

IZA DP No. 7929

## Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes

Jenny Jans  
Per Johansson  
J. Peter Nilsson

January 2014

# **Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes**

**Jenny Jans**  
*Uppsala University*

**Per Johansson**  
*IFAU, Uppsala University and IZA*

**J. Peter Nilsson**  
*IIES, UCLS and CHW*

**Discussion Paper No. 7929**  
**January 2014**

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes<sup>\*</sup>

On normal days, the temperature decreases with altitude, allowing air pollutants to rise and disperse. During inversion episodes, a warmer air layer at higher altitude traps pollutants close to the ground. We show how readily available NASA satellite data on vertical temperature profiles can be used to measure inversion episodes on a global scale with high spatial and temporal resolution. Then, we link inversion episode data to ground level pollution monitors and to daily in- and outpatient records for the universe of children in Sweden during a six-year period to provide instrumental variable estimates of the effects of air quality on children's health. The IV estimates show that the respiratory illness health care visit rate increases by 8 percent for each  $10 \mu\text{m}/\text{m}^3$  increase in PM10; an estimate four times higher than conventional estimates. Importantly, by linking the health care data to detailed records of parental background characteristics, we show that children from low-income households suffer significantly more from air pollution than children from high income households. Finally, we provide evidence on the importance of several mechanisms that could contribute to the difference in the impact of air pollution across children in rich and poor households.

JEL Classification: Q53, I1, I3, J24

Keywords: air pollution, health, inversions, environmental policy, instrumental variable, nonparametric regression, socio-economic gradient in health

Corresponding author:

J. Peter Nilsson  
Institute for International Economic Studies  
Stockholm University  
10691 Stockholm  
Sweden  
E-mail: [peter.nilsson@iies.su.se](mailto:peter.nilsson@iies.su.se)

---

\* Earlier versions of this paper have been presented at the ASSA Meetings in Chicago 2010, at Uppsala University (Spring 2010), and the IFAU Workshop in Labor and Public Economics in Öregrund (June 2010), SIEPR Stanford University (Fall 2010), and Statistics Norway (October 2013), Columbia University (November 2013), SOFI (December 2013). Special thanks to Douglas Almond, Ken Chay, and Timo Boppart for helpful suggestions and comments.

# 1 Introduction

Children in low- and high-income households differ in their health and wellbeing. The socio-economic health gap is present at birth and increases over the child's life cycle (Case, Lubotsky and Paxson, 2001). The capacity formation framework of Cunha and Heckman (2007) illustrates how childhood health could generate significant effects on subsequent health, educational attainment and labor market outcomes through dynamic complementarities and cross productivity with the development of cognitive and non-cognitive skills. An increasing number of design based studies have also found that early life health does not only influence adult health, but also educational attainments (Almond, Edlund and Palme, 2009) and labor market outcomes (Almond, 2006; Nilsson, 2008). By now, there is strong evidence on the link between parents' socio economic status and child health (c.f. Currie, 2008 and Almond and Currie, 2011); a link that suggests that parts of the intergenerational persistence in inequality are due to differences in childhood health conditions. Yet, why children from poorer socio-economic backgrounds experience more frequent and more severe health problems is far from clear.

Differences in environmental amenities have been suggested as one culprit believed to contribute to the socio-economic gap in child health. In this paper, we focus on identifying how one aspect of the environment, ambient air pollution, influences children's respiratory health and to what extent and why poor air quality affects children from different socio economic backgrounds differentially. Respiratory illnesses among children account for a large proportion of the health care costs and productivity loss due to parental work absence related to child health, and asthma is one of the most common chronic conditions among children, affecting around 10-15% of the children in developed countries. Air pollution is believed to be an important factor causing and aggravating respiratory illnesses.

The identification of the causal relationship between air quality and child health is, however, hampered by a number of factors. First, high-income parents tend to reside in larger metropolitan areas where pollution levels are, on average, higher and, at the same time, high income parents may have more resources available to mitigate the effects of poor air quality, which potentially leads to the health costs on the general population being understated. Such between city residential sorting may induce substantially biased estimates of the effects of poor air quality. Second, within city sorting may also lead to that families with low-income more often may live close to major roads than high income families, and may therefore be exposed to worse air quality. At the same time, houses in neighborhoods close to the major roads may have a worse standard in general. Thus, there is a clear risk using cross sectional data that the influence of ambient air pollution is overstated if, for example, we are not able to take housing standard, or other aspects of the home environment such as parental smoking, into account. Alternatively, families with children who suffer from e.g. asthma may choose to locate in areas with better air quality, which if not properly taken into account would bias the effect of poor air quality on respiratory illness downwards. Third, since individual monitoring of air pollution exposure is expensive, air quality measures typically stem from a few ambient air quality monitors located in the most polluted areas. Hence, measurement error in pollution exposure is likely to be a serious concern.

To address these issues, this paper uses NASA satellite data on vertical temperature profiles to provide causal estimates of the short term effects of poor air quality on children's health using inversion episodes as an instrumental variable. On normal days, the temperature decreases monotonically with altitude and the constant flow of air between warm and cool areas helps clear pollutants from the ground level air layer. During inversion episodes, the temperature follows a non monotonic pattern in altitude. The temperature first increases with altitude up to the inversion layer, and then

decreases with altitude. This leads to a sharp deterioration of air quality in the ground level air layer since pollutants are trapped under the inversion layer.

Using the vertical temperature profile data, which is readily available with high spatial resolution on a global scale, we identify daily temperature inversion episodes for six years. We merge the inversion episode data with ground level air quality monitor data, and provide nonparametric estimates showing that, conditional on ground level weather conditions, there is no apparent relationship between PM10 levels and inversion strength – the temperature difference between the ground level air layer and the air layer just above it – following normal nights. However, following inversion nights, there is a strong positive relationship between inversion strength and PM10 levels.

When combining these data sources with administrative in and outpatient health care records, we find a corresponding pattern between inversion strength and children's respiratory illnesses. On average, following inversion nights, the 24h PM10 levels are 30 percent higher than after normal nights and the respiratory illness rate is five percent higher. At the relatively low pollution levels considered, these estimates suggest that an increase of 10  $\mu\text{m}/\text{m}^3$  PM10 is associated with an increase of the respiratory health care visit rate by eight percent. This instrumental variable estimate is more than four times larger than the corresponding ordinary least squares estimate from the same data. The relationship between the OLS and the IV is almost identical to estimates in previous studies that also address measurement error, avoidance behavior, and other forms of endogeneity problems but use more context specific instrumental variables (see e.g. Moretti and Neidell, 2011).

The validity of the instrumental variable approach developed here hinges on the assumption that inversion episodes only influence respiratory health by their impact on air quality. This identifying assumption could, for example, be violated if inversion episodes also change e.g. children's outdoor activities. For example, if inversion episodes

are not only associated with pollution levels but also colder outdoor temperatures, changes in cloud coverage, or other weather conditions that reduce outdoor activities, our instrumental variable estimator would produce biased estimates of the effects of air pollution on health outcomes.

We address this important concern in two ways. First, all our baseline estimates control flexibly for ground level temperatures, cloud coverage, humidity and precipitation, wind speed and season of the year. Second, as a falsification exercise we make use of data on information on external causes for health care visits. Since accidents are related to children's activity patterns but arguably not directly to pollution levels, we check whether the changes in air pollution from inversion episodes are related to the risk of having to visit the health care providers due to an injury with external causes. We find no evidence of a correlation between inversion episodes and injuries due to external causes.

Our study contributes to the literature on air pollution and health in several ways. First, the instrumental variable approach we develop can be used with high resolution in both space and time on a global scale. Previous design-based studies have used instrumental variables that provide a valid inference in their particular contexts, typically in areas with relatively high pollution levels. The method can easily be extended to other countries, to other outcomes, and to other populations. Our approach opens up the possibility for comparative studies of the effects of air pollutants in various contexts using the same instrument. This could mitigate the concern that differing estimates of effects of air pollution are due to differences in estimation methods, rather than different effects on populations in e.g. developing vs. developed countries. The inversion data can potentially also be used in combination with more location specific instruments to improve estimation precision.

Second, as compared to other countries, Sweden has relatively low air pollution

levels, and previous studies have only shed limited light on health effects in low pollution areas. For example, the 24-hour mean PM10 level in our sample is  $20\mu g/m^3$ ,  $35\mu g/m^3$  in the United States, and  $67\mu g/m^3$  in Mexico City. Yet, we still find substantial effects from modest increases in PM10 levels. Our non parametric estimates indicate a linear relationship at least up to  $50\mu g/m^3$  (the current WHO 24 hour air quality guideline), and hence, our results suggest that increasing the Swedish 24-h average PM10 level to the United States level would increase children's respiratory health care visit rate by, on average, 12 percent. Clearly, even at levels well below current regulatory standards, significant benefits on child health can be expected from improvements in air quality. These benefits should be taken into account when considering the net benefits of stricter air quality standards.

Third, the richness of the data allows us to examine the effects of air pollution by background characteristics in a more detailed way than most previous studies. For example, while we find that parents' educational attainments seem to play a limited role, household income seems to matter much more. The estimated effects are about twice as large on children in households with an income below the median as compared to those above. We also assess the relevance of some of the potentially important mechanisms behind the SES-gap in the effects of air pollution. In particular, we find that the gap between children in rich and poor households decreases substantially if comparing children with relatively poor baseline health. We find no indication that the SES-gap is explained by strong differences in avoidance behavior across rich and poor households.

Combined with the absence of differential effects across households with differing educational attainment, and no indication of strong nonlinearities in the effects of PM10, these results suggest that economic resources via differences in children's baseline health play a bigger role than information differences in generating the SES-gap in the health

effects of air pollution in our low pollution setting. Since pollution exposure early in life has also been found to influence economic outcomes later in life (Nilsson, 2009; Sanders, 2012; Isen, Rossin-Slater, Walker, 2013), it seems that environmental policies could also play an important role in reducing inequality in economics outcomes.

Finally, the AIRS data can be incorporated in models predicting pollution levels at many more locations than what is currently possible using ground based vertical temperature sounders (e.g. available in 4 locations in Sweden, and 90 locations in the US). By developing methods to produce more precise location-specific pollution forecasts and effective means of disseminating such forecasts, the health care costs associated with respiratory symptoms could potentially be reduced by allowing sensitive populations to more effectively engage in defensive investments.

The rest of the paper is structured as follows: we start with a simple conceptual framework to guide our empirical exercise in section 2. In section 3, a background on air pollution and health is presented, and the data used is described in section 4. Section 5 discusses the econometric methodology, and the results can be found in section 6. The final section summarizes and concludes the paper.

## 2 Conceptual Framework

To fix ideas, and motivate our empirical exercise, suppose that respiratory illnesses induced by changes in air pollution are captured by the three key factors in equation (1)

$$Resp = f(P, A, H) \tag{1}$$

where respiratory illnesses ( $Resp$ ) are a function of ambient air pollution, ( $P$ ), parental awareness/avoidance behavior ( $A$ ), and baseline child health ( $H$ ).  $P$ ,  $A$ , and  $H$  can be viewed as functions of parents' income and/or education. In this paper, our

three primary objectives are (i) first to provide causal estimates of the direct biological effect of air pollution on respiratory health,  $\partial \text{Resp} / \partial P$ . For this purpose, our empirical strategy is designed to isolate the influences of pollution while holding avoidance behavior and baseline health fixed. (ii) Our second objective is to document to what extent the effects of pollution on child health may differ between children in different socio economic groups. Recent studies (Nilsson, 2009; Currie, 2008) find suggestive evidence that the reduced form effects of air pollution on children's health tend to be larger on children in low socio economic status (SES) households. i.e.:

$$\left| \frac{d\text{Resp}}{dP} \right|_{\text{LowSES}} > \left| \frac{d\text{Resp}}{dP} \right|_{\text{HighSES}}$$

However, conclusive evidence on the mechanisms behind the SES-gap in the effects is still missing. So our third objective is to (iii) provide insights on the key underlying mechanisms. To emphasize the different channels highlighted in previous work, assume that Equation (1) can be represented by the linear approximation that allows for interaction effects between  $P$  and  $A$ , and  $P$  and  $H$ :

$$\text{Resp} = \alpha_1 + \alpha_2 P + \alpha_3 A + \alpha_4 H + \alpha_5 P^2 + \alpha_6 A \times P + \alpha_7 H \times P \quad (2)$$

In Equation (2), respiratory illnesses are portrayed as a function of the three key factors displayed in equation (1), and capture three mechanisms that could contribute to differences in the marginal effects of air pollution on children across socio-economic groups. First, ambient air pollution affects child respiratory illnesses negatively through an increase in  $P$  via  $\alpha_2 + \alpha_5$ , where  $\alpha_5$  captures potentially nonlinear effects of ambient air pollution levels on child health. Second, the effects of an increase in  $P$  can be mitigated by parental avoidance behavior ( $A$ ) through the negative  $\alpha_6$ . Finally, the influences of marginal changes in pollution can also be affected by the child's health

stock. Children with a higher level of  $H$  are assumed to be more resilient to effects of changes in  $P$  and hence,  $\alpha_7 < 0$ .

This stylized framework suggests that children from poorer households can be more affected by changes in ambient air pollution than children from richer households for three reasons. First, as noted in the introduction, children from poorer households generally have poorer health than children in rich households. Second, parents in richer households generally have higher educational attainments and hence, may be more aware of effects of air quality on child health or, alternatively, parents in high income households may be more willing to engage in avoidance behavior to reduce the risk of children's respiratory illnesses since the parental costs of child illness could be higher in terms of lost parental labor earnings. Third, children in poorer households may be observed to be more influenced by average pollution levels within a municipality since the pollution levels vary even within small neighborhoods. If noise from traffic or ambient air pollution levels are reflected in housing prices, children from poorer households may more often tend to reside closer to pollution sources and hence, be exposed to higher levels of pollution. In other words, if children to poor and rich are equally affected but we are not able to observe individual levels of exposure, only e.g. municipality means, it may occur that children in poorer households are more affected given the same level of average pollution levels within a municipality.

Equation (2) shed light on the importance of the three mechanisms, however it can, of course, be extended to be made more complex and more realistic. For example, it is possible that the extent of parental avoidance behavior depends on the level of  $P$ , i.e. that parents in high pollution areas (such as Los Angeles) are more likely to be willing to engage in (potentially costly) avoidance behavior than parents in low pollution areas (such as in our setting). Similarly, parents of children with a lower health stock may also be more willing to engage in avoidance behavior if their child is more likely to be

affected by changes in pollution levels.

Below we provide evidence on the general effect, effects on children from differing socio economic backgrounds and also try to shed some light on the importance of the three mechanisms highlighted in equation (2); (i) non linearities in effects of air pollution, differences in (ii) avoidance behavior and/or (iii) baseline health across children in rich and poor households.

### 3 Background on the relationship between air quality and health

#### 3.1 Particulates, Health, and Air Quality Policies in the European Union

Epidemiological studies have attributed negative health effects from exposure to particulate matter (PM). PM is a general term used for particles where the major components are sulfate, nitrates, ammonia, sodium chloride, carbon, mineral dust and water. The particles are identified according to their aerodynamic diameter, as either PM10 (with a diameter of 10 micrometers or less<sup>1</sup>) or PM2.5 (with a diameter smaller than 2.5  $\mu m$ ). When breathed in, the particles are sufficiently small to penetrate to the thoracic region, where the finer fraction has a high probability of deposition in the smaller conducting airways and alveoli. Inhalation of PM has been found to trigger inflammation in the smaller airways, leading to the exacerbation of asthma and chronic bronchitis, airway obstruction and decreased gas exchange(Nel et al., 1998; Ghio and Devlin, 2001). Besides aggravated asthma and increased respiratory symptoms, inhalation of PM has also been associated with heart attacks were evidence suggests that the effects may be expressed through several, probably interrelated, pathways (WHO, 2006).

---

<sup>1</sup>By definition, PM10 thereby includes both ‘coarse particles’ and the finer PM2.5 particles. Due to data availability, we focus on PM10 in the current paper.

To mitigate the effect on public health and environmental protection, Sweden follows the air quality standards set by the EU-directive 2008/50/EG. For PM10, the limit values for short-term (24 hours) are  $50 \mu\text{g}/\text{m}^3$  (not to be exceeded more than 35 times per calendar year) and  $40 \mu\text{g}/\text{m}^3$  long term (annual) exposure.<sup>2</sup> However, the inability to identify a threshold below which adverse health effects are not observed implies that any limit value may leave some residual risk when exposed to PM. This has led the World Health Organization (WHO) to recommend more stringent air quality guidelines (WHO, 2006), with a 24-hour mean of  $50 \mu\text{g}/\text{m}^3$  and an annual mean of  $20 \mu\text{g}/\text{m}^3$ . The US EPA 24-hour PM10 standard is  $150 \mu\text{g}/\text{m}^3$  (not to be exceed more than 1 time per year).

In Europe, it is the short-term limit of PM10 that is most often exceeded in cities and urban areas. Countries such as Poland, Italy, Turkey, Latvia, Lithuania, Sweden, the United Kingdom and the Balkan region all exceeded the daily limit value in 2010. Sites that exceeded the 24 hour limit value were traffic sites (33 percent), urban background sites (29 percent), other, mostly industrial sites (17 percent) and rural sites (14 percent) within the EU (EEA, 2012). Figure A1 shows the attainment situation for PM10 in 2010 in the EU-27 countries.

In Sweden, the total emission of particles decreased by around 23 percent between 1990 and 2000, but has since remained constant (see Figure A2). In Sweden, the energy and transport sector contribute to more than 75 percent of total particle emissions (SEPA, 2012). The dominant factors contributing to the emissions are electricity and heat production, and in urban areas emissions from road vehicles including road wear.

### **3.2 Previous studies on Pollution and Respiratory Health**

#### *Epidemiology*

---

<sup>2</sup>Limits also exist for PM2.5, but so far only at the long-term annual level.

Epidemiological studies on the health effects of air pollution have to a large extent relied on cross sectional data and compared the prevalence of hospitalizations due to respiratory illness in cities with differing pollution levels at a single point in time, or used time-series data for e.g. PM10 levels and asthma for a particular city or region. Using these approaches, it is difficult to draw causal conclusions about the magnitude of the effects on health.

Cross-sectional studies are likely to confound effects of pollution with effects of unobserved factors that are correlated with pollution levels and respiratory illnesses in the cross section. It is not even clear in which direction omitted variables will tend to bias such estimates; a higher pollution level could signal better employment opportunities and income which, in turn, may mitigate the risk of experiencing respiratory health problems. Alternatively, higher air pollution levels could simply capture effects of unobserved factors such as a generally worse environment, housing standard, parental or child smoking patterns, etc. which, in turn, may result in overstated estimates of the effects of air pollution. Pure time series studies may, on the other hand, not only capture the effect of variations in pollution but also other unobserved factors that co-varies with pollution patterns, for example, weather conditions, seasonal variations in activity patterns. As discussed further below, temporal fluctuations in air quality are also, to the extent that they are captured in pollution forecasts, potentially related to defensive investments or avoidance behavior among sensitive populations such as asthmatics. Such behavioral changes are likely to lead to understated effects of poor air quality.

There are a few exceptions in the epidemiological literature on the effects of air pollution on health that employ a design based approach for inference. Pope, Schwartz and Ransom (1992) examine the effects of variations in PM levels following a temporary shutdown of production in a steel mill. By comparing respiratory related emergency

room visits in the valley where the mill was located to the neighboring valley, they find that when PM levels dropped following the plant shut down so did respiratory illnesses.

### *Economics*

Economists have contributed to the literature on effects of poor air quality on health in several ways during the last decade, primarily by highlighting potential identification problems and by using increasingly sophisticated empirical strategies designed to address these endogeneity problems.

First, air pollution is not randomly assigned across locations. Chay and Greenstone (2003b) note that air quality is capitalized in house prices where individuals with a higher income (which can be seen as a function of productivity and health) and/or individuals with preferences for clean air may sort into better air quality areas. Thus, exposure to pollution levels is typically endogenous. Failing to account for this kind of residential sorting, unobserved determinants of health may bias the estimation of the effect of pollution on health. In the absence of a randomized experiment, this has led to a rise in estimation techniques to isolate exogenous changes in pollution. For example, Chay and Greenstone (2003b,a) use the implementation of the Clean Air Act of 1970 and the recession of the early 1980s to exploit the induced temporal and spatial variation in TSP levels in the United States. Lleras-Muney (2010) use seemingly random allocations of military families across military bases in the US to estimate the effects of air pollution on children's hospitalizations.

Some studies use seasonal variations in pollution levels within residential areas to address endogenous sorting (e.g. Currie and Neidell (2005); Currie, Hanushek, Kahn, Neidell and Rivkin, 2009) . One potential problem with using seasonal variation is the risk of confounding by weather conditions, since weather directly affects health (Deschenes and Moretti, 2009) and pollution levels. Accounting for all possible weather

factors influencing *both* pollution and health is a challenging task. Knittel, Miller and Sanders (2011) show that including higher order terms for temperature and precipitation as well as second-order polynomials for some weather conditions, such as wind speed, humidity, and cloud cover, have a substantial impact on estimates of pollution on infant mortality.

A third complication arises from measurement error both in monitored pollution levels but even more importantly when assigning pollution exposure to the individual. In particular when using fixed effect models. Since individual exposure indicators are typically only available for small samples, the predicted pollution level is at best a noisy measure for true exposure. Studies typically assign data from ambient air pollution monitors to the residential location of the individuals, an approach that is likely to generate substantial measurement errors due to the spatial variation in pollution. If the measurement error in e.g.  $PM_{10}$  levels can be written as  $PM_{10} = u + PM_{10}^*$ , where  $PM_{10}$  is the observed exposure,  $PM_{10}^*$  is the true level of exposure and  $u$  is the measurement error, assumed to be independent from the true exposure level  $cov(u, PM_{10}^*) = 0$ , then the OLS estimates of the effects of air pollution will be biased towards zero. There are good reasons to suspect that measurement error is an important problem in pollution exposure studies and that accounting for measurement error is important.

A final problem is that the effect of pollution on health might be highly dependent on behavioral responses. For example, individuals might undertake defensive investments by purchasing preventive pharmaceuticals (Deschenes et al., 2012) or, perhaps particularly relevant in high pollution settings, engage in avoidance behavior and reduce their time spent outdoors (Neidell, 2009). Ignoring behavioral responses could generate downward biased estimates. To account for avoidance behavior, Moretti and Neidell (2011) estimate the health effects of ozone by employing data on daily shipping

traffic in the port of Los Angeles as an instrumental variable for ozone levels. Their findings are striking. The OLS estimates are significant but small; exposure to ozone causes \$11.1 million per year in annual hospital costs in Los Angeles. IV estimates, accounting for behavioral responses, measurement errors and potential confounders are considerably higher; indicating an annual cost of \$44 million from respiratory related hospitalizations. Their results underscore the importance of accounting for unobserved factors in order to understand the full welfare effects caused by air pollution.

Schlenker and Walker (2011) instrument air pollution using air traffic congestion in remote major airports to estimate the health impact of air pollution on populations living in the vicinity of 12 local major airports in California. They find that carbon monoxide (CO) leads to significant increases in hospitalization rates for asthma, respiratory, and heart related emergency room admissions that are an order of magnitude larger than conventional estimates. They assess the differential impact across age groups, but they do not examine whether the effects differ across socio economic groups.

### **3.3 Previous Studies on Temperature Inversions, Pollution, and Health**

As we will see in our analysis, the majority of inversion episodes are far from catastrophic; however, many of the worst pollution episodes on record have been found to coincide with inversion episodes.<sup>3</sup> Most recently, in January 2013, severe smog marked three straight days of dangerous air pollution in Beijing, China. The hazardous air-borne smoke, with alarming levels of PM matter exceeding  $900 \mu\text{g}/\text{m}^3$  in some districts, led the authorities to implement an emergency response plan for the first time in the Chinese history.

---

<sup>3</sup>E.g. Donora, Pennsylvania (1948), The London Fog (1952), and the Union Carbide plant disaster in Bhopal (1984).

Many studies have previously related inversion episodes to poor air quality. For example, a study by Kukkonen et al. (2005) finds that inversion periods in European cities coincide with levels of PM far above average. Likewise, in January 2004, Utah's Cache Valley (US) experienced an event of stagnant weather conditions that drove particulate concentrations to new levels, two times the 24-hour standard used at the time by the US EPA (Malek, Davis, Martin and Silva, 2006).

A few studies have also examined the effects of inversion episodes on public health. Abdul-Wahab, Bakheit, and Siddiqui (2005) suggest an association between the monthly number of inversion days and the monthly mean of daily visits to the emergency department in Oman. Beard et al. (2012) used weather balloon data in Salt Lake County and related inversion episodes to emergency room visits and found a positive correlation. Wallace, Nair and Kanaroglou (2010) first used the AIRS data to look at public health using cross section data on 674 asthmatics (on average 55 years old) in Hamilton, Ontario, Canada. They found evidence of an association between an indicator for inversion and sputum cell counts (an indicator of airway inflammation).

Methodologically, the most similar previous work is a recent study by Arceo-Gomez, Hanna and Olivia who uses information on inversion episodes measured in one (1) location over Mexico City<sup>4</sup>. Arceo-Gomez et al. exploit the number of thermal inversions over the city per week as an instrumental variable for weekly pollution levels in the municipalities within the city. Their result indicates that a 1 percent increase in PM10 over a year leads to a 0.42 percent increase in infant mortality, while a 1 percent increase in CO results in a 0.23 percent increase in infant mortality. They do not provide any results across socio economic groups.

Arceo-Gomez et al. compare the estimates from their developing country setting to

---

<sup>4</sup>The instrumental variable approach we use was developed independently and without knowledge of their paper. The first-stage results were presented at the ASSA Meetings in Chicago in 2010, Uppsala University (2010), and at SIEPR, Stanford University (2010).

estimates from US studies, but acknowledge that they are not fully comparable since the models differ across studies. By using the approach developed in the current study, comparative studies can be conducted on a global scale; in areas with high (such as Mexico City, with a PM<sub>10</sub> 24-h mean of  $67 \mu\text{g}/\text{m}^3$ ), medium (e.g. the United States, PM<sub>10</sub> 24-h mean of  $35 \mu\text{g}/\text{m}^3$ ), or relatively low levels of pollution (e.g. the Swedish cities in our sample, PM<sub>10</sub> 24-h mean of  $20 \mu\text{g}/\text{m}^3$ ) using the same identification strategy.

## 4 Data

We compiled a dataset on daily health care visits, weather conditions and pollution measures from September 2002 until September 2007. Each data source is described in detail below and summary statistics is provided in Table 1.

### 4.1 Data sources

#### *i) Identifying inversion episodes*

To identify inversion episodes, we exploit vertical temperature profile data available from NASA's Atmospheric Infrared Sounder (AIRS).<sup>5</sup> In 2002, the AIRS instrument was launched onboard the NASA satellite AQUA. AIRS produces a 3-D map of temperature and water vapors in the atmosphere. The primary mission of AIRS is to improve weather predictions, and collect a wide range of data twice a day.

NASA provides the AIRS data in three different forms. Level 1 data provides the highest resolution ( $1.5\text{km} \times 1.5\text{km}$ ) and is not yet available to researchers outside NASA. Level 2 data (L2) has a spatial resolution of approx.  $45\text{km} \times 45\text{km}$ . Level 3 data (L3), which we use, has a spatial resolution of  $1^\circ \times 1^\circ$  which corresponds to approximately

---

<sup>5</sup>As part of the activities of NASA's Science Mission Directorate and it is archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Centre (DISC).

100km×100km at the relevant latitude. The data is collected twice a day; at 02am and 02pm local time. The L3 standard products are the primary public L3 product that only contains well-validated fields and includes temperature and water vapor profiles that are reported for 24 pressure levels. Grid maps coordinates range from  $-180.0^{\circ}$  to  $+180.0^{\circ}$  in longitude and from  $-90.0^{\circ}$  to  $+90.0^{\circ}$  in latitude. In the current study, we use L3 data due to the easy access and its readiness for use by researchers. Downloading the L3 data for a particular region is straightforward and irregularities have been corrected by the NASA. Future studies could exploit additional spatial variation using L2 data.

The L3 temperature profile data contains temperatures for 22 layers, defined by average air pressure in the layer. To identify inversion episodes, we use the temperatures for the two pressure levels closest to the ground (1000 hPa and 925hPa).<sup>6</sup> The 1000hPa layer temperature corresponds to the surface conditions and 925hPa layer measure conditions at approx. 600m above the sea level. In the analysis, we use the temperature differences between these two layers to identify inversion episodes and inversion strength. Under normal conditions, the temperature decreases with altitude and hence, the temperature difference between the 925hPa and 1000hPa air layer is negative. Under inversion episodes, the difference is positive. As further motivated in section 4, we focus on night-time inversion episodes which occurred on around 25 percent of the days in our sample. The inversion strength is defined as the temperature difference between the two layers during inversion, with higher values corresponding to stronger inversions<sup>7</sup>.

We also use information on cloud coverage and humidity from the AIRS data. Cloud coverage is important since the AIRS instruments cannot retrieve temperature profiles if the grid cell is under complete cloud coverage. Missing values from the inability to

---

<sup>6</sup>AIRS Level 3 version 5 with spatial box: 55S, 10W, 70N, 24E.

<sup>7</sup>It is possible to calculate whether inversion occurs at higher altitudes as well. We abstained from doing so since we expected that the strongest effect on pollution would come from inversion episodes close to the ground.

measure temperature profiles is another reason why researchers may prefer to use L3 data over L2 data.<sup>8</sup> Humidity data is also important, since it has been linked to both air pollution level and health.

We also gathered wind and precipitation data measured every third hour at 119 weather stations around Sweden maintained by the Swedish Meteorological and Hydrological Institute (SMHI) as additional controls in our analysis. As a first step, daily means are calculated, after which the distance to the closest temperature grid point is obtained. Then, we assign the mean of the six nearest weather stations to each grid point, weighting by the inverse distance between the station and the grid centroid. Missing values were replaced by the monthly-municipality mean. We also present estimates without these additional weather variables included.

### *ii) Pollution data*

Pollution data for the period studied is obtained from the Swedish Environmental Research Institute, IVL. Pollution monitors collect data on either an hourly or a daily basis. We assign each pollution monitor to the nearest temperature grid by calculating the distance to the grid cell centroid (see Figure 1). Since some of the municipalities have more than one monitor, a daily municipality-pollution mean is calculated. Out of Sweden's 290 municipalities, 90 measured PM10 daily during the period studied. We concentrate on PM10 levels due to data availability; other pollutants are measured with much lower frequency, consistency, and spatial coverage. The PM10 levels are furthermore highly focused on in policy circles due to the health effects associated with PM exposure.<sup>9</sup>

---

<sup>8</sup>In our sample, on average 13.5 % (i.e. around 4 days per month) of the AIRS observations are missing due to full cloud coverage. The share of missing temperature profile days per month: Jan (.196) Feb (.160) March (.111) April (.093) May (.095) June (.113) July (.132) Aug (.107) Sept (.119) Oct (.149) Nov (.171) Dec (.178).

<sup>9</sup>Sweden follows the air quality standards set by the EU-directive 2008/50/EG. For PM10 there are limit values for short-term (24 hours) and long-term (annually) exposure. However, the consequent inability to identify a threshold below which adverse health effects are not observed implies that

However, a cautionary note is clearly in order before interpreting our results as the effects of PM10 on children's health outcomes. Since data on other pollutants is patchy in spatial and temporal coverage, we are not able to assess the influence of inversion on other pollutants in our setting; hence, our subsequent two stage least squares results of the effects of PM10 on health outcomes should be interpreted with care, because they may reflect the impact of e.g. nitrous oxides or particulates, or a combination of these pollutants. This is an endemic problem in the air pollution literature since all air pollutants are never measured. Because of this, we conservatively interpret the PM10 levels as a measure of air quality. Future studies with more diverse data on pollutants could examine which types of measured pollutants are most affected by inversion episodes. Notice that we also discuss the role of other pollutants further along with the reduced form results below.

We link the inversion data to the pollution data by assigning each pollution monitor to its closest AIRS grid centroid point located over land, and use the 24h average PM10 concentration as our air quality indicator.

### *iii) Health data*

Health measures were constructed utilizing both inpatient and outpatient data from the Swedish National Board of Health and Welfare (Socialstyrelsen), and cover all children living in Sweden in the age span of 0-18 years.<sup>10</sup> The inpatient data contains information on all visits to the health care providers which result in an overnight stay

---

any limit value may leave some residual risk when exposed to PM. This has led the World Health Organization (WHO) to recommend more stringent air quality guidelines (WHO, 2006).

<sup>10</sup>Young children are among the most susceptible to effects of air pollution (ALA, 2001; Kim et al., 2004). Compared to adults, children have higher breathing rates and therefore a higher intake of air pollutants per unit of body weight. Since children's lungs and immune system are not fully developed, exposure to air pollution opens up for the possibility of different responses than seen in adults. Furthermore, they also spend more time outdoors than adults when concentrations from air pollution are generally higher, thereby adding to their potential exposure. Since as much as 80 percent of alveoli are formed postnatally and the lung continues to develop throughout adolescence, exposure to air pollutants poses a serious risk to this population group(Schwartz, 2004).

at hospitals. The outpatient data captures visits where the patient does not stay overnight.<sup>11</sup> The data includes information on date of admission, type of diagnosis and municipality of residence. The ICD codes have been aggregated using the Clinical Classification Software (CCS) developed by the Agency for Healthcare Research and Quality (AHRQ).

Using individual identifiers, it is possible to link these administrative records to background data on individual characteristics, such as year and month of birth, place of residence, parents' income and educations from Statistics Sweden. Using these two data sources, we calculated the rate of health care visits due to respiratory illness by dividing the number of daily visits in each municipality by the total number of children residing in the municipality, multiplied by 10,000.

## 4.2 Summary statistics: Health, Weather and Pollution

Table 1 provides summary statistics of the variables described above. Panel A presents the means and standard deviations for the outcome variables used in the analysis. Information on the rate of health care visits is divided into cause of visit, and broken down by age. We also provide statistics on health care visits due to external causes. Since there exists no obvious causal pathway relating these accidents to air pollution or our instrumental variable, we later use the rate of visits caused by externally reasons in a check of the internal validity of our findings.

Panel B in Table 1 summarizes descriptive statistics for the key covariates used in our analysis. In order to shed some light on how pollution levels are affected by the occurrence of inversions, the weekly development of PM10 levels during night-time inversions as compared to normal nights is illustrated in Figure 2. As expected, the

---

<sup>11</sup>As all children in Sweden have access to free healthcare through taxation these measures of respiratory health problem should be highly valid.

pollution levels are significantly higher following inversion episodes. The level of PM10 accumulates over the week, increasing from Monday to Fridays and diminishes over the weekend. This weekend effect is due to decreased traffic volumes (Murphy et al., 2007), and occurs both during normal conditions and inversion episodes.

Figure 3 shows the monthly patterns of PM10 levels and shows peaks in March and April, much likely due to both residential and commercial heating, as these are in general cold months in Sweden. But this is also due to the fact that snow coverage on roads is largely absent during these months, while studded tires are used which increase road wear and which are also strongly associated with PM levels. During our sample period, most nocturnal inversions occur during the first (55 percent) and second (24 percent) quarter of the year. For the third and fourth quarter of the year, the corresponding frequencies were 6 percent and 15 percent, respectively. Moreover, on a monthly basis, the PM10 levels are on average higher during inversion episodes as compared to normal nights. From Figures 2 and 3, it is clear that the weekly and seasonal changes influence the behavior of pollutants under both inversions, as well as under normal conditions. This implies that both seasonal and weekly variations are important factors to take into account in our analysis since, for example, children's outdoor activity patterns are also likely to vary across weekdays and seasons.

For the 90 municipalities measuring PM10 in the period September 2002 to September 2007, there are 34,175 valid vertical temperature profiles, that is when temperature readings are non missing in both layers. Out of these, 8,608 night-time inversions were identified. Descriptive statistics for the key variables conditional on inversion status are provided in Table 2. Comparing normal nights with inversion nights, we conclude that the health care visit rate is, on average, 4.7 percent higher during the latter. Similarly, the PM10 level is on average 59.6 percent higher throughout inversion nights, confirming the pattern in Figures 2 and 3.

As expected, both daily and nightly temperatures are, on average, higher during normal conditions, since inversion episodes are more frequent during the winter months than during the summer months. A similar pattern can be observed for the remaining four weather variables. These mean differences in weather conditions across seasons highlight the importance of flexibly accounting for season of the year, day of week patterns, and weather conditions in the analysis for our identification strategy to provide valid causal inference. Our econometric specification is presented next.

## 5 Econometric framework

The purpose of this paper is to study the causal relationship between air quality ( $PM10_m$ ) on child health (the rate of hospital visits related to respiratory illnesses per 10 000 children). To motivate the use of instrumental variables, let us first consider the structural equation of the effect of PM10 on health, where we, for brevity, omit the time aspect:

$$Health_m = \beta_0 + \beta_1 PM10_m + e_m \quad (3)$$

$\beta_1$  is the parameter of interest and our primary objective is to test whether  $\beta_1 \neq 0$ . In other words, if exposure to air pollution has an effect on respiratory illness related hospital visits in municipality  $m$ .  $e_m$ , is the error term. Under the identifying assumption that the error term is uncorrelated with pollution exposure,  $PM10 \perp e_m$ , the ordinary least square estimate of  $\beta_1$  reflects the causal impact of one additional unit of PM10 per  $m^3$  on the rate of health care visits due to respiratory illnesses.

As pointed out in previous studies and as discussed above, several factors complicate the estimation of this relationship and we have strong reasons to question whether the exogeneity assumption holds. To deal with the endogeneity of residential location,

weather conditions, measurement error, and avoidance behavior we use the incidence of nighttime inversions as an instrumental variable in a municipality by date panel data framework. We estimate the following two equations by Two Stage Least Squares (2SLS):

$$Respiratory\ Illness\ Rate_{md} = \beta_0 + \beta_1 PM10_{md} + \alpha' W_{md} + \eta_d + \theta_m + e_{md} \quad (4)$$

$$PM10_{md} = \gamma_0 + \gamma_1 Inversion_{md} + \rho' W_{md} + \eta_d + \theta_m + v_{md} \quad (5)$$

where (4) is the second-stage and (5) is the first-stage equations.  $Inversion_{md}$  is our binary instrumental variable indicating whether an inversion occurred in municipality  $m$  at date  $d$ , or not. Both equations contain a vector of weather and time changing demographic characteristic  $W_{md}$  where  $\alpha_1$  and  $\rho_1$  are corresponding parameter vectors that will be estimated. We include in  $W_{md}$ , precipitation, wind speed, humidity, cloud cover, and their squared counterparts, together with daily and nightly temperature polynomials to account for a potential nonlinear relationship between temperature, pollution levels and respiratory illnesses. We also include timevarying variables such as the average age of the children in the municipality and the share of mothers with college degrees as additional controls in  $W_{md}$ .

$\eta_d$  is a set of common year by month effects and day-of-week effects that nonparametrically accounts for year-specific seasonal effects and weekday variations in pollution and respiratory illnesses. The year-month effects take into account, for example, year-specific seasonal patterns in respiratory illnesses associated with e.g. influenza outbreaks.  $\theta_m$  are municipality-specific effects which are included in order to account for permanent differences across municipalities affecting pollution concentrations and respiratory illnesses (e.g. time invariant demographic characteristics, industry composition, altitude and other geographic conditions).

In all estimations the errors are clustered at the municipality level, since all children in the same municipality are exposed to the same levels of predicted pollution, which are measured with error and are likely to be correlated over time within municipality.

### *Two Specification Issues*

Two general specification issues are worth noting before discussing the validity of the inversion episodes as an instrument variable. First, it is far from obvious that the linear specification as outlined in equation (2) is the appropriate functional form between PM10 and respiratory illnesses. Nonlinearities in the effects of PM10 are interesting for regulatory threshold purposes. But they are also of interest for the extrapolation of our low pollution setting results to higher pollution settings, since if strong nonlinearities are present, there is a clear risk that our estimates would underestimate the predicted effects in higher pollution regions. Nonlinearities could also be important when trying to understand the differences across socio-economic for the reasons discussed above.

Below, we provide non parametric estimates of the reduced form and first stage from a generalized additive model (Hastie and Tibshirani, 1986) using a local linear smoother with a narrow bandwidth to assess whether our baseline linear estimator is appropriate. This highly flexible model, which non parametrically takes daily weather conditions and seasonal patterns into account, provides estimates that are highly similar to the results from our 2SLS estimator outlined above. As shown below, the non parametric estimates provide no indication of strong nonlinearities in the effects of PM10 on respiratory illnesses in our setting.

Second, note that we focus on the contemporaneous effects of changes in air quality on health care visits due to respiratory illnesses. A potential concern with this daily specification is that temporary increases in pollution levels may simply displace the timing of respiratory illnesses forward. Such a short term forward shift in the timing

of health effects, so-called harvesting effects, would imply that while we may see an increase in health effects on high pollution days, the respiratory illness rate may fall and be fully compensated for over the following days. To check the validity of this concern, we estimate distributed lag model assessing whether a lagged increase in pollution following inversion episodes reduces the current respiratory illnesses rate. We also checked if instrumenting the PM10 levels during the past few days using the share of inversion nights during those days changes the results. We find no indications of strong displacement effects.

#### *Instrument validity*

Before proceeding to the results, we conclude this section with some additional remarks on identification with respect to inversion episodes as an instrumental variable. The exclusionary restriction may not hold if sensitive individuals are able to correctly understand and predict inversion occurrences and their effects on pollution levels. We believe this to be unlikely in general, and particularly in our setting. Information or predictions of inversion episodes are not available to the public. Neither information on inversion, nor inversion strength, are published in Swedish media or by local authorities, and vertical temperature profiles were not available on a large scale, nor was the data from the four Swedish ground level sounder stations used in pollution level forecasts.<sup>12</sup>

Yet, it is possible that particularly sensitive populations may have strong enough incentives to gather private information on inversion episodes since it is such a strong predictor of air quality. In some heavily polluted areas around the world, temperature inversions can sometimes be observed with the naked eye. In Sweden, this is generally not the case due to relatively low pollution levels and low humidity. But to address

---

<sup>12</sup>Personal communication with Michael Norman at IVL, where the pollution data is stored and provides the pollution prognosis for Stockholm, September 6 2013. Access to SMHI's and the Swedish Military's in total 4 weather balloon stations that measure vertical temperature profiles on a daily basis is under way according to SMHI but is not available at present (2013-11-04) .

this potential concern, we focus on the effects of night-time inversion episodes. Night-time inversion episodes are also more frequent and constitute a stronger predictor of PM10 levels in our setting.<sup>13</sup> As discussed above, previous studies from high pollution settings (Los Angeles) have suggested that avoidance behavior may underestimate the effects of pollution on health (Neidell, 2009; Moretti and Neidell, 2011). Hence, for our estimation strategy to provide estimates that hold avoidance behavior fixed, it is crucial that individuals cannot perfectly observe inversion status and inversion strength and adjust their behavior accordingly. By using night-time inversions, we believe that the risk of instrument observance is minimized. Even if individuals understand the meteorological relationship in question in general, it seems unlikely that they are able to correctly identify inversion episodes and inversion strength at 2 am in the morning. For these reasons, we believe that the instrumental variable approach used here should be able to provide causal estimates that hold avoidance behavior fixed.

However, we also provide indirect evidence to try to back up this assumption. Any test of the extent and prevalence of avoidance behavior relies on proxies (Graff Zivin and Neidell, 2012). We check whether children's activity patterns change during inversion episodes by examining whether the health care visit rate due to injuries with external causes changes during inversion episodes by estimating the following modified version of the second-stage equation

$$Rate\ External\ Causes_{md} = \beta_0 + \beta_1 PM10_{md} + \alpha' W_{md} + M_d + \theta_m + e_{md} \quad (6)$$

The idea is that if inversions are associated with substantial changes, not only in pollution levels but also in e.g. children's indoor/outdoor activity patterns, we may detect

---

<sup>13</sup>And second, as expected from the effects on air quality, when entered in the same equation as the night-time inversion indicator, the reduced form effects on respiratory illnesses are negligible in size and insignificant, unlike the night-time inversion indicator which remains strong and positive.

that injuries due to external causes also change, i.e. that  $\beta_1 \neq 0$  in equation (6). In other words, if inversion episodes do not affect our proxy of avoidance behavior, we would expect that  $\beta_1 = 0$  in equation (4).<sup>14</sup>

## 6 Baseline Results

We begin by providing an illustration of the identification strategy from non parametric estimates of the reduced form and the first stage, and then report the first stage and reduced form result from the fully specified linear model in Section 5.1. Throughout, we provide a comparison of the ordinary least square (OLS) and the two-stage least squares (2SLS) estimate. In section 5.2, we examine the robustness of the results to changes in the specification and in section 5.3 we look at the impact across socio economic groups and document a substantial gap in the effects between children in households with high and low income. In section 5.4, we examine the potential causes of this gap.

### 6.1 First Stage and Reduced Form

We begin by presenting nonparametric estimates of the first stage and the reduced form. Figure 4 shows the narrow bandwidth multivariate local linear regression estimates of the first stage (PM10 level on inversion strength) and the reduced form (respiratory illness rate on inversion strength) that flexibly control for ground level temperature

---

<sup>14</sup>Finally, instrument validity may theoretically also be compromised by reversed causality in the first stage equation. That is, if emissions levels are an important determinant of ground level air temperature, then we could find that higher pollution levels reduce the incidence of inversion episodes. However, first, it seems unlikely that local anthropogenic emissions have a strong differential impact on temperature in the two layers. Second even if so, our estimates would likely be downward biased since local emissions primarily heat the ground level air layer, reducing the occurrence of inversion. Third, we in Table A1 we test the likely severity of this concern by exploiting the well known variation pollution levels over the weekdays. See section 5.2 below.

and calendar month.<sup>15</sup> The x-axis displays the inversion strength, i.e. the temperature difference between the two air layers. Negative inversion strength values correspond to the relationship between the outcome variables and the temperature differences on normal days, and positive values to the relationship on inversion days. The left-hand side y-axis measures the average PM10 level and the right-hand side y-axis measures the respiratory health care rate. The kernel density estimate (dashed) shows the distribution of observations with respect to the instrument. Around 25 percent of the days in the sample are inversion days.

Figure 4 shows that, conditional on ground level temperature and the calendar month, under normal circumstances the rate of hospital visits (black) and PM10 levels is not strongly correlated with the temperature difference between the two air layers. However, once an inversion occurs, both PM10 levels and the respiratory illnesses rate increase almost linearly with the strength of the inversion. This first set of results and the implied elasticity between PM10 and health care visits summarize the main results of the paper. As we will see, even after adding a large set of additional control variables, the estimated elasticity never deviates substantially from that which can be inferred from this parsimonious, but highly flexible non parametric specification.

Table 3 reports the fully specified linear version of the reduced-form, the first stage, OLS, and the IV specifications. In column (1), the PM10 levels are regressed on the inversion dummy, and the other control variables, year by month time effects, and municipality fixed effects as specified above in equation (2). The point estimate suggests

---

<sup>15</sup>We begin by presenting nonparametric estimates of the first stage and the reduced form. Figure 4 shows the narrow bandwidth multivariate local linear regression estimates of the first stage (PM10 level on inversion strength) and the reduced form (respiratory illness rate on inversion strength) that flexibly control for ground level temperature and calendar month. The x axis displays the inversion strength, i.e. the temperature difference between the two air layers. Negative inversion strength values correspond to the relationship between the outcome variables and the temperature differences on normal days, and positive values to the relationship on inversion days. The left-hand side y-axis measures the average PM10 level and the right-hand side y axis measures the respiratory health care rate. The kernel density estimate (dashed) shows the distribution of observations with respect to the instrument. Around 25 percent of the days in the sample are inversion days.

that the PM10 level is, on average,  $6.15 \mu\text{m}/\text{m}^3$  higher following night-time inversion episodes. Hence, relative to the mean PM10 level, inversion episodes increase the PM10 levels by approximately 30 percent. In column (2), the health care visit rate is regressed on the inversion dummy and the other control variables. As expected, health care visits are positively and significantly related to inversions. During inversion episodes, the respiratory illness health care visit rate increases by close to 5 percent.

Column (3) reports the standard OLS estimates from the fully specified model and column (4) the instrumental variable estimate. Similar to previous studies using an instrumental variable approach that accounts for measurement error and avoidance behavior, the instrumental variable approach IV estimates are up to four times larger than the OLS estimates. Relative to the mean respiratory illness rate, the IV estimate implies that for each 10 microgram/ $\text{m}^3$  increase in daily PM10 levels, health care visits due to respiratory illnesses increase by 8 percent. The corresponding figure from a simple OLS estimate is 2 percent. These huge differences underscore the importance of properly accounting for measurement error, avoidance behavior and other endogeneity problems.

## 6.2 Robustness

The OLS and IV comparison continues in Table 4 where, for comparison, the baseline model estimates for OLS and the IV are once more reported in column (1). In column (2), we add average age and age squared of the children in the municipality and year to the baseline model, and in column (3) we add information on the share of mothers with college education in the municipality of residence. None of these modifications significantly changes the estimated effects. In column (4), we restricted the sample to children living within 2 km of the nearest pollution monitor. Notice that since pollution monitors are located in downtown areas, the mean of respiratory related hospital

admissions is higher. Moreover, notice that when restricting the sample to only those close to the monitors, the OLS estimate changes dramatically ( $\approx 33$  percent lower), while the IV estimate changes marginally ( $\approx 6$  percent lower). The differences in the robustness of the results between OLS and IV are an indication of the validity of the research design, since if unobservable covariates are uncorrelated to the instrument but not to PM10 levels, we would expect the OLS to be more sensitive to changes in the specification than the IV estimate. In column (5), we assess the robustness of the estimates with respect to the included weather variables. If we drop the precipitation and wind variables from the ground level weather stations, and instead only use the NASA satellite weather data, the estimates hardly change at all. Hence, the IV-approach could be implemented only using the NASA data.

Table A1 in Appendix A report additional specification checks intended to capture cumulative/displacement effects. Column (1) review the baseline IV estimates for comparison, Column (2) report the estimate when using the average PM10 levels between  $t$  and  $t-5$  using the share of days with inversion as an instrument on day  $t$  respiratory illness rates. Column (3) report the cumulative effect of PM10 over the last 5 days, i.e. the sum of the lagged coefficients from a distributed lag model where each lag is instrumented with the inversion status of that particular day. If harvesting effects were an important concern in the current setting we would expect to find that the cumulative effect over the recent past to be smaller than the effect from the baseline contemporaneous model (i.e.  $\sum_{k=0}^5 \beta_{t-k}^{full} < \beta_t^{base}$ ). The impact of the average PM10 level over the past 5 days should also be smaller than the contemporaneous effect. If anything both the estimates are larger, but they are not statistically distinguishable from the baseline estimates.<sup>16</sup>

---

<sup>16</sup>Except for the contemporaneous effect, none of the lagged coefficients in the distributed lagged model are statistically significant, and do not follow any obvious pattern. Note that missing values on any of the past days leads to that the day  $t$  is dropped from the distributed lag estimation, leading to a substantially smaller sample size. We also estimated a model using full set of contemporaneous

Finally, the last two columns of table A1 assess the relevance of reversed causality in the first stage equation. To gauge the severity of this concern, we check whether pollution levels affect inversion episodes by exploiting the well documented decrease in pollution levels during weekends. In particular we regressed the inversion dummy on a weekend dummy and the other variables in the main model (excluding the day of week effects) Column (4) show that PM10 decreases by around 23 percent (-4.754/20.8) during weekends. But despite this huge improvement in air quality, the frequency of inversion episodes are not affected; the point estimate is not statistically significant, positive, and close to zero. Thus, reversed causality in the first stage equation does not seem to be a major concern.

### **6.3 Heterogeneity: by respiratory diagnosis and child age**

So far, all respiratory illnesses have been lumped together, but it is possible that the effect of PM10 varies across type of respiratory illness. In Table 5, we supply estimates of the effect of PM10 separately across type of illnesses. In columns (2) to (5), we separately explore the effect of air pollution by type of respiratory illness. We split the respiratory illness diagnosis data into Pneumonia, Bronchitis, Asthma, and Other respiratory conditions. The first three conditions are likely to be exacerbated from current exposure as opposed to conditions such as Emphysema where the effect of exposure is cumulative over time. Therefore, we expect to find the largest effects for the first three conditions. This pattern is also largely confirmed by both the OLS and the IV estimates, with the highest percent effect for Asthma, Bronchitis and Pneumonia, and substantially lower but still significant for the other respiratory conditions.

Columns (6) and (7) provide separate estimates for pre-school children (0-5) and

---

weather controls but only lagged ground level temperature to minimized missing values. Results were very similar. The full set of estimates are available on request. Schenkler and Walker (2011) use a similar specification and find similar patterns.

school age children (6-18). Note that the mean respiratory illness rate is higher for children in the age span 0-5 years. Relative to the mean respiratory illness, the OLS estimates are higher for the older children than for the younger children. The IV estimates are positive and significant with the largest percentage effect on the increase of respiratory related hospital admissions for the older age group.

However, note the differences between the OLS and the IV estimates in the two age groups. The IV estimate is about 7 and 3 times larger than the OLS for the younger and the older children respectively, potentially indicating that avoidance behavior is more important among (the parents of) the younger children.

#### 6.4 Examining the Effects across Socio-Economic Groups

Next we report estimates by parental socio economic characteristics. In Table 6, we report the separate estimates for the sample of children with mothers having more than high school education and for those who do not. We do not find any substantial differences in the effects with respect to maternal education (columns 2-3). However, when turning to columns 4 and 5, we see clear evidence of the differential effect, depending on household income. The IV estimates for low-income households (below median income) suggest that the effects of a  $10 \mu g/m^3$  increase in PM10 are around twice as large as on children in high-income households (12 percent vs. 6 percent). Figure 5 illustrates the differences by SES-groups graphically.

The finding that children from low-income families are much more affected by changes in air quality than children from high income families clearly suggests that policies that reduce pollution levels may also reduce the inequality in childhood health. As highlighted in the conceptual framework outlined above, several potential mechanisms have previously been brought forth that could potentially explain SES differences in the effects of air pollution. A better understanding of the underlying mechanisms

could be informative for understanding which types of policies that are likely to not only reduce the effects of pollution, but potentially also reduce the socioeconomic differences in the impact. In the following section, we first discuss how our previous results square with the suggested mechanisms and use the data at hand to try to quantify the importance of the different mechanisms.

#### 6.4.1 Evidence on the mechanisms behind the SES-gap

##### (i) Non-linearities

An important difference between rich and poor households could be that residential segregation leads to differences in average levels of pollution exposure. In the case of road traffic, if environmental disamenities such as air pollution and noise from traffic are valued in the housing market, this would lead children from poorer households to be located, all else equal, closer to major roads. Since local road traffic contributes to a substantial share of the total PM10 exposure, a potentially important mechanism behind the SES-gap could be that children from poorer households are exposed to higher pollution levels than children from wealthier households. Since we measure average PM10 levels at the municipality of residence level, the observed SES differences could hence stem from nonlinearities in the effects of PM10.

However, this interpretation squares poorly with the results from the non parametric estimates in Figure 4 that give no indication of strong nonlinearities between PM10 and respiratory illnesses. Still, since we are only able to directly assess the effects of PM10 on health, careful readers may worry that inversion episodes could generate a substantial increase in other, unobserved, pollutants which, in turn, may be nonlinearly related to respiratory illnesses. However, again, if non linearities in other unobserved pollutants are important, we would expect that the reduced form relationship between inversion strength and respiratory illnesses should also be strongly non linear during

inversion episodes in Figure 4, which is not the case. In summary, at the relatively low pollution levels considered, our results do not provide any strong evidence in favor of the hypothesis that the differences in the observed effects between rich and poor households are driven by non linearities in the effects of air quality.

### *(ii) Avoidance Behavior*

Another reason why children from differing socio-economic backgrounds could suffer more from changes in pollution levels is differing parental responses to changes in pollution levels. As discussed above, avoidance behavior could influence estimates of the effects of air quality on respiratory illnesses. If high socio-economic status parents are more likely to engage in avoidance behavior, the SES-gap in the effects of ambient air pollution could stem from differences in avoidance behavior.

Note first that, as argued above, our identification strategy is explicitly designed to hold parental avoidance behavior fixed. Second, to the extent that knowledge about risks of high pollution episodes following inversion episodes is more prevalent among highly educated parents than low educated parents, we would expect to see differences in the effects across households with differing educational attainment. As shown by Table 6 and Figure 5, there is no indication of differential effects across families with different educational attainments.

However, we also try to directly assess the prevalence of avoidance behavior. Any test of the extent and prevalence of avoidance behavior relies on proxies (Graff Zivin and Neidell, 2012). Previous studies examining avoidance behavior have used visits to outdoor facilities (e.g Zoos or sporting events) in connection with warnings of predicted high ozone episodes (Neidell, 2009; Moretti and Neidell, 2011) in high pollution settings.

Here, we use changes in in- and out patient health care visits due to externally caused injuries to proxy for changes in children's activity patterns. The idea is that if

inversion episodes are associated with substantial changes in children's activity patterns, for example that parents are keeping their children indoors instead of allowing them to play outside to reduce exposure, we would expect that the risk of accidental falls, accidents with cyclists, or motor vehicles, etc. could change as well. Moreover, the externally caused injury rate per 10,000 children has a similar mean and distribution as the respiratory illness rate.

Since there is no obvious direct relationship between air quality and externally caused injuries except through changes in activity patterns, we would expect to see that changes in air pollution also change the risk of injuries due to external causes if avoidance behavior is highly prevalent. Moreover, for avoidance behavior to be a strong explanatory factor for the observed SES gap, children from richer households should experience a substantially stronger relative change in the externally caused injury rate.

The estimates in Table 7 column (1) show the effect for the full sample and columns (2) and (3) show the results for children in rich and poor households separately, a non significant and small correlation between changes in PM10 and externally caused injuries. The percent effect for respiratory illness is more than six times larger than the effect on externally caused injuries. Moreover, there are no differences between children in low- and high-income households.

These results first of all indicate that differences in avoidance behavior are not likely to be the cause of the SES-gap in our low-pollution context. This result is consistent with the absence of differential effects by maternal education. If knowledge of risks of high pollution episodes is more prevalent among highly educated parents, we would expect to see differences in effects across households with differing educational attainment; which we do not.

In addition, the fact that the rate of externally caused injuries does not create any change in response to high pollution episodes as instrumented by inversion episodes

can also been seen as a falsification test of our identifying assumptions. It indirectly validates that our instrumental variable estimates are not affected by e.g. changes in activity patterns caused by or other unobserved variables correlated with both inversion episodes and children's respiratory illnesses. In other words, it seems reasonable to assume that our instrumental variable estimates shift pollution exposure while parental responses are held fixed.

### *(iii) Differences in Baseline Health*

Children from poorer backgrounds have worse health in general (Currie et al., 2010). If children with poorer general health are more susceptible to pollution shocks, the SES gap in the effects of pollution could in part be explained by differences in children's baseline health across rich and poor households.

To assess the relevance of the hypothesis of differences in health capital between children in high- and low-income households, we make use of data on health at birth. Health at birth data is useful since it is a strong predictor of subsequent health in childhood and beyond, but also because this measure has been collected in a similar manner for all cohorts and is available for virtually all children in our sample. Specifically, we construct an index of initial health using the first principal component of gestational week at birth and birth weight. The children are then split into a good and poor health status group if they are above or below the median of the initial health index.

Table 8 presents the results for the full sample split into high and low health children in columns (1) and (2). Columns 3 through 6 then show the results after splitting the sample further into high- and low-income of the parents. The first two columns show that the IV effects of pollution are larger for children with poorer initial health than children with better initial health, while the OLS differences are negligible. An increase of  $10\mu g/m^3$  PM10 increases the respiratory illness rate by 4 percent for chil-

dren with good health and high-income parents, but are around three times larger for children with good health but who have poorer parents. The two final columns in Table 9 show that this huge SES gap decreases substantially if comparing children with worse initial health in rich (9 percent) and poor (11 percent) households. Note also that the standard errors are higher for children in low-income households with good initial health than for low-income household children with poor initial health, and vice versa for children from high-income households. This reflects the fact that children from low-income households are substantially more prevalent in the poor initial health group, and high-income household children are more prevalent in the high initial health group.

### **Summary of the evidence on the underlying mechanisms**

In summary, out of the mechanisms suggested to be able to explain differential effects of ambient air quality on children's health, differences in initial health across households with differing economic conditions seem to be key in explaining the differential impact of air quality on respiratory illnesses. However, since our identification strategy is designed to hold avoidance behavior fixed, our estimates reflect the "biological" effects. In high pollution settings where information differences across socio economic groups are strong, it is likely that avoidance behavior plays a more prominent role in generating differences in the effects of air pollution across socio economic groups.

Finally, an additional mechanism could be that parents with more resources may have greater opportunities for, or receive higher returns (e.g. by reducing the lost income due to work absence during children's illness spells) from, medical defensive investments. A higher level of defensive investments may reduce the need for visits to health care facilities in connection with high pollution days and hence, this could also contribute to parts of the SES-gap in the effect on health care visits. There is little

evidence of the importance of defensive investments with respect to air quality, but Deschenes et al. (2012) document that when ozone levels decrease, so does medication expenditure. To our knowledge, there are no studies examining whether defensive investments related to respiratory illnesses differ for children between socio-economic groups.

Difference in medical defensive investments is difficult to completely rule out as an additional explanatory factor given the lack of data on daily medical consumption. However, in Sweden all children have health insurance, and medical expenses of children (under age 18) are all fully subsidized if the sum of expenses of all children in the same family exceeds SEK 2200 ( $\approx$  USD 320) for 12 months after the date when the threshold is exceeded. Health care visits are free of charge for children aged below 19. Hence in our setting, it seems unlikely that differences in defensive investments constitute a major contributing factor to the SES gap in the effects of air quality on health.

## 7 Conclusions

The main contribution of this paper is the instrumental variable approach we develop for causal inference in short term effects of poor air quality. The method can be used on a global scale with a high spatial and temporal resolution. As an illustration of the usefulness of the data, we examine the short run effects of pollution on children's respiratory illnesses in Sweden. The results are robust to a number of alternative specifications and suggest that poorer air quality significantly increases health care visits due to respiratory illnesses for children aged below 18. There is a striking difference in the costs associated with high pollution levels as predicted by the OLS and IV estimators. OLS estimates of particle pollution on health care visits are statistically significant although small in magnitude. In contrast, IV estimates accounting for avoidance behav-

ior and measurement errors indicate a four times larger toll of pollution on respiratory illnesses.

Moreover, the differences in the response to air quality changes across children in high and low-income households clearly suggest that environmental policies that curb local pollution levels could reduce the persistence of inequalities in health. In our setting, despite heavily subsidized and universal access to preventive care, differences in economic resources seem to be more important than information differences across households. The non parametric estimates show no indication of strong nonlinearities in the effects of PM10. Hence, our examination of the channels through which the differential effects arise suggests that baseline health differences between children in rich and poor households are key. Since pollution exposure early in life has also been shown to influence cognitive ability, educational attainments and outcomes in the labor market (Nilsson, 2009; Sanders, 2012), it seems that environmental policies could also play an important role in reducing inequalities in economic outcomes. More research on long term effects of early life air pollution exposure and the interaction with parental income is of clear interest.

Note that since we only examine the impact on health care visits and disregard other outcomes, such as parental work absence or use of pharmaceuticals, they represent a lower bound of the total welfare costs of poor air quality. In addition, while the focus on short-term changes in air quality aids the identification, our results only provide limited insights into the effect of environmental policies that permanently improves air quality. Since health is a stock variable (Grossman, 1972), it is possible that the full health benefits of a permanently lower pollution level could be much higher than what is suggested by our short-term effect estimates. Further research is needed on health effects of policies that “permanently” change pollution levels. In addition, while the focus of the current paper is not about identifying potential heterogeneous SES effects

from different air pollutants, future studies with better data on different pollutants could examine the effects inversion episodes have on other air pollutants besides PM10.

But the method could also be used to shed some additional light on effects of air quality on e.g. labor supply on the extensive (Hanna and Oliva, 2011) and intensive margin (Graff Zivin and Neidell, 2009) among adults. These studies have used panel data methods to quantify the effects of poor air quality on labor supply in a setting with high pollution levels and for farm workers, respectively. The NASA vertical temperature profiles data opens up the possibility of comparative studies from areas with widely differing pollution levels, different economic circumstances, and in various occupations. Using similar identification strategies across studies could reduce the risks of mistaking treatment heterogeneity across locations and/or populations with differences in estimation strategies. The inversion data could also be used in combination with more location-specific instrumental variables to boost precision of estimates.

Finally, the AIRS data can be incorporated in models predicting pollution levels at many more locations than what is currently possible using ground based vertical temperature sounders (e.g. available in four locations in Sweden). By developing methods to produce more precise location specific pollution forecasts and effective means of disseminating such forecasts, the health care costs associated with respiratory symptoms could potentially be reduced by allowing sensitive populations to more effectively engage in defensive investments.

## References

- Abdul-Wahab, S. A., Bakheit, C. S., and Siddiqui, R. (2005). Study the relationship between the health effects and characterization of thermal inversions in the Sultanate of Oman. *Atmospheric Environment*, 39(30): 5466–5471.
- ALA (2001). Urban air pollution and health inequities: a workshop report. *Environmental Health Perspectives*, pages 357–374.
- Almond, D. (2006). Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population. *Journal of Political Economy*, 114(4): 672–712.
- Almond, D., and Currie, J. (2011). Human capital development before age five. *Handbook of labor economics*, 4: 1315–1486.
- Almond, D., Edlund, L., and Palme, M. (2009). Chernobyl’s subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden. *The Quarterly Journal of Economics*, 124(4): 1729–1772.
- Arceo-Gomez, E. O., Hanna, R., and Oliva, P. (2012). Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City. Working Paper 18349, National Bureau of Economic Research.  
URL <http://www.nber.org/papers/w18349>.
- Beard, J. D., Beck, C., Graham, R., Packham, S. C., Traphagan, M., Giles, R. T., and Morgan, J. G. (2012). Winter Temperature Inversions and Emergency Department Visits for Asthma in Salt Lake County, Utah, 2003–2008. *Environmental health perspectives*, 120(10): 1385.

- Case, A., Lubotsky, D., and Paxson, C. (2001). Economic Status and Health in Childhood: The Origins of the Gradient. Working Paper 8344, National Bureau of Economic Research. URL <http://www.nber.org/papers/w8344>.
- Chay, K. Y., and Greenstone, M. (2003a). Air Quality, Infant Mortality, and the Clean Air Act of 1970. Working Paper 10053, National Bureau of Economic Research. URL <http://www.nber.org/papers/w10053>.
- Chay, K. Y., and Greenstone, M. (2003b). 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.
- Cunha, F., and Heckman, J. (2007). The Technology of Skill Formation. Working Paper 12840, National Bureau of Economic Research. URL <http://www.nber.org/papers/w12840>.
- Currie, J. (2008). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. Working Paper 13987, National Bureau of Economic Research. URL <http://www.nber.org/papers/w13987>.
- Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M., and Rivkin, S. G. (2009). Does pollution increase school absences? *The Review of Economics and Statistics*, 91(4): 682–694.
- Currie, J., and Neidell, M. (2005). Air pollution and infant health: What can we learn from California’s recent experience? *The Quarterly Journal of Economics*, 120(3): 1003–1030.
- Currie, J., Stabile, M., Manivong, P., and Roos, L. L. (2010). Child health and young adult outcomes. *Journal of Human Resources*, 45(3): 517–548.

Deschenes, O., Greenstone, M., and Shapiro, J. S. (2012). Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program and Ozone Reductions. Working Paper 18267, National Bureau of Economic Research. URL <http://www.nber.org/papers/w18267>.

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

EEA (2012). Air Quality in Europe - 2012 report. *European Environment Agency Report*, 4.

Ghio, A. J., and Devlin, R. B. (2001). Inflammatory lung injury after bronchial instillation of air pollution particles. *American journal of respiratory and critical care medicine*, 164(4): 704–708.

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

Graff Zivin, J., and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102: 3652–3673.

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy*, 80(2): 223–255.

Hanna, R., and Oliva, P. (2011). The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City. Working Paper 17302, National Bureau of Economic Research. URL <http://www.nber.org/papers/w17302>.

Hastie, T., and Tibshirani, R. (1986). Generalized additive models. *Statistical science*, pages 297–310.

Isen, A., Rossin-Slater, M., and Walker, W. R. (2013). Every Breath You Take - Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970. Working paper.

Kim, J., et al. (2004). Ambient air pollution: health hazards to children. *Pediatrics*, 114(6): 1699.

Knittel, C. R., Miller, D. L., and Sanders, N. J. (2011). Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. Working Paper 17222, National Bureau of Economic Research. URL <http://www.nber.org/papers/w17222>.

Kukkonen, J., Pohjola, M., S Sokhi, R., Luhana, L., Kitwiroom, N., Fragkou, L., Rantamäki, M., Berge, E., Ødegaard, V., Håvard Slørdal, L., et al. (2005). Analysis and evaluation of selected local-scale PM10 air pollution episodes in four European cities: Helsinki, London, Milan and Oslo. *Atmospheric Environment*, 39(15): 2759–2773.

Lleras-Muney, A. (2010). The Needs of the Army Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health. *Journal of Human Resources*, 45(3): 549–590.

Malek, E., Davis, T., Martin, R. S., and Silva, P. J. (2006). Meteorological and environmental aspects of one of the worst national air pollution episodes (January, 2004) in Logan, Cache Valley, Utah, USA. *Atmospheric research*, 79(2): 108–122.

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

Murphy, J. G., Day, D. A., Cleary, P. A., Wooldridge, P. J., Millet, D. B., Goldstein, A. H., and Cohen, R. C. (2007). The weekend effect within and downwind

of Sacramento—Part 1: Observations of ozone, nitrogen oxides, and VOC reactivity. *Atmospheric Chemistry and Physics*, 7(20): 5327–5339.

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

Nel, A. E., Diaz-Sanchez, D., Ng, D., Hiura, T., and Saxon, A. (1998). Enhancement of allergic inflammation by the interaction between diesel exhaust particles and the immune system. *Journal of Allergy and Clinical Immunology*, 102(4): 539–554.

Nilsson, J. P. (2009). The long-term effects of early childhood lead exposure: Evidence from the phase-out of leaded gasoline. Working paper, Institute for Labour Market Policy Evaluation (IFAU).

Nilsson, P. (2008). Does a pint a day affect your child's pay? The effect of prenatal alcohol exposure on adult outcomes. Cemmap Working Paper CWP22/08, London. DOI 10.1920/wp.cem.2008.2208. URL <http://hdl.handle.net/10419/64679>.

Pope III, C. A., Schwartz, J., and Ransom, M. R. (1992). Daily mortality and PM10 pollution in Utah Valley. *Archives of Environmental Health: An International Journal*, 47(3): 211–217.

Sanders, N. J. (2012). What doesn't kill you makes you weaker prenatal pollution exposure and educational outcomes. *Journal of Human Resources*, 47(3): 826–850.

Schlenker, W., and Walker, W. R. (2011). Airports, Air Pollution, and Contemporaneous Health. Working Paper 17684, National Bureau of Economic Research. URL <http://www.nber.org/papers/w17684>.

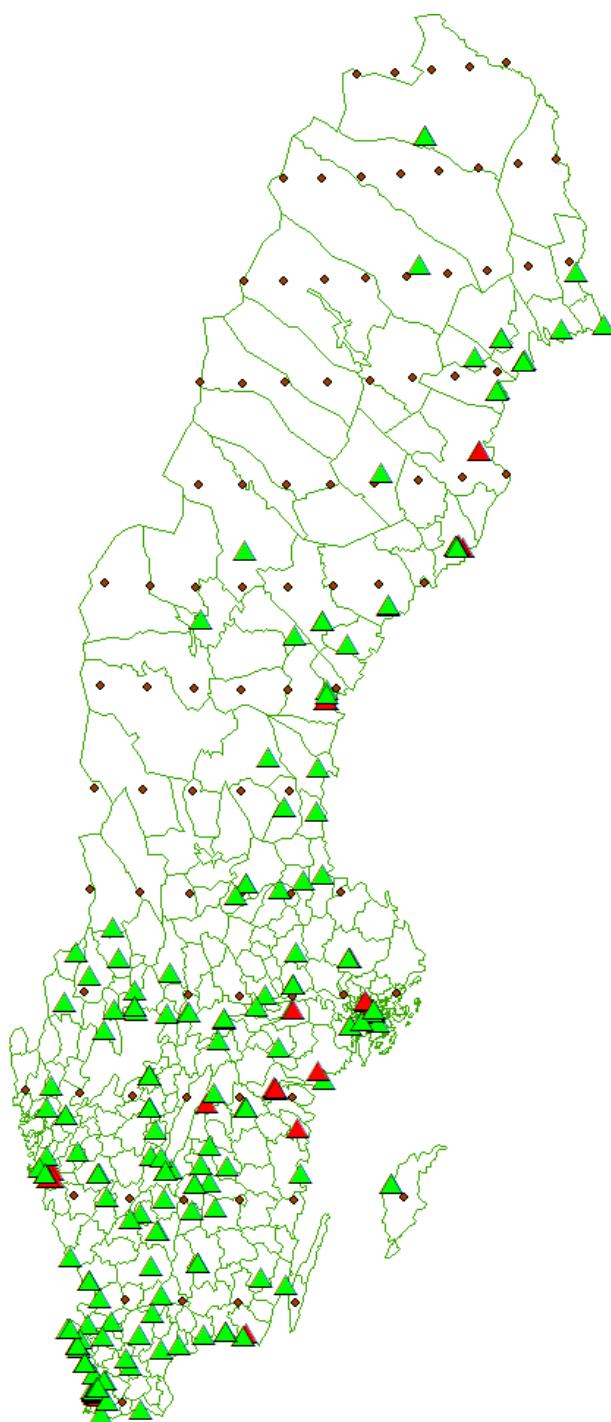
Schwartz, J. (2004). Air pollution and children's health. *Pediatrics*, 113(Supplement 3): 1037–1043.

SEPA (2012). The Swedish Environmental Protection Agency. Accessed December 2012.

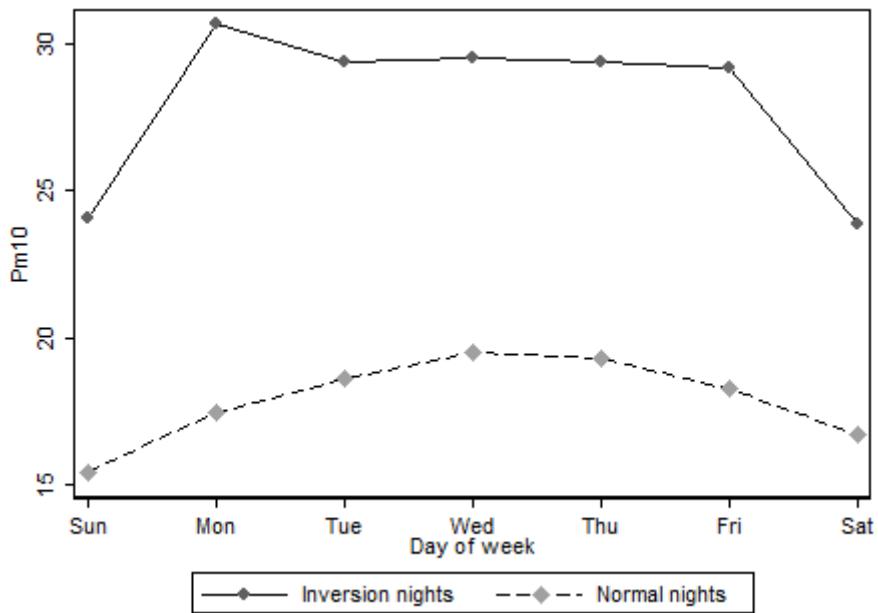
URL <http://www.naturvardsverket.se/Sa-mar-miljon/Statistik-A-0>.

Wallace, J., Nair, P., and Kanaroglou, P. (2010). Atmospheric remote sensing to detect effects of temperature inversions on sputum cell counts in airway diseases. *Environmental research*, 110(6): 624–632.

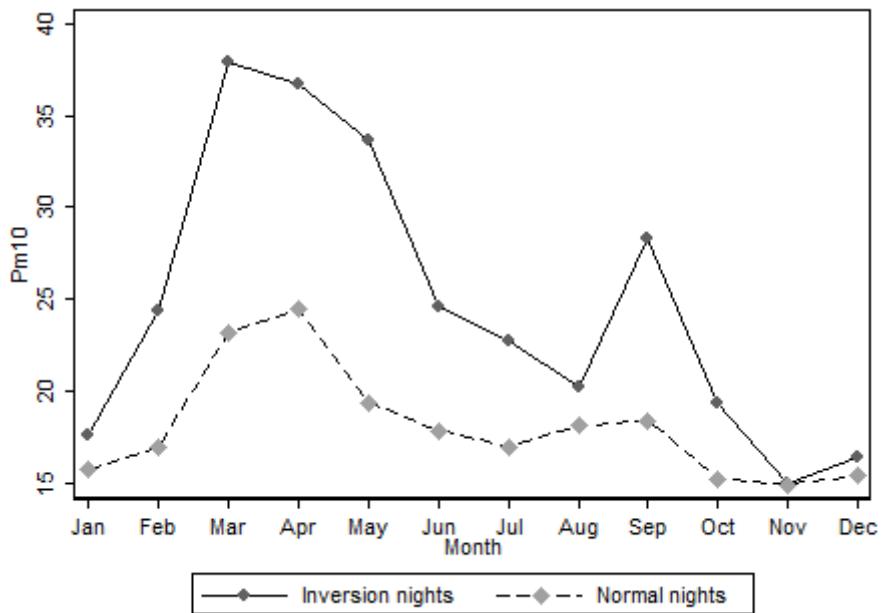
WHO (2006). *Air Quality Guidelines: Global Update 2005: Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide*. World Health Organization.



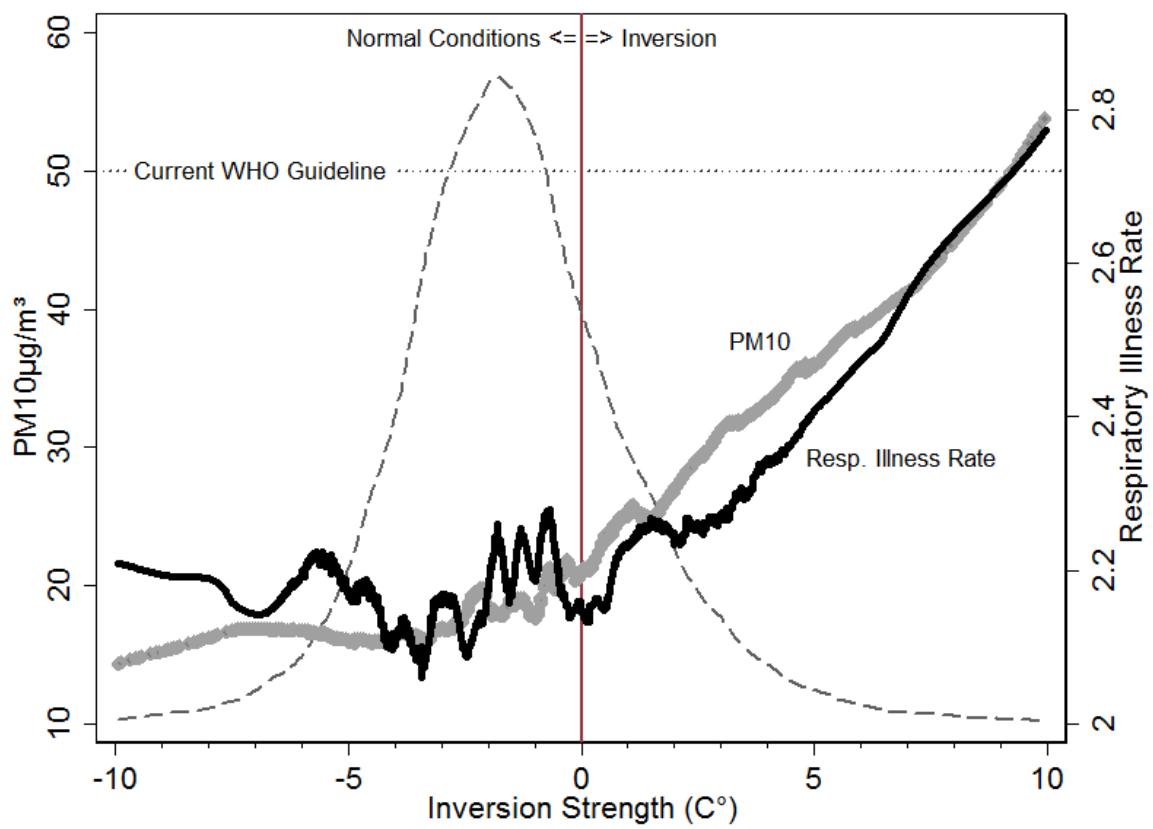
**Figure 1:** Municipalities, Temperature Grid Centroids ( $\diamond$ ) and Air Quality Monitors ( $\Delta$ ).



**Figure 2:** Comparison of 24h-mean PM<sub>10</sub> levels during normal and inversion episodes

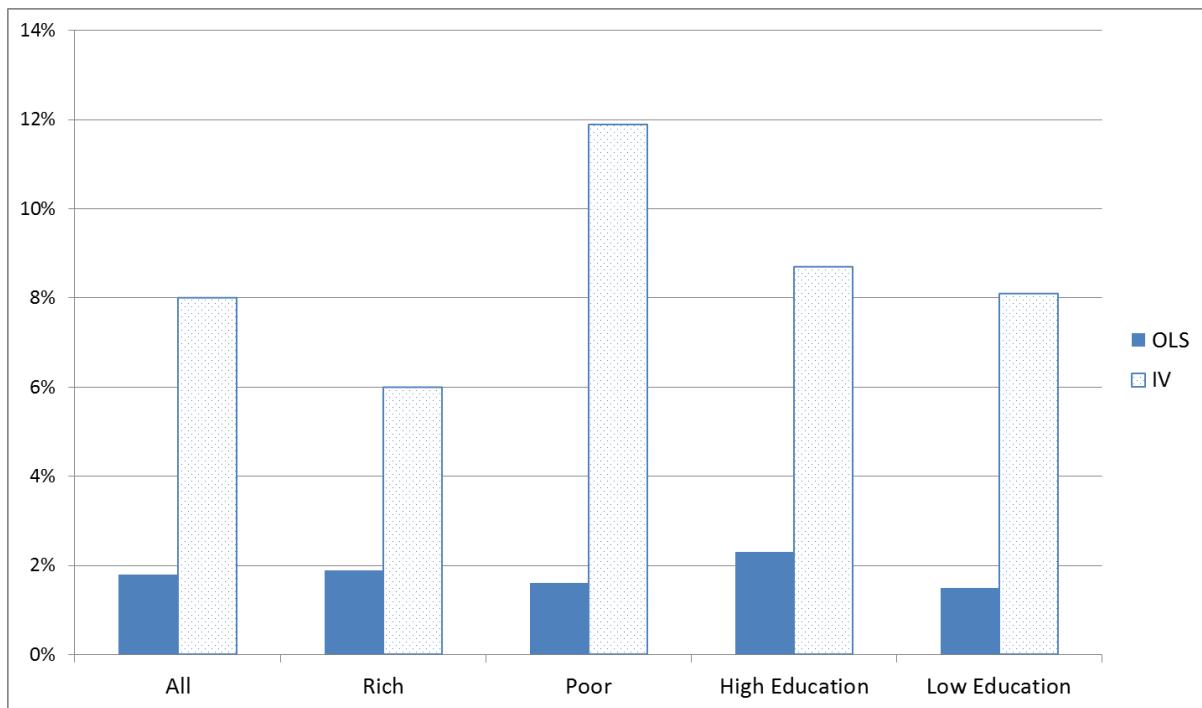


**Figure 3:** Seasonal comparison of 24h-mean PM<sub>10</sub> levels during normal and inversion episodes



**Figure 4:** Non-parametric Reduced Form and First-stage Estimates

Generalized additive model estimates of respiratory health care visits per 10,000 children (black) and PM10 level (gray) on inversion strength using a narrow bandwidth local linear smoother, controlling for calendar month and an extensive set of weather variables (see the text for details). Kernel density distribution in the background (dashed), current WHO 24-h PM10  $\mu\text{g}/\text{m}^3$  guideline for reference (dotted).



**Figure 5:** OLS and IV Estimates by Socio-Economic Status (10  $\mu\text{g}/\text{m}^3$  PM10 increase)

**Table 1:** Summary statistics for Health Care Visits,  $PM_{10}$ , and covariates

	Mean	Standard deviation
<b>A. Dependent variables</b>		
<i>Rate of daily respiratory related hospital visits per 10,000 children</i>		
Any respiratory illness	2.16	2.55
Age 0-5	4.45	5.77
Age 6-10	1.47	3.40
Age 13-18	1.03	2.19
Asthma	1.68	2.55
Pneumonia	0.56	1.52
Bronchitis	0.78	1.88
Other respiratory illness	1.82	2.51
External causes	2.44	2.33
<b>B. Independent variables</b>		
PM <sub>10</sub> ( $\mu\text{m}/\text{m}^3$ )	20.77	17.57
Temperature (Kelvin)	277.57	9.31
<i>Daytime (ground level)</i>		
Temperature (Kelvin)	275.04	7.99
<i>Nighttime (ground level)</i>		
Precipitation (mm) (N=33,079)	0.51	1.14
Windspeed (m/s) (N=33,079)	3.26	1.75
Daily Cloud cover ratio	0.43	0.28
Nightly Cloud cover ratio	0.45	0.27
Share of Inversion days	0.25	0.43
Inversion strength	-1.35	2.80
Number of Observations	34 175	

**Table 2:** Descriptive statistics for key variables conditional on inversion status

Normal Nights N=25,567	Mean	Standard Deviation .
Rate of health care visits per 10,000 children:		
<i>Any respiratory illness</i>	2.14	2.51
PM <sub>10</sub> ( $\mu\text{m}/\text{m}^3$ )	18.06	13.09
Temperature (Kelvin)	277.80	8.90
<i>Daytime (ground level)</i>		
Temperature (Kelvin)	276.06	7.48
<i>Nighttime (ground level)</i>		
Precipitation (mm) (N=24,862)	0.63	1.24
Wind speed (m/s) (N=24,862)	3.48	1.79
Daily Cloud cover ratio	0.47	0.27
Nightly Cloud cover ratio	0.49	0.26
Inversion nights N=8,608		
Rate of health care visits per 10,000 children:		
<i>Any respiratory illness</i>	2.24	2.68
PM <sub>10</sub> ( $\mu\text{m}/\text{m}^3$ )	28.82	25.08
Temperature (Kelvin)	276.87	10.40
<i>Daytime (ground level)</i>		
Temperature (Kelvin)	272.01	8.68
<i>Nighttime (ground level)</i>		
Precipitation (mm) (N=8,217)	0.15	0.59
Windspeed (m/s) (N=8,217)	2.61	1.46
Daily Cloud cover ratio	0.33	0.27
Nightly Cloud cover ratio	0.32	0.26

**Table 3: Baseline Estimates**

	<i>First-stage</i>	<i>Reduced form</i>	<i>OLS</i>	<i>IV</i>
	Pm <sub>10</sub>	Respiratory related health care visits per 10,000 children	Respiratory related health care visits per 10,000 children	Respiratory related health care visits per 10,000 children
	(1)	(2)	(3)	(4)
Inversion	6.154*** (0.892)	0.106*** (0.0275)		
PM10			0.00398*** (0.000908)	0.0172*** (0.00253)
Number of clusters	90	90	90	90
Mean of dependent variable	20.78	2.164	2.164	2.164
Observations	34 156	34 156	34,156	34,156
Weather controls (Nasa+SMHI):	Yes	Yes	Yes	Yes
First-stage F-stat.				47.99

**Notes:** Robust standard errors in parentheses clustered at the municipality level, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. This table presents the first-stage and reduced-form estimates of the effect of pollution on respiratory related hospital visits for children aged below 18 in Sweden. The rate of hospital visits is calculated by dividing the number of hospital visits each day in each municipality by the total number of children in the municipality and then multiplying by 10,000. Thus, the coefficient can be interpreted as the number of hospital visits per day and 10,000 children aged below 18. This model includes year-month fixed effects, municipality fixed effects and day-of-week fixed effects. Temperature controls (polynomials of degree up to four) of the daily and nightly temperatures, cloud coverage, cloud coverage<sup>2</sup>, humidity, humidity<sup>2</sup>, H2O levels, precipitation, precipitation<sup>2</sup>, wind speed, wind speed<sup>2</sup> are included in all regressions. Each regression is weighted by the total number of children in the municipality. The percent effect is (PM10 coefficient\*10)/(mean of dependent variable) in the last two columns.

**Table 4:** OLS and IV Regression results for the effect of PM<sub>10</sub> on respiratory illnesses

Dependent variable: <i>Specification:</i>	Respiratory related health care admissions per 10,000 children				
	Baseline specification: + age	Baseline specification: + age + maternal education	Full Baseline specification: + age + maternal education	Full Baseline specification: within 2km of pollution monitor	Full Baseline specification: without weather station data
	(1)	(2)	(3)	(4)	(5)
<b>OLS estimates</b>					
PM <sub>10</sub>	0.00398*** (0.000908)	0.00398*** (0.000914)	0.00394*** (0.000926)	0.00262*** (0.000744)	0.00401*** (0.000901)
<b>IV estimates</b>					
PM <sub>10</sub>	0.0172*** (0.00253)	0.0173*** (0.00252)	0.0174*** (0.00251)	0.0166*** (0.00421)	0.0171*** (0.00225)
Observations	34,156	34,156	34,156	34,156	34,156
Other Controls	No	Yes	Yes	Yes	Yes
SMHI weather	Yes	Yes	Yes	Yes	No
NASA weather	Yes	Yes	Yes	Yes	Yes
Mean outcome	2.164	2.164	2.164	2.215	2.164
F-stat 1 <sup>st</sup> Stage	47.67	47.67	47.99	49.88	44.80
# of cluster	90	90	90	90	90
Effect OLS	1.8%	1.8%	1.8%	1.2%	1.9%
Effect IV	7.9%	8.0%	8.0%	7.5%	7.9%

See notes to Table 1. Column (4) represents children living within 2km of the nearest pollution monitor. The percent effect is (PM10 coefficient\*10)/(mean of dependent variable)

**Table 5:** Effects by child age and by sub-diagnosis.

Dependent Variable:	Respiratory related health care admissions per 10,000 children						
	(1) Full baseline specification	(2) Asthma	(3) Bronchitis	(4) Pneumonia	(5) Other respiratory	(6) Pre-School Kids	(7) School-age Kids
<u>OLS estimates</u>	0.00394*** (0.000926)	0.00283*** (0.000601)	0.000385** (0.000192)	0.000052 (0.000087)	0.000670*** (0.000221)	0.00422** (0.00163)	0.00343*** (0.0007)
<u>PM<sub>10</sub></u>							
<u>IV estimates</u>	0.0174*** (0.00251)	0.00956*** (0.00181)	0.00190* (0.000986)	0.00108** (0.000483)	0.00485*** (0.00155)	0.0280*** (0.00618)	0.0123*** (0.00193)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,156	34,156	34,156	34,156	34,156	34,156	34,156
F-stat first-stage IV	47.67	47.99	47.99	47.99	47.99	46.92	48.63
Number of clusters	90	90	90	90	90	90	90
Mean respiratory illness	2.164	0.890	0.188	0.115	0.971	4.445	1.184
Percent effect (OLS)	1.8%	3.2%	2.0%	0.5%	0.7%	0.9%	2.9%
Percent effect (IV)	8.0%	10.7%	10.1%	9.4%	5.0%	6.3%	10.4%

Notes: The percent effect is (PM10 coefficient\*10)/(mean of the dependent variable)

**Table 6:** Examining the Effects across Socio-Economic Groups.

Dependent Variable:	Respiratory related health care admissions per 10,000 children				
	(1) Full baseline specification	(2) Mother's Education: > High School	(3) Mother's Education: $\leq$ High school	(4) Family Income: > Median	(5) Family Income: $\leq$ Median
<u>OLS estimates</u>	0.00394*** (0.000926)	0.00497*** (0.00113)	0.00333*** (0.000897)	0.00392*** (0.000833)	0.00422*** (0.00117)
PM <sub>10</sub>					
<u>IV estimates</u>	0.0174*** (0.00251)	0.0190*** (0.00319)	0.0183*** (0.00374)	0.0120*** (0.00441)	0.0307*** (0.00365)
PM <sub>10</sub>					
Weather controls	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Observations	34,156	34,156	34,156	34,156	34,156
F-stat first-stage IV	47.67	45.02	44.27	46.26	42.72
Number of clusters	90	90	90	90	90
Mean respiratory illness	2.164	2.186	2.266	2.011	2.584
Percent effect (OLS)	1.8%	2.3%	1.5%	1.9%	1.6%
Percent effect (IV)	8.0%	8.7%	8.1%	6.0%	11.9%

*Notes:* The percent effect is (PM10 coefficient\*10)/(mean of the dependent variable)

**Table 7:** Examining the Causes of the SES-gap in the Effects: Avoidance Behavior

Dependent variable: <i>Specification:</i>	Injuries due to External Causes per 10,000 children		
	All (1)	High income (2)	Family income Low Income (3)
<b>OLS estimates</b>	0.000317	0.000258	0.000362
PM <sub>10</sub>	(0.000970)	(0.00107)	(0.000961)
<b>IV estimates</b>	0.00448	0.00407	0.00473
PM <sub>10</sub>	(0.00586)	(0.00588)	(0.00700)
Weather controls	Yes	Yes	Yes
Other controls	No	No	No
Observations	34,156	34,156	34,156
F-stat first-stage IV	49.13	50.40	47.44
Number of clusters	90	90	90
<b>Mean resp. illness rate</b>	2.437	2.482	2.393
Effect on outcome variable OLS	0.1%	0.1%	0.2%
Effect on outcome variable IV	1.8%	1.6%	1.8%

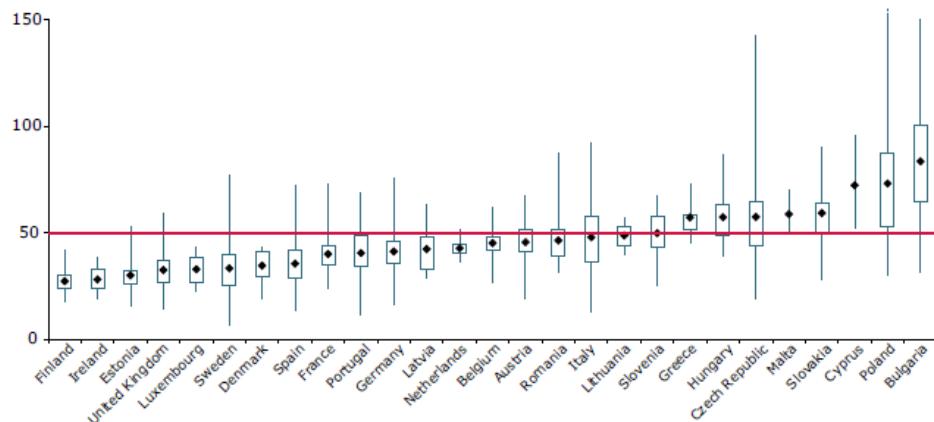
*Notes:* The percent effect is (PM10 coefficient\*10)/(mean of the dependent variable)

**Table 8:** Results for subsamples based on initial health index and family income.

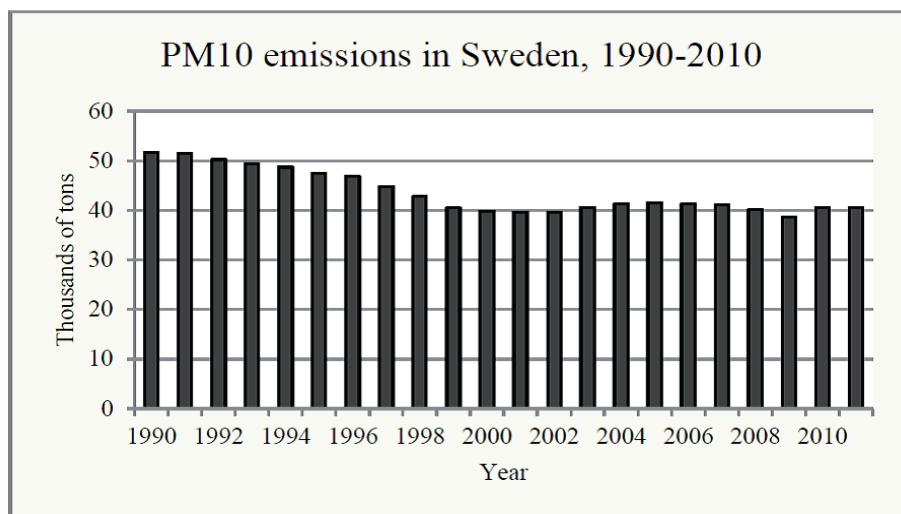
Dependent Variable:	Respiratory related health care admissions per 10,000 children					
	Health Index		Good Initial Health (High index)		Bad Initial Health (Low index)	
	High	Low	High income families	Low income families	High income families	Low income families
	(1)	(2)	(3)	(4)	(5)	(6)
<u>OLS</u> Pm10	0.00432*** (0.00107)	0.00427*** (0.00112)	0.00441*** (0.00109)	0.00424*** (0.00123)	0.00393*** (0.000895)	0.00478*** (0.00161)
<u>IV</u> Pm10	0.0183*** (0.00362)	0.0257*** (0.00437)	0.00851** (0.00396)	0.0337*** (0.00744)	0.0201** (0.00867)	0.0330*** (0.00551)
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,231	33,231	33,231	33,231	33,231	33,231
Number of clusters	88	88	88	88	88	88
Mean resp. illness rate	2.178	2.490	1.920	2.539	2.161	2.885
F-stat 1st Stage	43.93	44.57	44.55	42.87	45.33	43.34
Effect on outcome OLS	2.0%	1.7%	2.3%	1.7%	1.8%	1.7%
Effect on outcome IV	8.4%	10.3%	4.4%	13.3%	9.3%	11.4%

Notes: The health index is based on the first principal component using birth weight and gestational weeks at birth. The children are split into high and low initial health based on whether they are above or below the median of the index. See Table 4 for full specification details. The percent effect is (PM10 coefficient\*10)/(mean of the dependent variable)

## Appendix A



**Figure A1:** The graph illustrates the range of concentrations from all types of stations in  $\mu\text{m}/\text{m}^3$  officially reported by the EU member states and how the concentrations relate to the limit value set by EU legislation (marked by the red line). *Source:* EEA (2012).



**Figure A2:** PM<sub>10</sub> emissions in Sweden between 1990 and 2010, measured in thousands of ton. *Source:* SEPA.

**Table A1:** Additional Specification Checks

Dependent variable:	Respiratory Illness Rate at day $t$			PM <sub>10</sub>	Inversion
Specification:	(1)	(2)	(3)	(4)	(5)
PM <sub>10</sub> ( $t$ ) ( <b>baseline</b> )	0.017*** (0.003)				
PM <sub>10</sub> (avg. past 5 days)		0.020*** (0.003)			
Cumulative PM <sub>10</sub> Effect ( $\sum_{k=0}^5 \beta_{t-k}$ )			0.019*** (.007)		
Weekend Dummy				-4.754*** (0.993)	0.0042 (0.003)
Weather controls	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes
Observations	34,156	34,156	9,591	34,156	34,156
Number of clusters	90	90	88	90	90
Mean outcome var.	2.2	2.2	2.1	20.8	0.25

*Notes:* All reported estimates are IV estimates, except last two columns. Robust standard errors in parentheses clustered at the municipality level, \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Column (1) repeats fully specified model IV model results (see Table 4). Column (2) use the share of inversions nights between t to t-5 to instrument average PM10 levels during the same period. Column (3) presents the cumulative effect from a instrumented distributed lag model from t to t-5. Column (4) show the drop in PM10 levels on weekends, while column (5) show that the sharp drop in pollution levels on weekends do not influence inversions.