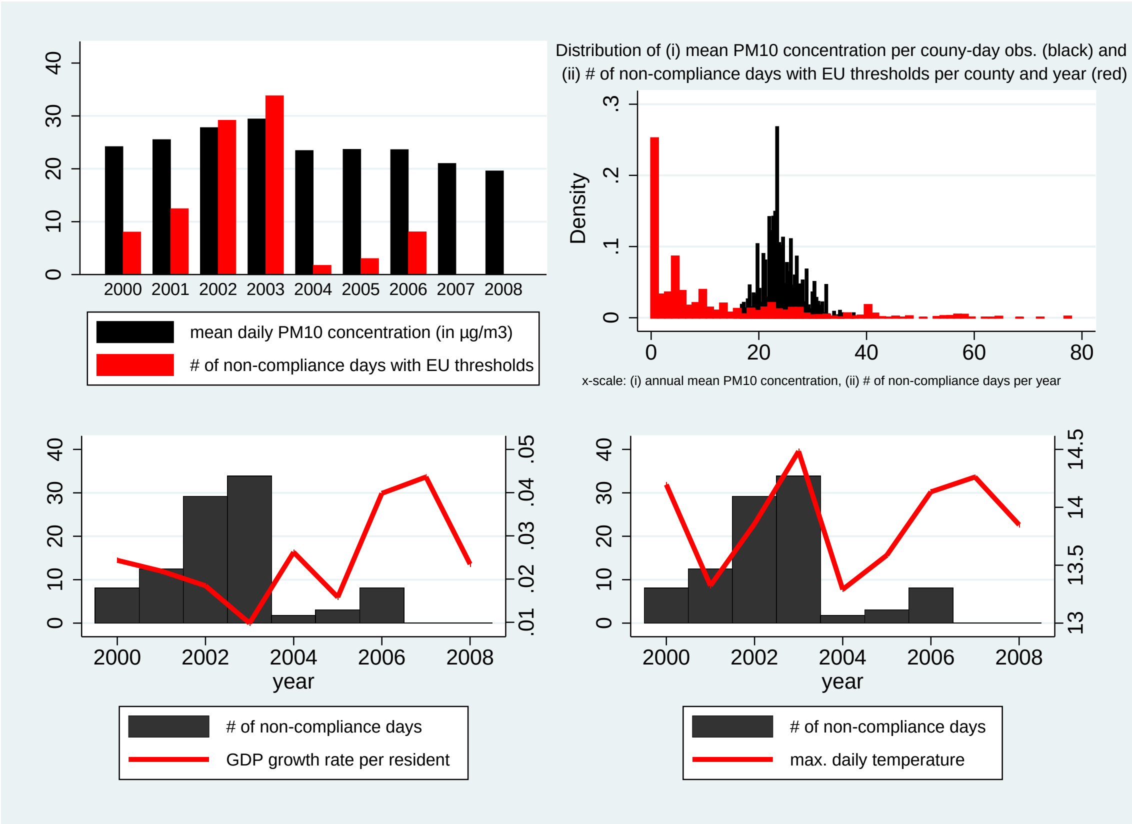

dust, pollen or ash from volcanos. However, PM may also result from fuel combustion, e.g., burning of wood, domestic heating, road dust due to traffic, or power generation. Then it is typically formed from oxidation and transformation of "primary" pollutants such as $SO_2$ or $NO_2$ (European Environment Agency, 2012).

Health effects of PM are caused through lung inhalation, and physical as well as chemical reactions with lung cells. A plenitude of epidemiological studies demonstrate a strong link between PM exposure and cardiovascular mortality in particular (cf. Pope III et al., 2002; Li et al., 2011). For example, Abbey et al. (1999) find a signficant impact of $PM_{10}$ on respiratory deaths as well as lung cancer. However, studies that intend to measure the effects of long-term exposure to PM suffer from various methodological challenges, such as selection into regions and a high permanent correlation with other pollutants.

The EU short-term limit value is a 24 hour concentration of 50 $\mu g/m^3$. Effective January 2005, this concentration ought not to be exceeded on more than 35 days per year. However, various European cities regulary exceed that threshold (European Environment Agency, 2012). The WHO sets the same daily air quality guideline value in addition to a maximum annual mean value of 20 $\mu g/m^3$ and states: "The aim is to achieve the lowest concentration possible. As no threshold for PM has been identified below which no damage to health is observed [...]" (World Health Organization, 2011). The Environmental Protection Agency (EPA) (2013) defines the $PM_{10}$ threshold as an 24 hour average concentration of 150 $\mu g/m^3$, i.e., three times larger than in Europe.

Table D1 shows that the average daily $PM_{10}$ concentration is indeed relatively high in Germany, namely 24.3 $\mu g/m^3$ and thus lies above the WHO annual guideline value. However, it is twice as low as in the US. The maximum daily mean is 64.6 $\mu g/m^3$. Nevertheless, plotting the daily $PM_{10}$ concentrations over a decade, it becomes clear that concentrations decreased between

**Figure 16:** Distribution of Particular Matter ($PM_{10}$) Concentration and Non-Compliance Days: Identifying Variation




1999 and 2008 (graph not shown). Interestingly, there are only very weak seasonal $PM_{10}$ trends (not shown). In the US, in 2010, the average measured $PM_{10}$ concentration is twice as high as in Germany and about 60 $\mu g/m^3$. Ten percent of all sites measure *average* concentrations of more than 90 $\mu g/m^3$, despite a 30% decrease in average national concentrations since 2001 (Environmental Protection Agency (EPA), 2011).

Figure 16 is the equivalent to Figures 4 (for $O_3$) and 13 (for $NO_2$) for $PM_{10}$. As for $O_3$ and $NO_2$, one sees that a large set of German counties contributes to the identifying variation in the $PM_{10}$ concentration. Between 1999 and 2008, all German counties exceeded the thresholds on between 8 and 558 days per county, i.e., even the least $PM_{10}$ polluted county did not comply to regulatory thresholds on 8 days within a decade. The bottom graphs of Figure 16 also show that high temperatures rather than GDP growth are correlated with high levels of particular matter. As for high ozone levels, heat is an input factor for the formation process—through oxidation—of this secondary pollutant (Arya, 1998; World Health Organization, 2003; European Environment Agency, 2012). The relationship between daily $PM_{10}$ levels and the daily mean temperature is U-shaped with $PM_{10}$ levels increasing strongly when temperatures exceed 20$^\circ$C (68$^\circ$F). For maximum daily temperatures above 20$^\circ$C (68$^\circ$F), the correlation between the maximum daily temperature and the maximum daily $O_3$ concentration is 0.7, for the maximum daily $NO_2$ concentration it is 0.2, and for mean daily $PM_{10}$ concentration it is 0.3. Hence, it is reasonable to think of exogenous heat shocks triggering high pollution levels.

### D1.5 Carbon Monoxide (CO)

Carbon monoxide is a colorless odorless gas that is toxic to humans in higher concentrations. The typical concentration in the atmosphere is about 0.1 parts per million (ppm). Incomplete burning of carbon-containing materials, such as smoke from fire, is one main source of high CO concentrations. However, in industrialized countries, automobile fuel combustion is responsible for a large fraction of CO concentration in the air. CO concentrations of more than 100 ppm are considered health damaging, although individual tolerance levels vary significantly (Stewart et al., 1970; Anderson et al., 1973; Penney, 2000; Omaye, 2002; Mayr et al., 2005).

CO decreases the blood oxygen transmission. According to the Centers for Disease Prevention and Control (CDC), in the US, about 450 people die every year from "accidental, non-fire related exposure to this toxic gas." CO poisoning would require medical care for thousands more (Centers for Disease Control and Prevention, 2012). Omaye (2002) notes that CO poising may be the main cause of more than 50% of all fatal poisonings in industrialized countries and that many situations would remain un- or misreported. The EU and WHO 8-hour threshold values are 10 $\mu g/m^3$ (or 8.7 parts per million (ppm)) (European Environment Agency, 2012). The US threshold is very similar and an 8-hour average concentration of 9 ppm (Environmental Protection Agency (EPA), 2013). In 2010, the average actual concentration in the US was 2 ppm.

Table D1 above shows that the daily mean ambient carbon monoxide (CO) concentration in parts per million (ppm) is 0.43, ranging from 0.002 to 1.31. The daily mean minimum concentration is 0.23 and the maximum concentration is 0.81. The latter varies between 0.03 and 2.8. A boxplot of daily CO levels shows the typical seasonal variation with lower CO levels during the summer month. Over the last decade, average CO concentrations have slightly decreased, but the standard deviation remains high.

Note that we do *not* generate binary "non-compliance" indicators for carbon monoxide, simply because the EU alert threshold was never exceeded during the period of observation in Germany.

### D1.6 Sulphur Dioxide ($SO_2$)

Sulphur dioxide is a colorless toxic gas emitted by sulphur containing fuels when burned. Industrial processes lead to $SO_2$ emissions as do domestic heating and transportation. For example, coal contains sulphur and thus coal combustion releases $SO_2$ unless the sulphur components are removed before the burning process. Oxidation of $SO_2$ may lead to $H_2SO_4$ and acid rain. $SO_2$ is also a precursor for particular matter. While $SO_2$ is still one of the main air pollutants in developing countries, due to environmental regulation, $SO_2$ emissions decreased significantly over the last decades in industrialized countries (World Health Organization, 2000; European Environment Agency, 2013).

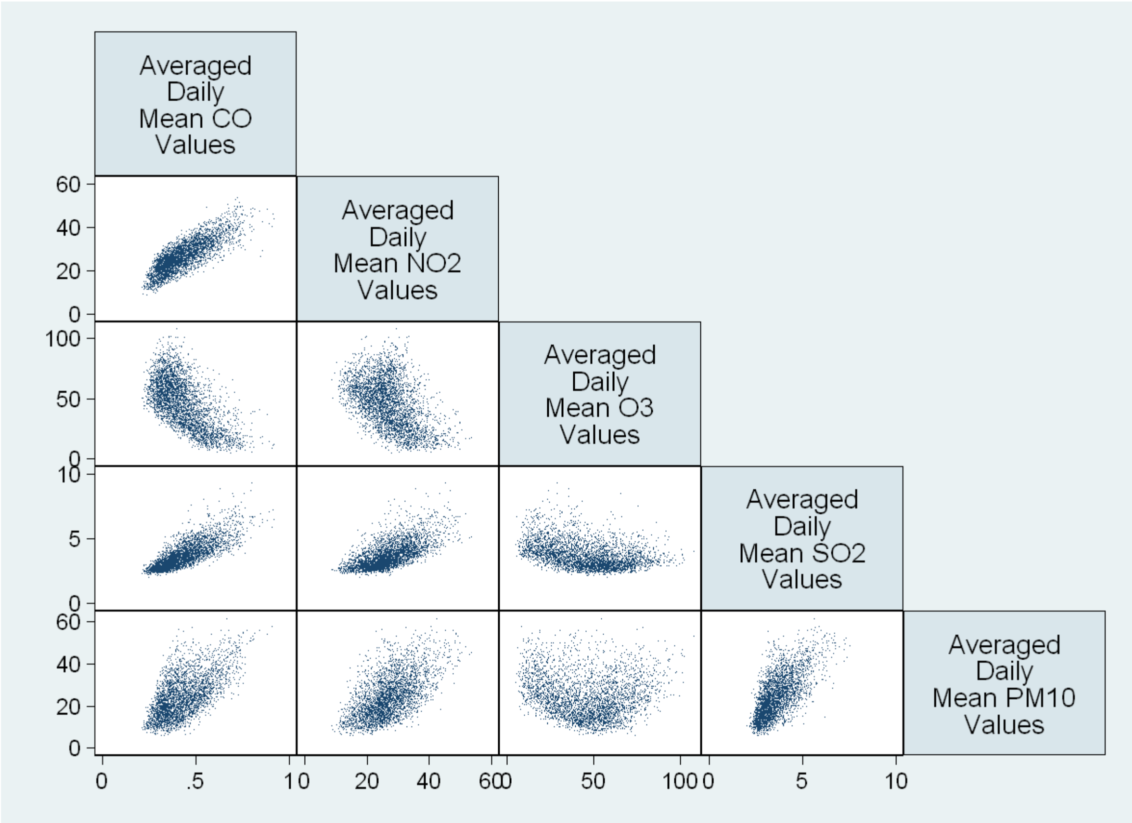
Epidemiological and experimental studies with small numbers of volunteers show that $SO_2$ concentrations may primarily result in adverse respiratory health effects. It disrupts the ciliary function, slows the ciliary transport of mucus and may lead to coughing, asthma and chronic bronchitis. Moreover, for people with heart diseases and among vulnerable populations, $SO_2$ shocks my lead to hospitalizations, premature birth, and deaths (Lawther et al., 1975; Horstman et al., 1988; Shah and Balkhair, 2011).

Natural $SO_2$ concentrations in rural areas are around 5 $\mu g/m^3$. The EU threshold for daily $SO_2$ concentrations is 125 $\mu g/m^3$. The hourly alert threshold is 500 $\mu g/m^3$ and action plans have to be implemented when this threshold is exceeded in three consecutive hours. The US thresholds are significantly larger. The "primary" threshold is a one hour concentration of not more than 75 ppb (=2,120 $\mu g/m^3$) and the "secondary" threshold a three hour concentration of not more than 0.5 ppm (= 14 $\mu g/m^3$) (Environmental Protection Agency (EPA), 2013). The average concentration measured across the US was about 2.5 ppb (= 71 $\mu g/m^3$) (Environmental Protection Agency (EPA), 2011).

As Panel A of Table D1 illustrates, all $SO_2$ concentration values measured in all German counties from 1999 to 2008 are significantly below these thresholds. The average concentration is 3.7 $\mu g/m^3$ and its maximum 12.5 $\mu g/m^3$. Thus, as in case of CO, we do not generate binary non-compliance indicators for $SO_2$. Boxplot graphs (not displayed) show significant variation across counties with average values slightly lower in the summer months. Plotting values over time illustrates a significant decline in $SO_2$ concentrations from 1999 to 2008.

In principle, pollution regulation in the US is similar to in the EU: the US Environmental

**Figure 17:** Scatter Matrix Illustrating Associations Between Pollutants




Protection Agency (EPA) implements pollution concentration thresholds and requires the US states to comply. However, the EPA thresholds are significantly less strict: The $PM_{10}$ threshold is an 24 hour average concentration of 150 $\mu g/m^3$. The $O_3$ threshold is an 8 hour average concentration of 160 $\mu g/m^3$. And the $NO_2$ threshold is an annual average concentration of 107 $\mu g/m^3$ or a maximum daily hourly concentration of 203 $\mu g/m^3$ (Environmental Protection Agency (EPA), 2013).[26] Thus, the threshold levels for $NO_2$ and $PM_{10}$ are 2 to 3 times larger in the US, which should be kept in mind when comparing the results of this study to related US studies. In Germany, from 1999 to 2008, the US regulatory thresholds for $PM_{10}$, $O_3$ and $NO_2$ were *never* exceeded (see Table D1). The actually measured average concentrations for $O_3$ and $PM_{10}$ are three and two times larger in the US than in Germany, respectively, while average $NO_2$ concentrations are—despite larger regulatory thresholds—very similar.

### D1.7 Associations Between All 5 Pollutants

Lastly, Figure 17 shows the associations between all five air pollutants discussed above. $NO_2$ is positively correlated with $SO_2$ and $PM_{10}$, but negatively correlated with $O_3$. The same is true for CO. $O_3$ exhibits only very noisy and weak associations with $SO_2$ and $PM_{10}$. However, $SO_2$ and $PM_{10}$ show a strong and positive association.

[26] The original scales for $NO_2$ and $O_3$ are expressed in "parts per million (ppm)" and have to be converted to "micrograms per cubic meter of air $\mu g/m^3$". The annual threshold for $NO_2$ is 0.053ppm and the hourly maximum 100ppm. For $O_3$, the "annual fourth-highest daily maximum 8 hours concentration, averaged over 3 years," must not exceed 0.075ppm.

## Appendix F: Annual Socio-Economic County-Level Data

Finally, this paper makes use of yearly county-level data provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (2012) (*Bundesinstitut für Bau-, Stadt- und Raumforschung*) in their INKAR (*Indicators and Maps on Spatial Development*) database. The data vary by year.[27] To normalize the hospitalization and death rate dependent variables, we use county-level total population counts. In addition, we use information on the *unemployment rate* and *GDP per capita*. Supply-side constraints are captured by the *# hospitals per county*, *hospital beds per 10,000 pop.* and *physicians per 10,000 pop.*

On average, about 190,000 residents live in each German county. The average per capita income is €25,000 p.a.[28], but varies between €11,282 and €86,728 across counties and over years. A similarly strong variation is observed for the county unemployment rate which varies between 1.6 and 29.3% with an average of 10.5%.

An average county has 5 hospitals. However, in some counties there exist no hospital and one county counts a staggering 76 hospitals. Consequently, the number of hospital beds per 10,000 residents and county varies between 0 and 24,170. The outpatient physician density varies between 69 and 394 doctors per 10,000 residents of a county.

**Table F1:** Descriptive Statistics Other (County-Level, 1999-2008, Annual)

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Unemployment rate | 10.47 | 5.28 | 1.6 | 29.3 | 4,356 |
| GDP per capita | 24971 | 10146 | 11,282 | 86,728 | 4,354 |
| # hospitals per county | 4.84 | 5.49 | 0 | 76 | 4,354 |
| Hospital beds per 10,000 pop. | 1211.19 | 1593.88 | 0 | 24,170 | 4,354 |
| Physicians per 10,000 pop. | 152.72 | 52.59 | 69 | 394 | 4,358 |
| Total population | 189,450 | 219,753 | 34,525 | 3,431,675 | 4,361 |
| Male 0 to 2 years | 2,575 | 3,034 | 331 | 47,489 | 4,361 |
| Male 3 to 5 years | 2,697 | 2,968 | 328 | 42,964 | 4,361 |
| Male 6 to 9 years | 3,776 | 3,972 | 409 | 60,320 | 4,361 |
| Male 10 to 14 years | 5,151 | 5,277 | 525 | 92,611 | 4,361 |
| Male 15 to 17 years | 3,280 | 3,323 | 366 | 55,698 | 4,361 |
| Male 18 to 19 years | 2,241 | 2,323 | 383 | 38,669 | 4,361 |
| Male 20 to 24 years | 5,613 | 6,704 | 987 | 111,475 | 4,361 |
| Male 25 to 29 years | 5,708 | 7,926 | 1,007 | 134,581 | 4,361 |
| Male 30 to 34 years | 6,628 | 9,117 | 881 | 164,445 | 4,361 |
| Male 35 to 39 years | 7,991 | 10,168 | 1,056 | 172,517 | 4,361 |
| Male 40 to 44 years | 8,089 | 9,634 | 1,347 | 164,928 | 4,361 |
| Male 45 to 49 years | 7,195 | 8,082 | 1,157 | 149,742 | 4,361 |
| Male 50 to 54 years | 6,274 | 7,021 | 926 | 116,102 | 4,361 |
| Male 55 to 59 years | 5,589 | 6,749 | 845 | 129,022 | 4,361 |
| Male 60 to 64 years | 5,745 | 6,929 | 817 | 119,554 | 4,361 |
| Male 65 to 74 years | 9,210 | 10,096 | 1,108 | 187,669 | 4,361 |

*Continued on next page...*

[27] The hospitalization and mortality data contain the county of residence according to the county codes and boundaries of the specific year. In contrast, the INKAR database contains all information according to the county codes and boundaries as of January 1, 2012. From 1999 to 2008, various county reforms, mostly mergers between two counties, led to changes in the county codes and boundaries. Consequently, the number of counties varies across years from 442 (1999) to 413 (2008). For counties with county reforms, we imputed pre-reform values using the post-reform boundary data as of January 1, 2012. In addition to reforms, not all information listed above have been collected in every single calendar year. We imputed missing values for these cases. See notes to Table F1 for more details.

[28] In 2012 values.

*... Table F1 continued*

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Male > 75 years | 4,882 | 5,087 | 658 | 81,884 | 4,361 |
| | | | | | |
| Female 0 to 2 years | 2,442 | 2,882 | 295 | 44,660 | 4,361 |
| Female 3 to 5 years | 2,561 | 2,824 | 313 | 41,049 | 4,361 |
| Female 6 to 9 years | 3,584 | 3,770 | 406 | 57,060 | 4,361 |
| Female 10 to 14 years | 4,887 | 4,997 | 492 | 88,234 | 4,361 |
| Female 15 to 17 years | 3,109 | 3,147 | 358 | 52,753 | 4,361 |
| Female 18 to 19 years | 2,135 | 2,275 | 377 | 37,463 | 4,361 |
| Female 20 to 24 years | 5,431 | 7,071 | 939 | 117,108 | 4,361 |
| Female 25 to 29 years | 5,516 | 8,044 | 828 | 137,220 | 4,361 |
| Female 30 to 34 years | 6,331 | 8,559 | 699 | 152,632 | 4,361 |
| Female 35 to 39 years | 7,578 | 9,364 | 1,046 | 158,939 | 4,361 |
| Female 40 to 44 years | 7,714 | 9,012 | 1,204 | 153,034 | 4,361 |
| Female 45 to 49 years | 6,998 | 7,868 | 1,270 | 140,548 | 4,361 |
| Female 50 to 54 years | 6,232 | 7,188 | 906 | 117,351 | 4,361 |
| Female 55 to 59 years | 5,634 | 6,939 | 855 | 127,897 | 4,361 |
| Female 60 to 64 years | 5,959 | 7,239 | 838 | 123,874 | 4,361 |
| Female 65 to 74 years | 10,689 | 11,874 | 1,952 | 214,713 | 4,361 |
| Female > 75 years | 10,006 | 11,110 | 1,964 | 164,217 | 4,361 |

*Source:* Federal Institute for Research on Building, Urban Affairs and Spatial Development (2012). The information varies across counties and over years on an annual basis. Some information has not been surveyed in every calendar year. In addition, in contrast to the register databases in Appendices A and B, the INKAR data refers to the county codes and boundaries as of January 1, 2012. Since various county reforms were implemented between 1999 and 2008, we had to impute information for pre-reform counties with post-reform data (if possible). For example, if counties A and B simply merged to county C and we only had the GDP per capita for county C, we would impute the GDP per capita values for A and B using the population information on A and B which is available for all years and counties. If, as another example, data was surveyed in every other year, we took the mean value of $t_0$ and $t_2$ to impute information for $t_1$. However, we were unable to impute values for all measures and all counties in every year according to the boundaries of that specific year, which is why the number of observations slightly varies between the measures.

# Appendix G: Cross-Validation of Weather and Pollution Interpolation

**Table G1:** Cross-Validation of IDW Interpolation

| Variable | Raw Correlation | | Time and Season-Adjusted Correlation | |
|---|---|---|---|---|
| | IDW Method | NN Method | IDW Method | NN Method |
| CO Mean | 0.477 | 0.363 | 0.149 | 0.082 |
| CO Max | 0.413 | 0.301 | 0.131 | 0.069 |
| CO Min | 0.607 | 0.522 | 0.227 | 0.182 |
| NO2 Mean | 0.562 | 0.450 | 0.407 | 0.321 |
| NO2 Max | 0.531 | 0.423 | 0.400 | 0.313 |
| NO2 Min | 0.606 | 0.497 | 0.434 | 0.349 |
| O3 Mean | 0.862 | 0.797 | 0.435 | 0.362 |
| O3 Max | 0.929 | 0.896 | 0.373 | 0.328 |
| O3 Min | 0.671 | 0.555 | 0.473 | 0.371 |
| SO2 Mean | 0.616 | 0.532 | 0.306 | 0.265 |
| PM10 Mean | 0.837 | 0.814 | 0.239 | 0.212 |
| Cloud | 0.874 | 0.821 | 0.585 | 0.508 |
| Humidity | 0.876 | 0.826 | 0.643 | 0.566 |
| Vapor Pressure | 0.979 | 0.970 | 0.735 | 0.678 |
| Temperature | 0.981 | 0.972 | 0.733 | 0.661 |
| Air Pressure | 0.549 | 0.579 | 0.239 | 0.257 |
| Wind Speed | 0.497 | 0.478 | 0.219 | 0.156 |
| Min Temperature | 0.968 | 0.953 | 0.713 | 0.637 |
| Max Temperature | 0.977 | 0.966 | 0.659 | 0.587 |
| Precipitation | 0.788 | 0.740 | 0.688 | 0.634 |
| Sunshine | 0.934 | 0.922 | 0.556 | 0.535 |

*Source:* German Meteorological Service (*Deutscher Wetterdienst (DWD)*) and German Federal Environmental Office (*Umweltbundesamt (UBA)*). The table represents the cross-validation of the weather and pollution interpolation as described and discussed in Section 2.5. The underlying data stems from up to 1,044 ambient weather monitors and up to 1,317 ambient pollution monitors between 1999 and 2008. Columns (1) and (3) display the Pearson's Correlation Coefficient between the orginal values of monitior X and its predicted values solely using all surrounding monitors and Inverse Distance Weighting (IDW). Columns (2) and (4), in contrast, simply use the Nearest Neighbor (NN) method and thus predict values of monitor X with the measurement of its nearest neighbor monitor. Columns (3) and (4) are based on values that have been non-parametrically adjusted for all 3,650 day effects, i.e., the nationwide daily mean of a specific measure was first removed from all monitor measurements. This exercise removes time trends, but likewise the "true" correlation in measurements between monitors and has to be regarded as a very conservative test. More details are in Section 3.3.

**Table G2:** Share of Correctly Predicted Extreme Weather Indicators

| **Panel A: IDW** | Overall Correct Predicted | Positives Correct Predicted | Zeros Correct Predicted | Reliability Ratio |
|---|---|---|---|---|
| *Hot Day* | 0.9904 | 0.8133 | 0.9939 | 0.8071 |
| *Heat Wave Day* | 0.9983 | 0.8003 | 0.9989 | 0.7993 |
| *Cold Day* | 0.9927 | 0.7680 | 0.9954 | 0.7634 |
| *Cold Wave Day* | 0.9982 | 0.5812 | 0.9989 | 0.5801 |

| **Panel B: NN** | Overall Correct Predicted | Positives Correct Predicted | Zeros Correct Predicted | Reliability Ratio |
|---|---|---|---|---|
| *Hot Day* | 0.9881 | 0.7286 | 0.9937 | 0.72233 |
| *Heat Wave Day* | 0.9978 | 0.7089 | 0.9989 | 0.7079 |
| *Cold Day* | 0.9908 | 0.6699 | 0.9951 | 0.6651 |
| *Cold Wave Day* | 0.9965 | 0.3063 | 0.9991 | 0.3054 |

*Source:* German Meteorological Service (*Deutscher Wetterdienst (DWD)*). The underlying data stems from up to 1,044 ambient weather monitors between 1999 and 2008. Panel A tests the predictive quality of the Inverse Distance Weighting (IDW) interpolation method into the county space and Panel B the Nearest Neighbor (NN) method. All numbers are shares of predicted relative to actual values. The predicted value for monitor X are calculated using solely all surrounding monitors and assuming that monitor X is non-existent. Column (1) reports the overall share of correctly predicted positive or negative extreme weather indicator values. Column (2) reports the share $\chi$ of correctly predicted positives and column (3) the share $\delta$ of correctly predicted zero values. Consequently, 1-$\chi$ represent false positives and 1-$\delta$ false negatives. Column (4) shows the Reliability Ratio (RR) $\alpha$ which indicates the ratio between OLS and IV estimates and thus assesses the size of the potential attenuation bias (Hyslop and Imbens, 2001). More details are in Section 3.3.