Bringing About Pro-Environment Behavior Through Policies and Social Norms

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Ashwini Sankar

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Jay S. Coggins, Paul Glewwe

July 2019

Acknowledgement

I have been lucky enough to come across some fantastic teachers, who I consider as towers of intellect. In that, I would like to thank my advisors, Profs. Jay S. Coggins and Paul Glewwe, without whom working on my dissertation would not have been this joyful. Jay, I can never thank you enough for the number of hours you have spent with me refining my thesis, and all the laughter. Paul, I will forever be in debt for your kindness and interest in all aspects of my career. I would like to thank Prof. Marc F. Bellemare, for providing such thoughtful comments on so many of my drafts. I would also like to thank Prof. Andrew L. Goodkind for being so generous with his time and knowledge, and Prof. Kristen Nelson for the new perspective she brought to my research. I am ever grateful to Prof. Elizabeth Davis for believing in me from the very beginning and being a fantastic boss. Thank you, Prof. Joe Ritter, for going above and beyond and teaching me so many little details that made the biggest difference. I would like to express my sincerest gratitude to Profs. Steve Miller, Steve Polasky and Frances Homans for editing my papers and showing enormous interest in my career.

I wouldn't be here without my family. I am forever grateful for my mother and father, Usha Sankar and G.V. Sankar, who are solely responsible for my sense of wonder and curiosity. I am so grateful for my husband, Balaji Krishnan, because he taught me how to keep on working with a smile. I am grateful to God for giving me my daughter, Anika, who provides the light in my life. I am so grateful for my time with my late grandmother, Narmada Arunachalam, whose enthusiasm and impossibly high standards are the reasons why I am here.

I was lucky enough to find a superb set of friends to help me move forward in the program. Specifically, I would like to thank Bixuan Sun, Lindsey Novak, Vinicius and Claudia Vidigal, Joel Cuffey and Shuoli Zhao for the many hours of problem solving as well as much needed levity.

Lastly, I would also like to thank the staff in the Department of Applied Economics, Jenna Mead, Elaine Reber, Christina Olson, Linda Eells and Karen Routhe, whose industry and genuine concern for students keeps the program going without a hitch.

Dedicated to my mother

Abstract

As human population and consumption have risen, waste generation and air pollution have also increased leading to steady environmental degradation. To stem overflowing landfills and combat air pollution, policies and social norms have often been used as tools to bring about change. My dissertation analyzes the impact of these tools to achieve pro-environment behavior. The first chapter tries to understand two such behaviors, i.e., how to increase recycling and reduce waste. They are key to protecting natural resources, but households probably do not derive any benefit from recycling other than social approval. Based on a theoretical model I built for households, I show that when the social norm of recycling increases, the recycling rate of the household rises and waste per capita falls. My paper is one of the first to test these propositions empirically for Minnesota data using an instrumental variable setup. I show that while waste per capita declines significantly with an increase in social capital, recycling rate does not seem to be influenced by social capital. My second chapter studies the impact of environmental regulations in India on mortality that includes all causes and all ages (or mortality). We know that chronic exposure to air pollution is more harmful to adults than babies and hence focus on mortality as the outcome, for the first time for India. Using a difference-in-differences framework, in the first part of the paper, I show that environmental regulations in India have led to a significant drop in mortality. The second part analyzes the effect of different pollution types on mortality, where I show that $PM_{2.5}$ exposure is more harmful to mortality (but not infant mortality) than TSP. This further strengthens the claim that policies should focus on adults and shift its focus from TSP to $PM_{2.5}$ to get greater gains in health. The last chapter studies the functional form of the relationship between $PM_{2.5}$ concentrations and mortality for the first time for India. The shape of this concentration-response curve will determine if the air in India affects public health at a different or the same rate as the U.S. baseline rate. My paper is one of the first studies to analyze this relationship using panel data for India, without simply extrapolating coefficients from U.S. or European data, following a rigorous identification strategy. I then arrive at the relative risk of mortality estimates at higher pollution concentrations as well as the estimated lives saved due to the reduction in pollution exposure.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Reducing ambient air pollution and reducing the flow of waste into landfills restores the Earth's ecosystem, yet individuals do not find it profitable to act on their own to improve the environment. Legislation and policies, therefore, have been used to reduce air pollution and to induce people into recycling household waste; but to enable pro-environment behavior, sometimes other tools such as nudges, and social norms can help where policies cannot. My dissertation helps inform us of the different scales at which formal and informal institutions can operate to improve the environment and increase human well-being.

As human population and consumption per capita have risen, environmental degradation has steadily increased too due to increased waste generation. Recycling can be key to slowing this degradation and protecting natural resources. The U.S. is one of the top waste producing countries of the world, but only 35 percent of its waste is being recycled or composted, as per the 2014 U.S. EPA Fact Sheet. When policy is inadequate, informal institutions such as social norms can reinforce pro-environment behavior such as recycling and waste reduction to contain overflowing landfills, a major global environmental concern. My first chapter studies how social capital can be used to increase recycling behavior and reduce the generation of waste in households.

Pro-environment behavior such as recycling is expensive and requires active participation from individuals in the household, while the returns to the person involved are difficult to measure and likely to be small. Some studies have explained this behavior as a desire to be altruistic (a pure motive) while others describe it as an impure motive, namely the wish to fit in and follow the social norm.

In this chapter, I develop a theoretical framework where the household decides to recycle and produce waste based on both monetary as well as non-monetary incentives. In that, the model

examines the household's concern for the environment (a purely altruistic motive), the "warm-glow" from doing good—i.e., recycling—and the benefit from following a norm (a more impure motive). Following Azar (2004), I claim that we need both these motives in the utility function in order to understand recycling and waste behavior. The optimization of this problem yields two propositions: The first one proves that when the social norm increases, recycling rate increases The second one proves that when the social capital norm increases, waste per capita declines.

The two propositions from the theoretical model become the basis for the empirical framework. My aim in the empirical section is to use county-level data from Minnesota from 1996 to 2013 to show that these propositions hold in reality. Using an instrumental variables regression framework, I show that when social capital increases, both recycling per capita and waste per capita decline significantly. However, the analysis shows that the social capital has no significant impact on recycling rate. These results hold even when there are deviations from the strict exogeneity assumption.

These results show that social capital can be used as a tool to bring about change in recycling and waste behavior. As the authors in Nyborg et al. (2016) state, policies can act as agents that bring shifts in social norms that can ultimately bring about change in behavior such as recycling.

The second chapter of my dissertation is influenced by the most recent Global Burden of Disease Study (IHME, 2016) which estimated that one million deaths are attributable to ambient air pollution in 2016 in India. This impact is an underestimate since it does not account for the significant threat that air pollution poses to general human well-being and to ecosystems (UNEP, 2002). In this chapter, I investigate whether environmental regulations in India have been successful in reducing the overall mortality rate from all causes and for all ages. I also go further by comparing the impact of different pollution types on mortality to make policy recommendations for future Indian regulations.

In a departure from the traditional emphasis (in the economics literature) on infants, my second chapter estimates the effectiveness of India's environmental regulations on mortality. One of my three main contributions in this chapter is in addressing a gap in the literature by being one of the first studies to estimate this effect for India. While we know that policies have led to a significant reduction in total suspended particles (TSP) levels (Greenstone and Hanna, 2014), I am able to show in this chapter that these regulations were also effective in reducing mortality. But not all pollutants are created equal. Upon comparing the effects of TSP and fine particulate matter that is less than 2.5 microns in size ($PM_{2.5}$), using recently released $PM_{2.5}$ data, we found that $PM_{2.5}$ is more dangerous to human health than TSP. This result, my second major contribution, emphasizes the importance of the need for a shift in policy focus towards targeting $PM_{2.5}$ in India, rather than TSP. My third contribution is the new measure of mortality that I built using two sources of data within the Census. It is well known that the Crude Death Rate (CDR) that is recorded by the Census

in India is not complete, and hence it underestimates the number of deaths across India. Combining the CDR from the Census and estimated registration rate from the Sample Registration System (SRS), that measures the completeness of data, I created an adjusted mortality rate (AMR) that adjusts for the measurement error that exists in these data and also includes deaths from all causes and for all ages.

The previous economics literature has primarily focused on the impact of environmental policies on infant mortality because it is easier to calculate cumulative years of pollution exposure for infants, making it a more suitable endpoint and estimate the causal effects of measures designed to combat pollution. However, we know from earlier reports that the impact of ambient air pollution on adults is a few hundred times more than its effect on infant mortality (U.S. EPA, 2012), making mortality for all ages crucial for public health and policy.

To test the key hypotheses in this chapter, I compiled data for a wide-ranging set of variables that include mortality rates and pollution levels for 119 districts (out of a total of 466 in India, as per the 1991 Census of India) from 1987 to 2007. The Supreme Court of India introduced two major air pollution policies, the Supreme Court Action Plan (SCAP) and Catalytic Converter (CC) policies, in a phased manner across different cities over a period that enabled an experiment setup. I use the staggered rollout and a difference-in-differences model to estimate the differential impact of Indian pollution regulations in districts that were affected by policy in contrast to those that were never under any pollution regulation. As part of my first contribution, I show that the areas that were under regulation experienced a 19% reduction in the AMR.

The aim of pollution regulation policy in India, much like the rest of the world, is to keep air pollution from exceeding certain standards so as to curb pollution-related illnesses and deaths. My second main contribution in this chapter is to address that objective directly by investigating the impact of several types of pollutants on human health. While the policies in India aim at reducing coarse particulate matter less than 100 microns in size (TSP levels) or other pollutants such as sulfur dioxide (SO_2), a toxic gas released mainly from industrial activity and vehicular emissions, and nitrogen oxides (NO_x), gases produced from combustion of fuels, typically from motor vehicles, no policies to date consider the dangers of fine particulate matter that is less than 2.5 microns in size ($PM_{2.5}$). Even though $PM_{2.5}$ can penetrate through the lungs with severe health impacts that are well documented in the context of the U.S. and Europe (U.S. EPA, 2009), few existing studies have estimated the effect for India, where the stakes are so high.

In addressing this gap, I examine the effect of $PM_{2.5}$ (in comparison with other pollutants) on the AMR. I show that while a 10 percent reduction in $PM_{2.5}$ concentrations leads to a 2.5% decline in the AMR, the effect of TSP levels and other pollutants on mortality is negligible, which indicates that $PM_{2.5}$ is more dangerous than the coarser TSP. From this recommendation, I make

the recommendation that targeting $PM_{2.5}$ as part of policy would result in greater gains in health.

My third and last chapter builds on some of the results in the second chapter to embark on a discussion of the burden of disease attributable to pollution in India. It estimates the shape of the concentration-response curve, a powerful policy tool, and calibrates the disease attributed mortality rate for the first time in India using actual mortality rate data. The estimated curve can accurately measure the number of avoided premature deaths due to a reduction in pollution.

Previous studies have established the existence of a causal relationship between pollution and premature mortality, predominantly for the U.S. or Europe where $PM_{2.5}$ is less than 30 $\mu g/m^3$. Very few papers have tried to understand this relationship for higher pollution levels. The 2010 Global Burden of Disease (GBD) risk assessment study was the first to estimate the relative risk of mortality for a broader range of pollution levels. The study combined many of the models from various U.S. and Europe-based studies that use different pollution sources (namely, active smoking, household cooking fuel, and others) to predict the relative risk of mortality for higher pollution levels.

My paper uses a new panel dataset with $PM_{2.5}$ levels and mortality rates for India spanning 119 districts from 1998-2015, instead of an amalgam of model coefficients that previous research used to arrive at GBD-like relative risk estimates of the burden of disease at each pollution concentration level.

The estimated shape of this concentration-response function shows that as baseline pollution levels rise, mortality increases are smaller for a 10 $\mu g/m^3$ increase in concentration levels, which is counterintuitive to the prevailing idea that a 10 $\mu g/m^3$ increase in pollution levels in dirtier places causes more damage to human health than in cleaner places. Therefore, to achieve the benefits from a given number of avoided deaths, pollution in India should decline substantially as compared to places with lower baseline pollution levels. I use this estimated shape to calculate the extent of ambient pollution reductions that would be required to decrease premature mortality to specific standards. My model assessment as of 2015 is that an aggressive $PM_{2.5}$ target of 35 $\mu g/m^3$ from the 2015 country average (for the districts in the data) of 55 $\mu g/m^3$ would lead to a reduction of 134,489 deaths annually using the preferred log-log specification of the concentration response curve. This result is well within the 95% confidence interval estimates of relative risk of mortality in the recently developed Global Exposure Mortality Model.

Chapter 2

Changing Recycling Behavior through Social Capital: Evidence from Minnesota

Abstract

As human population and consumption per capita have risen, environmental degradation has steadily increased due to many causes, including increased waste generation. Recycling can slow this degradation and protect natural resources. The economics literature show that public goods provision, such as recycling, can be influenced by social norms. In this paper, I build a theoretical model where a household optimizes the amount it recycles given that its preferences are based on monetary as well as non-monetary motives. Since pro-environment behavior can be expensive, especially in terms of time, and has very little return to the household, non-monetary factors such as peer approval and social norms can play a large role in the household's decision to recycle. Based on this model, I show that when the social norm of recycling increases, the recycling rate of the household rises and waste per capita falls. This provides an empirical framework to test whether social capital leads to higher recycling rate and lower waste per capita in the counties of Minnesota. Using instrumental variable regression, and panel data from 87 counties of Minnesota spanning 18 years, I show that while waste per capita declines significantly with an increase in social capital, the recycling rate does not seem to be influenced by social capital. These results are robust to movements from the strict exogeneity assumption.

Keywords: social norm, social capital, recycling, waste, environmental degradation

2.1 Introduction

Global ecosystems have been changing at an accelerated pace in the last 50 years, all in response to the growing demand across the world for food, fresh water, timber, fiber, and fuel (Assessment, 2005). The Assessment (2005) report says that there is considerable degradation of ecosystem services due to waste that ends up in landfills. With this current state of consumption and the resulting waste generation, the act of recycling and recovering materials in order to re-use them in a sustainable manner is key to minimizing the impact of human activities on environmental resources and slowing climate change (U.S. EPA, 2016).

There has been some improvement in recycling rates over time in the United States. The U.S. EPA reports that the United States recycles 35% of its municipal solid waste in 2015, an increase from a mere 7% in 1960. More locally, Minnesota recycles its waste at a rate of 41% in 2013. The need for more recycling, however, is apparent in the many initiatives and policies that are proposed to increase recycling activities even more in the state. By law, counties in Minnesota are required to provide their residents opportunities to recycle and access to recycling facilities. The 2014 legislative recycling goals for 2030 are to increase recycling to 75% in the metropolitan counties [1] of Minnesota and keep the same goal of 35% for rest of Minnesota. Clearly, recycling is a priority for local and national governments.

Reducing, reusing and recycling change the "lifecycle" of waste and so could ultimately help protect the environment (Office of the Legislative Auditor, 2015). Unlike some economic activities, however, recycling is often expensive, especially in terms of time invested by households, with no clear return to the people engaged in it other than the larger, common good of a cleaner environment (Abbott et al., 2013). There is also no clear monetary incentive or sanction to recycling that would ensure that all trash is separated such that recyclable materials do not directly enter the landfill. Since the act of recycling may not lead to an increase in direct utility for people, the decision to recycle might be influenced by the "warm glow of doing good" (Andreoni, 1990). Andreoni (1990) provides a model where public goods provision is dependent not only on pure altruism but also on "impurely altruistic" motives, namely peer approval and social norms.

Such recycling behavior is similar to tipping behaviors and other such provision of public goods, where the behavior may not be consistent with "selfish consumers" (Azar, 2004). Even if the social norm of recycling may be costly to comply with and there are no clear returns (in the present or the future) from recycling, people do recycle waste. If there is no clear economic incentive, then why do people recycle? In a similar dilemma with tipping, Bodvarsson and Gibson (1999) show that

[1] The term "metropolitan counties" refers to the seven Twin Cities counties of Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington

people tip because they would like to conform to the social norm of tipping.

Following the model of tipping by Azar (2004), my theoretical model of recycling behavior specifies that the decision to recycle and generate waste is affected by both economic incentives and the desire to follow the social norm of recycling. As in the case of Abbott et al. (2013), I introduce a constraint in terms of the time used for recycling in the model. I focus on two phenomena in this model: (i) a change in the recycling rate due to a change in social norms; and (ii) the change in municipal solid waste (MSW) per capita due to a change in social norms. The theory shows that recycling rates rise with an increase in social norms, and that MSW falls with an increase in social norms.

The two propositions from the theoretical model become the foundation for the empirical framework. My paper is one of the first studies to use aggregate county-level data for Minnesota spanning 18 years to understand whether social norms do indeed make an impact on recycling and waste generation behavior. Using an instrumental variables estimation, I show that MSW falls significantly by 184 tons per 1000 persons, following a one unit increase in social capital, and that this result is consistent with the theoretical result, whereas social capital does not have a significant impact on recycling rate as a percentage of MSW. In addition to these two main results, I also show that the impact of social capital leads to a significant decline in recycling per capita. These set of results leads me to believe that when there is an increase in social capital, it drives the household to waste less per person, which in turn leads lower materials to recycle per capita. These two foregoing results might be the reason that recycling rate as a percentage of MSW does not change significantly due to a change in social capital.

Social norms are important if they cans bring about pro-environment behavior such as recycling. As such, changes in expected social reactions and social sanctions could potentially shift social norms (Nyborg et al., 2016), which could in turn shift recycling behavior. Policy changes and other government actions could bring about this shift in social norms, which could increase recycling activity and thereby make this an important part of the solution to this environmental problem. Given the results from this paper, a policy that could shift the social capital could ultimately act as a way to change the way households produce waste.

The rest of this chapter proceeds as follows: Section 2.2 presents the theoretical foundation of how consumers behave with regards to recycling and waste generation. The propositions at the end of this section form the basis of the empirical section that follows. Section 2.3 describes all the data that have been used, and their various sources. Sections 2.4 explains the empirical framework and the identification strategy, after which the results are discussed in Section 2.5. Section 2.6 concludes.

2.2 A Model of Recycling Behavior

An economy with a representative household in county i recycles at the rate of R_{it} at time t, which is the amount recycled as a percentage of the total amounted wasted or MSW, i.e., X_{it}. Following Azar (2004), the utility function is additively separable with respect to both its traditional and its social norm components. Specifically, the utility of the household is given by:

$$U(R_{it}, X_{it}; R_t^*, \theta) = B(Q_{it}) + \nu(R_{it} - R_t^*) - \theta D(X_{it} * (1 - R_{it})) \tag{2.1}$$

Households get utility, $B(Q_{it})$, from consumption Q_{it}. ν is the utility that the household derives from following the norm (R_t^*), D is the disutility that the household gets from contributing waste to the landfill, and θ measures the extent of feeling that households have for recycling and against wasting.

B is the utility derived from consumption that increases as consumption increases, but increases at a decreasing rate as consumption rises. This can be summarized as the following assumption:

Assumption 1. $B(Q_{it})$ is twice differentiable and concave, i.e., $B' \geq 0$ and $B'' \leq 0$.

For a given price and wage, a household can choose to spend its time in working or recycling (Abbott et al., 2013). Hence, more time spent on recycling would mean that the household will spend less time working, which would ultimately lead to fewer resources that can be spent on consumption of goods. Following Abbott et al., 2013, the budget constraint can be stated as:

$$Q_{it} = w(1 - t_R - \bar{l}) \tag{2.2}$$

where Q_{it} is the composite consumption good whose price is normalized to 1, w is the wage, t_R is the time taken to recycle, and $\bar{l}$ is a constant time that is devoted to leisure. Hence, there is a constant amount of time net of leisure that a household has to devote to either work or recycling activities. I assume, in this paper, that t_R is a linear function of the recycling rate, R_{it}. In other words, $t_R = t \times R_{it}$, where $t \geq 0$. Substituting this functional form into Equation (2.2) yields:

$$Q_{it} = w(1 - tR_{it} - \bar{l}) \tag{2.3}$$

Consumption of goods, however, ultimately turns into garbage, X_{it}. We can think of garbage as a portion of consumption, i.e., $X_{it} = \gamma Q_{it}$ where $0 \leq \gamma \geq 1$. So, if Equation (2.3) can be expressed in terms of X instead of Q:

$$X_{it} = \gamma w(1 - tR_{it} - \bar{l}) \tag{2.4}$$

Since time spent on recycling rises as recycling amount per waste generated (R_{it}) increases, we can say that quantity consumed will also decline, also leading to less garbage generated. In other words, Equation (2.4) shows that as the recycling rate increases, garbage declines.

Eventually, X_{it} adds to the total municipal solid waste that all the households together generate, which in turn contributes to the landfill. Although households enjoy consumption, they would not like to see waste accumulate as a result of their consumption. R_{it} is recycling as a percentage of the household's solid waste (X_{it}). For a given amount of household garbage, X_{it}, households can decide to recycle more or less, in proportion to X_{it}. Since this is the amount of recycling per unit of garbage generated, I assume that R_{it} can be really close to 1 but never be equal to 1 because there will always some garbage that the household will generate. Hence, the amount recycled per unit of waste generated, $R_t < 1$.

It should be noted, however, that R_{it} does not distinguish between two households that generate different levels of garbage (X_{it}), as long as they recycle the same proportion of their waste.

Households derive utility from comparing their recycling behavior to their peers. In other words, they could recycle as a result of peer pressure from a prevailing social norm, R_t^*. $D(X(R_{it}) \times (1 - R_{it}))$ is the disutility that the household derives from seeing waste generated that is not recycled, due to the household. It follows that $D' \geq 0$. The model allows for $\theta \geq 0$ by which at least some households dislike having their waste go into the landfill and get pleasure from recycling their garbage. This does not mean that *all* people care about their waste going into the landfills. It is possible that people do not care about recycling, in which case $\theta = 0$. The households are heterogeneous with respect to θ, which has cumulative distribution $F(\theta)$.

Just as in Azar 2004, the function ν can take on different forms. It can represent the desire to conform to the social norm or even show a tendency to surpass your neighbor's efforts, i.e., a desire to recycle more than the norm (Clark and Oswald, 1998). We assume that the household derives higher utility from recycling close to the norm. Mathematically, $\nu'(x) > 0$ for all $x < 0$ and $\nu'(x) < 0$ for all $x > 0$ (where $x = R_{it} - R_t^*$). I assume that ν is continuously differentiable implying that $\nu'(0) = 0$. Specifically, since the notion of the recycling norm is vague to the household, a small change in the recycling norm does not cause a discrete jump in utility (Azar, 2004). In addition, I also assume that the function ν is concave because $\nu(x)$ increases at a decreasing rate for all x. In other words, as the household gets farther away from the social norm, the increase in utility derived from surpassing their neighbors keeps declining or, the drop in utility is more drastic as the household recycling level keeps falling below the social norm. The following assumption summarizes what is stated above:

Assumption 2. The function ν is continuously differentiable, $\nu'(x) > 0$ for all $x < 0$, and

$\nu'(x) < 0$ for all $x > 0$ (it follows that $\nu'(0) = 0$). ν is twice continuously differentiable, $\nu''(x) \leq 0$ for all x

In this economy, following Conrad and Clark 1987, the representative household in county i also derive disutility, D, from accumulated waste X_{it}. It follows that the function D is convex in X_{it}, i.e., as the household generates more waste, disutility keeps increasing at an increasing rate. As recycling rate, R_{it}, increases, this disutility declines. We assume here that the cost of recycling is negligible compared to what the household wealth. However, as recycling rate keeps increasing the time taken to recycle and also the cost might rise which means that $D'' \geq 0$ or $-D'' \leq 0$.

Assumption 3. D is continuously differentiable, $D'(x) \geq 0$ for all $x > 0$, and $D''(x) \geq 0$ for all $x > 0$

The social norm evolves such that the norm in each period is the average recycling rate of all persons in the economy in the previous period. Given that the optimal recycling rate of household i at time $t-1$ is $R_{t-1}(\theta, R^*_{t-1})$, we can say that the social norm, R^*_t, evolves in the following manner:

$$R^*_t = E_\theta[R_{t-1}(\theta, R^*_{t-1})] = \int R_{t-1}(\theta, R^*_{t-1})]dF(\theta) \tag{2.5}$$

where the integration is over all θ types of households.

To obtain the optimal recycling action of the representative household for county i at time t, it will maximize Equation (2.1) subject to the constraint, Equation (2.4) (dropping the i and t subscripts):

$$\text{Maximize } U(R, X; R^*_t, \theta) = B(X) + \nu(R - R^*_t) - \theta D(X(1 - R))$$
$$\text{Subject to } X = \gamma w(1 - tR - \bar{l})$$

The Lagrangian for this optimization problem is then:

$$\mathcal{L} = B(X) + \nu(R - R^*_t) - \theta D(X(1 - R)) + \mu(X - \gamma w(1 - tR - \bar{l})) \tag{2.6}$$

Assuming inferior solutions, i.e., $X_{it} > 0$ and $0 < R_{it} < 1$ for the representative household, the first-order conditions from Equation (2.6) are

$$\frac{\partial \mathcal{L}}{\partial R} = \nu'(.) + \theta X D'(.) + \mu\gamma w t = 0 \tag{2.7}$$

$$\frac{\partial \mathcal{L}}{\partial X} = B'(.) - \theta(1 - R)D'(.) + \mu = 0 \tag{2.8}$$

The solutions to this problem, using the first order conditions 2.7 and 2.8, can be denoted as $R^0(R_t^*, \theta)$ and $X^0(R_t^*, \theta)$. The parameters, R_t^* and θ are given exogenously, and different values of R_t^* and θ will give us different solutions. Substituting the different values of these solutions into the maximand, U, will give us the following value function:

$$U^0((R_t^*, \theta) = U(R^0(R_t^*, \theta), X^0(R_t^*, \theta), R_t^*, \theta) \tag{2.9}$$

which is the corresponding maximized value of the consumer utility for the optimal solution.

Proposition 1 Consider a representative household who satisfies the assumptions 1 to 3 that were specified above, and with the consumption levels and recycling rates satisfying the conditions in Equations (2.7) and (2.8):

(i) For a given value of recycling social norm, R_t^*, an increase in θ would increase the value of consumer utility.

(ii) For a given θ, an increase in recycling social norm R_t^* would increase or decrease the value of consumer utility depending on where R^0 is with respect to R_t^*.

The proposition $1(i)$ demonstrates that under equilibrium conditions, when there is a shift in the value of θ and households derive more pleasure from recycling, there is higher equilibrium utility. In a similar vein, proposition $1(ii)$ establishes that when there is an exogenous shift in the social norm of recycling, the utility at equilibrium increases or decreases depending on where the household is with respect to the social norm. The proofs of all the formal results are shown in the Appendix.

Proposition 2. Consider a representative household with utility and constraints satisfying all the assumptions stated above. Then, the following hold at equilibrium:
(i) a small increase in R_t^* causes R to increase
(ii) an increase in R_t^* causes X to decrease

For all households with $\theta \geq 0$, $\frac{dR_t}{dR_t^*} \geq 0$ since p'' as well as ν'' are less than 0, and B'' and D'' are positive. This means that every time the recycling norm rises, we also see an increase in recycling rate. For all types of households, the proof of proposition $2(ii)$ tells us that as the social norm of recycling increases, the amount of garbage generated by the household declines.

The intuition behind these 2 results in proposition 2 is that when the social norm of recycling increases, households feel the social pressure to recycle more and waste less in order to gain social approval from others around them. Therefore, the household recycle more and waste less in response to the higher social norm. These 2 results provides the structure to my empirical analysis, in which I explore the plausibly causal relationship using Minnesota county data spanning over 2 decades in

the following sections.

2.3 Data and Sources

The Waste Management Act was enacted in Minnesota in 1980 in order to preserve the natural resources of the state, as well as protect public health. In keeping with this broad goal, one of the objectives of this Act was to separate and recover materials. Once this objective was achieved, the aim was to recycle these materials if they could not be reduced or re-used.

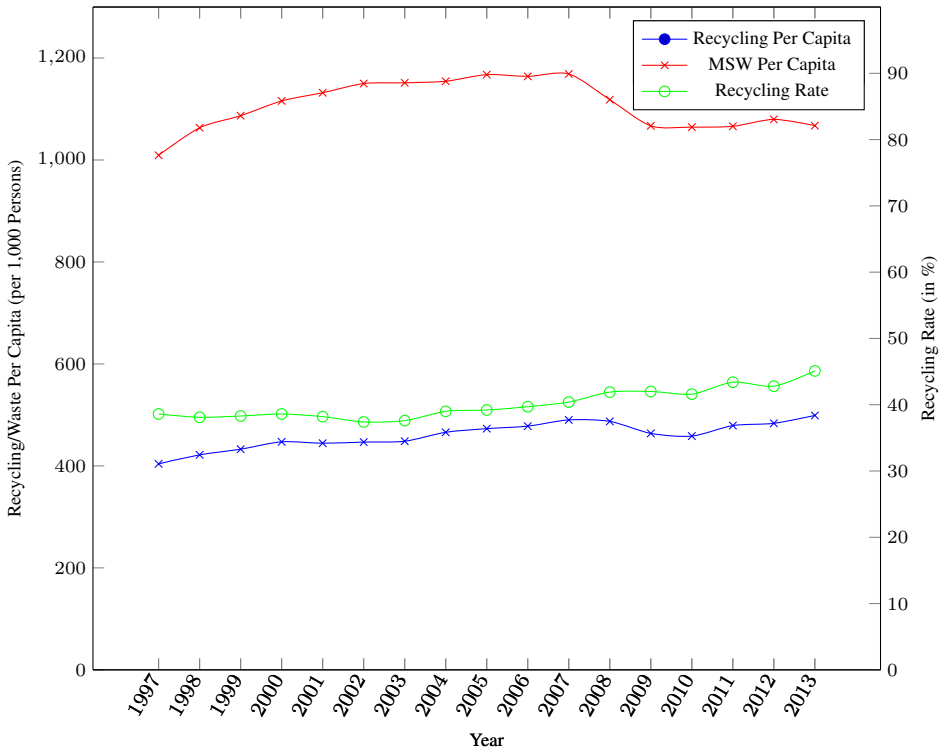
In 1989, the Governor's Select Committee on Recycling and the Environment (SCORE) recommended that the Legislature adopt a set of laws related to waste management, commonly referred to as SCORE. This was also the beginning of state funding for a number of recycling programs among other waste management programs. This initiative also included a collection of data on recycled materials. This endeavor was started by the State in 1989 and then transferred into the counties' hands in 1991. The data collected as part of this initiative were supplied by all 87 counties of Minnesota, and were published as part of the SCORE reports.

One major task of SCORE was to publish data on different types of recycling materials. Recycling materials can be classified into four major material types: glass, metal, paper, and plastics. Within these broad categories, data have been collected on several sub-categories over the years. Glass mainly consisted of glass collected from food and beverage containers. The metal category consists of aluminum, ferrous and other metals. Within paper, there are several reported types namely cardboard, magazines, newspaper, mixed paper, and others. Plastics are classified into many types including film, high-density polyethylene (HDPE) and mixed plastics among others. In addition to these categories, data have also been collected for organics that include food to people and livestock, source separated compost, yard waste and others. Lastly, statistics on problem materials (or banned items) such as antifreeze, electronic devices, major appliances and others have also been consistently collected for the past three decades.

In addition to data on the separated recycled materials, SCORE reports also report the total mixed municipal solid waste (MSW) generated [2] every year for all 87 counties of Minnesota. This includes all waste, including recycled materials, that ends up in landfills.

In order to measure the recycling rate for each county in each year, I first added the tons recycled in all the above categories to get the total amount recycled for each of the 87 counties of Minnesota.

[2] "Mixed municipal solid waste" is defined in Minnesota Statutes 2014, 115A.03, subd. 21(a), as garbage, refuse, and other solid waste from residential, commercial, industrial, and community activities that the generator of the waste aggregates for collection.



Source: Minnesota Pollution Control Agency, U.S. Census Bureau
Note: MSW stands for mixed municipal solid waste.

Figure 2.1: Recycling and Waste Trends

I then calculated the recycling rate as follows:

$$R_{it} = \frac{\text{Total amount recycled in county } i \text{ in year } t}{\text{MSW in county } i \text{ in year } t} \times 100 \tag{2.10}$$

From 1997 to the most recent data in 2013, the total recycled amount increased by 33%. During this time, however, MSW across the counties (from the Minnesota Pollution Control Agency's (MPCA) SCORE report data) of Minnesota have increased by over 20%. The trend in the per person figures is relatively stable. Both the recycled amount and the MSW per 1,000 persons have risen, by 5.0% and 5.8% respectively. Recycled amount per waste generated (or recycling rate) has also been steadily increasing in this time period, increasing by 6 percentage points (see Figure 2.1).

In addition to data on recycled amounts and waste generated, the SCORE report also collects data annually from the counties that relate to all the programs and ordinances that the county has in place for separating waste materials and recycling. Among the several types of questions in this survey, I have made use of data that records the "surcharge for county waste per ton (in $)", and "cities in a county with a population greater than 5,000 that provide curbside recycling to residents".

The average surcharge per ton of waste was $11 in 1997, which had increased to $26 by 2013. This surcharge varies widely across the counties, with standard deviation of $16 in 1996 and $44 in 2013.

The Minnesota statutes in 1989 required that cities with a population greater than 5,000 must provide curbside recycling to residents. Collection of these data started in 1996, with 48 counties offering curbside recycling. It stayed relatively stable, increasing slightly to 51 in 2004, but dropping slightly to 49 in 2013.

In addition, socio-economic factors such as education levels and income can affect recycling as well as consumption (and therefore, garbage) behavior. To account for these factors, I collected data for income and education level from the U.S. Census. The instrument that I use in this analysis is the unemployment rate, which was also obtained from the Census. The summary of all of these control variables that might correct a potential omitted variable bias is listed in Table 2.1.

2.3.1 Index of Social Capital

To understand the effect of social norms on recycling and waste generation behavior, an effective way of measuring social norms is required. In the absence of data on the social norm of recycling, I used the more general measure of social capital to estimate the impact on recycling and waste generation. In the definition used in Putnam (1993) for social capital, he asserts that activities within group associations promotes cooperation, which can solve collective action problems within the community. Since this study uses county as the unit of analysis, we can use the county-level social capital measure that I construct here as a way to understand how a community comes together to engage in a civic activity like recycling. We use this social capital index in this chapter as a statistic that measures the social norm of recycling.

To obtain data for such a measure, I use the methodology used for the county-level measure of social capital as published by Northeast Regional Center for Rural Development, at the Pennsylvania State University (Rupasingha et al., 2006). The Center uses data on the number of religious, civic and social, business, political, professional, labor, bowling, fitness, golf and country club, sports club establishments, non-profit organizations, voter turnout, and census response rate to develop the social capital index. They apply principal component analysis to these data to arrive at a comprehensive index of social capital. Thus far, however, the Pennsylvania State University has released this index only for the years 1990, 1997, 2005, 2009, and 2014.

As per the methodology prescribed in Rupasingha et al. (2006), the focus in creating this social capital index is mainly on organizations that are not rent-seeking. In doing so, the index includes organizations that promote trust and cooperation among the community of people but does not

include those that are created for financial reasons. Being in such organizations involves repeated interactions leading to an atmosphere of greater trust and information sharing. The increased sense of cohesiveness and information sharing could be of great assistance while promoting behaviors such as recycling. The advantage that this social capital index has over others that rely on surveys such as the General Social Survey and World Views Survey is that the data are from secondary sources and are not based on individuals' views on trust or related concepts.

Since this index is constructed at the county-level for 18 years, we can see variations in the index across the counties as well as over time. Rupasingha et al. (2006) show in their paper that this social capital index is strongly associated with ethnic homogeneity, income inequality, education, age, community attachment, and female labor force participation. Hence, any changes in these factors within the counties, the states or the nation, and over time could cause a change in the social capital index. Rupasingha et al. (2006) notes that this is one of the first measures of social capital at the county level that combines multiple dimensions of the complex concept of social capital.

Since the Pennsylvania State University ("Penn State") data have only limited number of years of data, I used the principal component analysis methodology used by Rupasingha et al. (2006) to create a social capital index for all the years for which I have recycling data, for every county in Minnesota. The sources for the data to create this index consists of four types and are as follows. The first type is the total number of group organizations in each county, which include religious, civic and social, business, political, professional, labor, bowling, fitness, golf and country club, sports club establishments, and were obtained from the U.S. Census Bureau from the County Business Patterns data. For each establishment, as listed above, the Census has the numbers for all U.S. counties from 2005 to 2016.

The second type is the census response rate, or the Census Participation Rate, and is defined as the percentage of questionnaires mailed back by households that received them. The data for these rates were published for the years 2000, 2005, 2010, and 2015. In my analysis, I used the 2000 Census Participation Rates for the years 1996 to 2000. Similarly, I used the 2005 Census Participation Rates for the years 2001 to 2005, and the 2010 Census Participation Rates for the years 2006 and onwards. For voter turnout (the third type), I used the presidential voter turnout from the years 2004, 2008, and 2012 from the data published by Penn State (Rupasingha et al., 2006). This data is from Dave Leip's Atlas of U.S. Presidential Elections. Similar to the Penn State methodology, I used the 2004 data for the years 2004 to 2007, the 2008 turnout data for the years 2008 to 2011, and the 2012 data from 2012 to 2015.

The fourth type is the number of non-profit organizations, excluding those with an international approach, which is obtained from Urban Institute's National Center for Charitable Statistics database. They have county level data for the number of non-profit organizations for the United

Table 2.1: Summary statistics: County Averages from 1996 to 2015

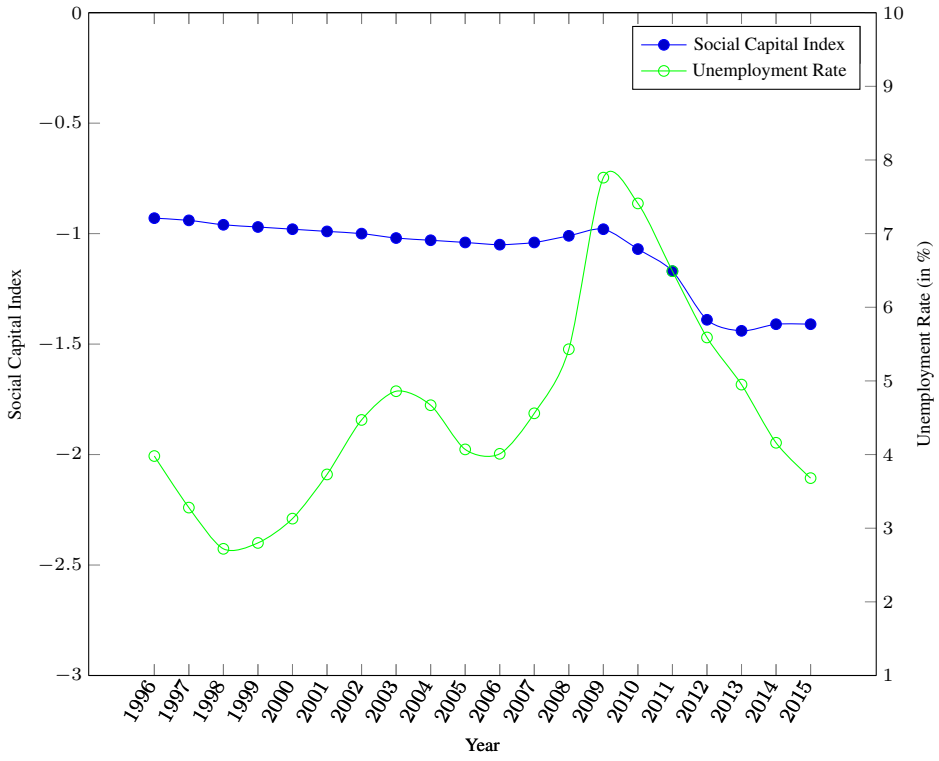
	Measures	Observations	Average	Standard Deviation	Minimum	Maximum
Outcomes	Recycling rate (in %)	1,394	40.09	9.60	5.67	80.17
	MSW per 1,000 persons	1,394	1108.1	291.3	307.8	3069
	Recycling per 1,000 persons	1,763	456.2	212.1	34.8	2444.3
Main variables of interest	Social capital index	1,762	-1.1	1.1	-3.3	7.8
	Non-profit organizations per 10,000 persons	1,763	60.2	30.5	21.4	372
	Total establishments per 10,000 persons	1,007	1.3	.6	.5	4.4
	Environmental organizations per 10,000 persons	938	1.9	1.2	.3	17.9
	Voter Turnout (in %)	1,007	69.7	22.3	58.8	91.2
	Census Response Rate (in %)	1,680	76.6	7.7	24	88
Control variables	Unemployment rate (in %)	1,759	4.57	1.7	1.7	15.1
	Median household income (in $)	923	60,919.4	11,850.1	33,834	94,057
	Surcharge per ton of waste (in $)	1,596	.82	7.1	0	121
	Persons 25 or older, with bachelor degree (%)	1,703	30.8	10.5	9.1	47.6

Sources: Minnesota Pollution Control Agency, U.S. Census Bureau
Notes:
The values of all (except median household income and voter turnout) the above descriptive statistics are for the years 1996 to 2015. The values for Median household income and voter turnout are for the years 2005 to 2015 because those are the years for which data exists. The averages are all population-weighted averages.

States from 2005 to 2016.

As stated in the methodology in Penn State's social capital index, I standardize the four types listed above to have a mean 0 and standard deviation of 1. I then apply the principal components method to combine these four factors to arrive at the first principal component, which I then use as the index of social capital. This methodology gives me data for all counties of Minnesota from 2005 to 2015. Since recycling data goes all the way back to 1996, I used linear extrapolation so that I can extend the data. I will use this index as a measure of the social norm of recycling rate. As part of my robustness checks, I verify that my results of the impact of social norms in later sections, I also use the number of non-profit organizations (separately) as a measure of social norms.

This social capital index has fallen by over 50% from 1996 to 2015. A large part of this decline happened from 2009 to 2013 when it dropped by 46 %, possibly due to the Great Recession. In the beginning of this period, from 1996 to 2009, the decline was more moderate with a decrease of 6 %. The total number of non-profit organizations remained stable in this period with a modest increase of 2%, with an average of 61 non-profits per 10,000 persons across MN counties. However, the number of non-profits also saw a fall of 9% from 2009 to 2013. The average number of establishments (that included many types of social organizations as listed in earlier paragraphs) per 10,000 persons decreased by 7% in this period, consistently declining the whole time (see Table 2.1 for a summary). The voter turnout and census participation, however, increased over time. The voter turnout rate increased by 7 percentage points from 2005 to 2015, and the census participation or response rate increased by 7 percentage points from 1996 to 2015.



Sources: Bureau of Labor Statistics, U.S. Census Bureau,
Dave Leip's Atlas of U.S. Presidential Elections, Urban Institute

Figure 2.2: Social Capital Index and Unemployment Trends

2.4 Empirical Framework

Trends and in the last 20 years show a decline in social capital in the state of Minnesota, falling over 50% from 1997 to 2013 (Figure 2.2). At the same time, there was an increase in the recycling rate of 16.9%, and a 5.8% rise in municipal solid waste (MSW) per 1000 persons in the state (Figure 2.1). The total MSW generated by the state, however, stayed relatively stable through the last 2 decades, declining slightly by around 2%.

These trends do not, however, indicate that social norms have a causal impact on recycling behavior or waste generation. From 2.5, we know that social norms are endogenous since they evolve over time based on the recycling behavior in previous periods. In order to comment on the causal relationship, therefore, I will be using an instrumental variable framework to comment on the direction of the impact. My instrument will be unemployment rate, where I posit that any change in the unemployment rate is a plausibly exogenous shock and that it will affect waste per capita through the social norm.

Following the hypotheses that has emerged from the Proposition 2 from Section 2.2, my main equation of interest is:

$$Y_{it} = \alpha + \gamma R^*_{it} + \beta Z_{it} + \tau_t + \delta_i \times t + \kappa_i \times t^2 + \epsilon_{it} \tag{2.11}$$

where Y_{it} is the outcome of interest. Following the two parts of proposition 2 (as shown in Section 2.2), I am interested in the impact of social capital on three outcomes, i.e., R_{it}, the amount recycled per waste generated (or recycling rate), the amount recycled per capita, and the waste generated per 1000 persons (X_{it}). R^*_{it} is the social norm in county i at year t. The control variables, which consist of county demographics characteristics and other variables are represented by Z_{it}. The variables include the percentage of people who 25 years or older with a bachelors' degree, income, and also some of the ordinances and statutes that prevail in the county at that time. The terms $\delta_i \times t$ and $\kappa_i \times t^2$ represent the county-specific linear and quadratic trends, which control for systematic trends in each county, and τ_t is the year fixed effect, which captures the impact of any unobserved state-wide event in a given year that is common to all counties.

In Equation (2.11), the coefficient of interest is γ. If social norms were exogenous or randomly assigned, the estimate of γ in Equation (2.11) would be the causal impact of norms on recycling rate. In other words, the main hypothesis to be tested is $H_0 : \gamma = 0$. However, R^*_{it} is not randomly assigned since social norms clearly evolve over time and depend on a number of factors that may be correlated with the error term in Equation (2.11).

As part of my identification strategy, I enumerate all the possible issues I might face and how I address them in order to estimate this causal impact. Broadly, some of the sources of endogeneity issues can be classified into (i) reverse causality; (ii) unobserved heterogeneity and omitted variable bias; (iii) measurement error; and (iv) violation of the SUTVA. The following paragraphs explain how I address all these issues with the model, after which the instrumental variable identification strategy is described.

It is possible that recycling rates and social norms are jointly determined, with both variables affected by a common set of unobserved factors. This reverse causality could potentially bias the main coefficient of interest, γ. Specifically, there could even be a situation where households that particularly care about recycling migrate to certain counties and thereby determine the social norm of that county. Data from Current Population Survey (CPS) of the U.S. Census Bureau indicate that all migration in the U.S. due to climate change has been less than 1% for all the years from 1998 to 2018. Miscellaneous reasons other than employment, health, familial, housing, education, climate change or natural disasters form around 5% in the last 2 decades. We could, perhaps, conclude from this that migration due to environmental reasons form a negligible part of geographic mobility.

Equation (2.11) includes fixed effects in the model, both county-specific fixed effects as well as county-specific linear and quadratic trends. Since year fixed effects are not included, any variation within counties could come from the counties' responses to national, state or county events for that year. Rupasingha et al. (2006) show in their paper that changes in inequality, education, age, ethnic homogeneity, and female labor force participation can cause changes in social capital within counties.

If reverse causality as specified in the preceding paragraph causes certain counties to have more recycling oriented households leading to higher baseline recycling rates that is invariant across time, this will be captured by the county-specific fixed effects. In addition, I have also included literacy rates as well as median household income as control variables to capture some of this sorting behavior.

Similarly, the county fixed effects will capture any unobserved heterogeneity causing higher recycling rates in a county that is invariant across time. The linear time trends will capture any overall trends for the state of Minnesota that occurs across all the counties of the state. For instance, the Minnesota statute in 1989 established that counties shall ensure that residents get an opportunity to recycle. Such laws that are applicable to the entire state could raise the recycling rates for all counties in that year and keeps increasing for the years after the year of implementation.

In addition, there could also be unobserved heterogeneity that varies systematically over time. In order to account for that, I have also included county-specific linear trends as well as county-specific quadratic time trends. The assumption that I make here is that once we include all these fixed effects, the remaining unaccounted and unobserved heterogeneity will not substantially bias the main coefficient of interest.

Another source of endogeneity is measurement error. Since recycling rate is the dependent variable in my analysis, measurement error in the dependent variable that is not systematic does not bias the coefficient on the main variable of interest. However, there could be measurement error in the data that are particular to specific counties or specific years. If there are such errors, the battery of fixed effects in these models will capture some of these county-specific or year-specific quirks in the data.

There could also be measurement error in the social capital index. I argue here that if this measurement error present, it is minor since the index combines data from 15 different variables collected from the Census, NCCS and David Leip's Atlas data. Hence, whatever measurement error is present in any one of the variables would be reduced because of the amalgamation of variables.

The other identification challenge is the violation to the stable unit treatment value assumption

(SUTVA). The SUTVA states that social norm in a particular county in some year will not affect the recycling rate or waste in another county-year. It is possible, however, that a resident in a town with a higher prevailing recycling social norm might be close enough to the border of the county that she might cross over and recycle in the neighboring county, thus raising the average recycling rate in the neighboring county. Over time, though, the neighboring county's recycling rate could ultimately affect the social norm of recycling in its own area, which will affect its own recycling behavior. Since it will affect the social norm before it affects the recycling behavior, this may not be a big issue. Nevertheless, it remains a limitation of this analysis.

While the identification challenges in the preceding paragraphs may not be serious, my method of bringing out the causal impact of social norms is the two-stage least squares (2SLS). Following Equation (2.11), I instrument the social capital with the unemployment rate in each county for that year.

The 2SLS model uses county fixed effects, linear trend, county-specific linear and quadratic trends. So, the variation that is left off of the estimation once all these fixed effects are included, can be reasonably concluded as plausibly exogenous. After taking into account within-county and within-year averages, and also systematic trends, the goal is to measure the impact of within county-year shocks to unemployment rate on social norms.

There are papers that have pointed out that individuals who enter into unemployment will see a change in their well-being (Clark, 2003). We can say that this might have positive or negative impact on the way these unemployed people interact with society. Therefore, a change in unemployment is likely to change the social capital of the area that the individual lives, which in turn can affect the way the individual recycles or wastes. Of course, there can be ways in which a change in unemployment might have a direct impact on the recycling behavior, which is the reason we employ the Conley et al. (2012) technique to test the strict exogeneity assumption of the instrument. In effect, the assumption is that these shocks to unemployment rate are plausibly exogeneous and will affect recycling and waste generation behavior only through social norm.

In addition to all these components to this analysis, the residuals from the recycling and waste regressions could be correlated within each county even after controlling for fixed effects and the variables mentioned above. Following Bertrand et al. 2004, and as is standard practice, I cluster standard errors for all the regressions at the county level so that they are robust to heteroscedasticity and serial correlations between error terms.

It must be noted that the estimated effect of the impact of social norms on recycling rate as well as the waste generated per capita, from the model described above, is the local average treatment effect (or LATE). In other words, the estimated effect of social norms on recycling rate (or waste per

capita) measures the impact for only those counties whose unemployment shocks lead to greater or less cohesiveness among the people in those counties, causing social capital to rise or fall. This is the main difference between the estimates from the 2SLS and the OLS models, where the estimated OLS coefficient is the average treatment effect (or ATE).

It should be mentioned here that I tried many possible instruments in addition to unemployment rate, in attempting to estimate the impact of social norms. In order to pin down variables that could potentially cause shocks to the social capital of a certain area, I used the number of refugees [3] per capita, population, and population growth as possible instruments that may explain the variations in social capital. I obtained the data on refugees from Minnesota Department of Health, and the population data from Census. These variables, however, were not strong instrumental variables. Unemployment rate was the only variable that had a high F-statistic, i.e. a strong explanatory power in the first stage regression. I used only contemporaneous variables as possible instruments, following Bellemare et al. 2017, where they find that lagged instruments can often lead to more endogeneity issues than less.

While there could be a number of challenges to identification, as mentioned above, there could also be issues related to internal validity when it comes to instrumental regressions. Specifically, Cole and Fletcher (2009) prove that, in the case of the impact of social networks, one can often show that a relationship exists when it really does not. Since my paper deals with social norms, which can potentially impact household behaviors in several ways, I check for internal validity of the IV methodology through a falsification test to verify that a variable that should not have a relationship does indeed show no significance. I use data on a fake treatment (i.e., waste tires recycled as a percentage of MSW in every county and year) to test whether (i) it impacts the recycling rate and MSW per capita; and (ii) social norms along with the control variables affects recycled tires.

Even though recycled tires are part of total recycled materials, residents cannot discard tires as part of their trash. Every county provides a list of companies that residents have to contact in case they need to dispose off waste tires. Hence, disposal of tires (and how to recycle them) is not a choice that could possibly be affected by changes in social norms or other household demographics. Ideally, therefore, it should not be impacted by total recycled amount or total waste. More importantly, it should not be affected by social norms at the margin, after being controlled for by fixed effects and other control variables.

[3] Arrival of Primary refugees that includes "Primary Refugee", "Asylee", "Parolee", "Amerasian", and "Victim of Trafficking", by initial county of resettlement

2.5 Do social norms impact recycling and waste generation behavior? Results

Using the strategy described in the previous section to deal with endogeneity, I estimate the results for the three outcomes of interest: recycling rate, recycling per capita, and MSW per capita. The results are presented for different combinations of fixed effects.

Table 2.2: Ordinary Least Square Regression Results for Recycling Rate

	[1]	**[2]**	**[3]**	**[4]**
Social capital index	0.146	-0.090	-1.771	1.581
	(1.495)	(1.408)	(1.631)	(1.913)
Persons with bachelor degree	-0.437	-0.751**	-0.929***	0.305
	(0.437)	(0.342)	(0.349)	(0.444)
Cities that provide curbside recycling				1.273
				(1.293)
Observations	1345	1345	1345	1345
Adjusted R^2	0.210	0.197	0.487	0.626
County Fixed Effects	Yes	Yes	Yes	Yes
Linear Trend	No	Yes	No	No
Year Fixed Effects	Yes	No	No	No
County-Specific Linear Trends	No	No	Yes	Yes
County-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Recycling rate (as a proportion of MSW); clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county. Constants are not displayed. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Tables 2.2, 2.3, and B1 (in the Appendix) show the results for the ordinary least square (OLS) regressions, showing the impact of social capital (the ATE if social capital were exogenous) on the three outcomes—recycling rate, recycling per capita and MSW per capita. Each column in these tables represents a model with different combinations of fixed effects. The column [4] is the model with county fixed effects, county-specific linear and quadratic fixed effects. In addition, it also controls for literacy as well as the presence of curbside recycling.

Column [1] in Table 2.2 shows that with county and year fixed effects, a change in social capital does not have significant impact on recycling rate. Column [2] shows the result of the impact of social capital when county fixed effects and a linear trend are added as control variables, whereas Column [3] shows the impact when county fixed effects and county-specific linear trends. Both Columns [2] and [3] show that the impact of social capital remains insignificant. Column [4] shows

Table 2.3: Ordinary Least Square Regression Results for MSW per Capita

	[1]	**[2]**	**[3]**	**[4]**
Social capital index	15.178	-5.339	6.749	-123.210**
	(38.485)	(38.233)	(39.608)	(54.027)
Persons with bachelor degree	-36.058**	5.886	48.913***	29.334**
	(14.298)	(8.183)	(11.722)	(11.940)
Cities that provide curbside recycling				-32.088
				(21.369)
Observations	1345	1345	1345	1345
Adjusted R^2	0.198	0.000	0.448	0.646
County Fixed Effects	Yes	Yes	Yes	Yes
Linear Trend	No	Yes	No	No
Year Fixed Effects	Yes	No	No	No
County-Specific Linear Trends	No	No	Yes	Yes
County-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: MSW per capita per 1,000 persons; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county. Constants are not displayed. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

the model that includes county-specific linear or quadratic trends along with county fixed effects, where the impact of social norms is positive although insignificant. I will consider the fourth column with this specific battery of fixed effects as my final specification because it offers the best fit in terms of R^2. The corresponding column in Table B1 shows that the impact of social norm on recycling per capita is also insignificant.

The impact of social capital on waste generated is, however, negative and significant following the results from Proposition 2 in Section 2.2. From Column [4] in Table 2.3 we can see that social norms has a negative and significant impact on MSW per capita, implying that an increase in the social capital making the county more cohesive leads to lesser waste generation.

Table 2.4: Endogeneity Checks for the Impact of Social Capital on Recycling and Waste Generation

Dependent Variable:	Recycling rate			Recycling per capita			MSW per capita		
	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**
Social capital index	-0.767	-4.640*	-2.810	-43.509	-121.827	-149.649**	-131.437**	-188.695***	-183.899***
	(1.906)	(2.396)	(2.366)	(48.489)	(75.277)	(71.948)	(64.922)	(71.706)	(63.455)
Cities that provide curbside recycling	1.547	-0.688	2.193	33.687	9.750	16.643	15.838	-44.969***	-35.595**
	(2.418)	(1.794)	(1.690)	(25.285)	(20.569)	(20.416)	(28.609)	(17.402)	(16.213)
Persons with bachelor degree		-0.614**	-0.183		5.293	5.789		20.213	12.712
		(0.286)	(0.488)		(6.867)	(5.425)		(14.902)	(14.337)
Observations	1394	1345	1345	1595	1539	1539	1394	1345	1345
Adjusted R^2	0.009	0.375	0.527	-0.005	0.211	0.506	-0.258	0.195	0.500
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Trend	Yes	No	No	Yes	No	No	Yes	No	No
County-Specific Linear Trends	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
County-Specific Quadratic Trends	No	No	Yes	No	No	Yes	No	No	Yes
F-Statistic (Instrumental Variable)	64.05	126.11	305.88	59.30	78.48	157.32	64.05	126.11	305.88

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county.
Unemployment rate is used as the instrument for social capital. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

OLS regressions may result in biased estimates because social capital is not exogenous and may not even be plausibly exogenous. Based on the identification strategy described in 2.4, we run instrumental variable regressions for the three outcomes. As a first step to the instrumental variables setup, I assess the strength of the impact of the instrumental variable, i.e., the unemployment rate, on the social capital index in Table B2 in the Appendix. In other words, the first stage regression results in Table B2 shows the impact of unemployment rate is positive and significant at 99% significance level. I choose model [3] as the final specification because it has the highest R^2 and the highest F-statistic, making it the best suited specification for the first stage regression.

The results of the 2SLS estimation are shown in Table 2.4, and are mostly similar to the OLS estimates. We see that the effect of social capital on recycling rate is insignificant as can be seen in Column [3] in Table 2.4. Column [9] shows that MSW per capita falls with an increase in social capital, and this effect is significant at 1% level, reiterating the result that we got in Table 2.3 and also the result from the proposition 2 in Section 2.2. The result from the 2SLS model shows a decline of 184 tons per person of MSW, and the OLS model shows a decline of 123 tons per person. Since the two results are both negative and very similar in magnitude, we can say that endogeneity may not be a big issue when it comes to the impact of the social capital on waste per capita.

I also show the results of the impact of social capital on recycling per person in Column [6] in Table 2.4. Here, the estimates show that the impact of social capital on recycling per capita is negative and significant at 5% level, which is different from the impact of social capital on recycling rate. This result could be key in explaining the above two estimates of the impact of social capital on recycling rate and MSW per capita. I claim here that MSW is the main driver of the results. When the social capital increases, increasing the cohesiveness of the group, households get more conscious of the waste they generate leading to lower waste per capita. In addition, they consume less recyclable materials and hence reduce the recycling per capita too, as shown in Column [6] in Table 2.4. Since the recycling rate is total recycling as a percentage of total MSW, it is possible that when both recycling per capita as well as MSW per capita decline, the rate of recycling remains unchanged.

In the 2SLS regressions in Table 2.4, we consider Columns [3], [6] and [9] as the main regressions that show the impact of social norms. We see from the results that while social capital does not affect recycling rate, it significantly affects both MSW per capita and recycling per capita negatively. The recycling amount per person could have declined because the environmentally aware households might be consuming fewer recyclable materials and more environmentally friendly products that need not be recycled.

Although these results could be plausibly exogenous and the estimates from Table 2.4 may measure the impact, the social capital index could be a noisy indicator of social norms. In this case,

it becomes necessary that we run robustness checks using another indicator of social norms.

As a robustness check, I use one of the components that I use to construct the social capital index, i.e, non-profit organizations per 10,000 persons as another indicator of social norms. The reason for choosing to run the robustness check with this variable is because non-profit organizations have larger numbers per 10,000 persons than the other type, i.e., total group organizations, and are therefore prone less volatility. In addition, it also has data for more number of years than voter turnout and census participation. I ran the 2SLS model with this variable instead of the social capital index. The results of these regressions, once again using unemployment rate as the instrument, are shown in Table B3. The results are in the same direction as with the social capital index, where the impact of social capital on both recycling per capita and MSW per capita are negative and significant.

All the 2SLS results are valid under the assumption that the changes in IV (i.e., unemployment rate) are strictly exogenous. This may, however, not be true. As robustness checks, therefore, I then check whether these results hold when there are small departures from the strict exogeneity assumption. I use the method developed by Conley et al. (2012) to address the fact that the instrument I am using in this analysis could be plausibly exogeneous.

To run these robustness checks, I use the intermediate local-to-zero (LTZ) method built by Conley et al. (2012), which requires a prior on the mean and standard deviation of γ, where the prior measures the extent of departure from the strict exogeneity assumption. For the prior, I assume a mean of 0 and standard deviation of 5, which assumes that while strict exogeneity holds in expectation, there can be big departures from strict exogeneity.

The LTZ checks of the 2SLS results from Table 2.4 show that all of the estimates hold. The impact of social capital on MSW per capita, recycling per capita as well as recycling rates in the LTZ checks are similar to estimates in Table 2.4 showing that the results are robust to departures from the strict exogeneity assumption.

The confidence interval estimated by the LTZ method for the 2SLS estimate for the impact of social capital on MSW per capita is [-296.7, -62.2] at the 95% level. Similarly, the estimated impact of social capital on the recycling rate and recycling per capita using the LTZ method are [-90.8, 85.3] and [-265.5, -33.8] 95% confidence intervals respectively. The estimates are significant at 1% for the impact on MSW per capita, significant at 5% for the impact on recycling per capita and not significant for the impact on recycling rate, which are similar to the results in Table 2.4. This shows that the results are robust to movements from the strict exogeneity assumption.

Of course, social capital or norms could use several pathways, different from the one described

above, to affect recycling or waste generation behavior. Since waste tires cannot be disposed off as part of the rest of the garbage, it is not a household choice variable and hence, should not be affected by social norms even though it moves with recycling rate and MSW figures. So, as stated in Section 2.4, I performed placebo and falsification tests to show that the estimates in Tables 2.4 and B3 are not spurious. Results in Columns [1] and [2] in Table B4 show that the effect of social capital is small and insignificant, confirming that the main coefficient of interest is not spurious, confirming that the results are internally valid.

2.6 Conclusion

Social capital and social norms can be used as a tool to bring about change in behavior such as recycling and waste generation. It is possible that there is no clear incentive for a household to recycle other than social approval, in which case policy can be used to improve the social capital of a community in order to change such behaviors (Nyborg et al., 2016). Since a framework for social capital and its determinants was not established until very recently, this idea was not used as a policy tool effectively (E.L.Glaeser, 2001). Rupasingha et al. (2006) built the methodology to create a comprehensive measure of social capital and also examined its many determinants. Using their methodology, I created panel data on social capital for the 87 counties of Minnesota from 1996 to 2013 that I then use in this paper to measures its impact.

I examine the impact of social norms and capital on recycling and waste generating behaviors in two parts in this paper. Firstly, using a theoretical model, I show that increasing social norm of recycling increases recycling rate and decreases waste per capita. This follows some of the literature (such as Bodvarsson and Gibson, 1999) that focuses on the public goods provision that states that people engage in recycling or waste reducing behaviors mainly to comply with the prevailing social norms. Secondly, these two propositions that emerged from the first part of the paper are the basis for my hypotheses in the empirical framework.

The theoretical model in this paper arises from the fact that recycling behavior may not entirely be based on monetary incentives alone, and can be based on certain forms of altruism such as peer approval and social norms. Using tipping behavior as a model (Azar, 2004), I claim that social norm for recycling can affect how a household recycles as well as generates waste, in a model where households decide to consume, produce waste and recycle based on both economic incentive as well as in comparison to the social norm. From this setup, two propositions arise. When social capital increases, firstly, recycling rate increases, and secondly, waste per capita declines.

Using instrumental regression method in the empirical section, I measure the impact of social

capital on recycling rate and waste per capita using aggregate data for Minnesota for the years 1996 to 2013 for the first time. Here, I would like to impress upon the large dataset that was collected from various sources. Specifically, I created a new social capital index that combines several dimensions of social capital, following the methodology drawn out by Rupasingha et al. (2006). The recycling and waste data are from MPCA, and other socio-economic factors are from the Census. In addition, I also collected policy data from the SCORE reports as published by the MPCA. Using the comprehensive dataset for the 87 counties of Minnesota spanning 18 years, I set out to answer the two propositions from the theoretical model.

Following the empirical framework, I show that waste per capita declines significantly by 184 tons annually with a 1 unit increase in social capital, which is in line with the results from the theoretical model. The impact of social capital on recycling rate, however, is not significant. Even though this is not a result from an experiment (i.e., the ideal case), the results are robust with different measures of social capital, alternative specifications, falsification and placebo tests. These 2SLS estimates are also robust to deviations from the strict exogeneity assumption.

Given the instrumental regression framework, this seemingly causal results are encouraging as far as policy is concerned. The findings in this paper suggest that social norms could act as a potential policy tool. It might be possible, therefore, to use social norms in conjunction with policy to bring about change in recycling behavior (Nyborg et al., 2016). It must be noted, however, that the social capital index is a noisy one in which mechanisms of how households use social norms to make recycling decisions are still unclear. It would definitely be interesting, therefore, to understand this behavior using a primary survey of households. This endeavor would be useful to policy and is still open to further research.

Chapter 3

Combating Particle Pollution: Estimating the Impact of Pollution Standards on Mortality in India[1]

Abstract

India experiences some of the highest air pollution levels globally, with 13 of the 20 most polluted cities in the world. In this paper, we estimate the impact of air pollution policies in India on mortality for people of all ages and all causes ("mortality") . Our focus is on (overall) mortality, rather than infant mortality, because the overwhelming burden of air pollution is borne by older populations. We use a differences-in-difference setup to estimate the impact of two major air pollution regulations, the Supreme Court Action Plan (SCAP) and the Catalytic Converter (CC) policy, on mortality in India. Using new mortality data, we find evidence that the CC policy reduces mortality in the urban areas of India by 19%. In the second half of the paper, we compare the direct impacts of different pollution measures on mortality. Using new satellite data on fine particulate matter ($PM_{2.5}$), we find that $PM_{2.5}$ exposure is associated with increased mortality, but TSP exposure is not. A 10% increase in $PM_{2.5}$ leads to a 2.5% increase in mortality, while the impact of TSP is not significantly different from zero. We suggest, therefore, that policies in India that focus on $PM_{2.5}$ rather than TSP are likely to lead to greater gains in health.

Keywords: regulation, mortality, pollution

[1] co-authored with Jay S. Coggins and Andrew L. Goodkind

3.1 Introduction

Anthropogenic air pollution continues to be a serious threat to human health around the world, and perhaps nowhere more seriously than in India. The most recent Global Burden of Disease (GBD) study, from 2016, estimates that more than one million deaths in India that year could be attributed to ambient fine particulate pollution (IHME, 2016, Lim et al., 2012), a number that amounts to more than one quarter of the world's deaths due to this cause. A more recent study by Burnett et al. (2018) estimates that the number of early deaths in India caused annually by fine particulate pollution less than 2.5 microns in size ($PM_{2.5}$) could be more than two million.

These estimates of premature mortality due to air pollution, included in the GBD, are based upon scientific evidence from the epidemiological literature. This literature is based largely on observational cohort studies from the US and Europe, with relatively low $PM_{2.5}$ exposures, which are then projected out and calibrated to high-exposure countries. Two of the leading studies, both observational in nature and tracking large populations of U.S. subjects over decades, are the Harvard Six Cities study (Lepeule et al., 2012) and the American Cancer Society study (Krewski et al., 2009). Another cohort study (Beeler et al., 2014) assesses the impact of long-term exposure to pollution in Europe, where pollution levels are smaller than in high-exposure countries such as China or India. Of the 15 cohort studies included in Burnett et al. (2018), only one includes populations exposed to high $PM_{2.5}$ concentration (Yin et al., 2018) which examines cause-specific mortality of men in China with $PM_{2.5}$ exposures between 15 and 84 $\mu g/m^3$. There is currently no cohort study of the health effects of air pollution based upon Indian data. Understanding the relationship between high exposure to $PM_{2.5}$ (as high as 120 $\mu g/m^3$) and mortality in India is a void in the literature that this paper seeks to fill.

Another growing literature, distinct from but complementary to the epidemiological studies, is to be found in economics journals.[2] This literature exhibits a focus on causality, with various quasi-experimental designs aimed at identifying the causal connection between health outcomes such as mortality, on one hand, and either a policy intervention or air quality itself, on the other. It also exhibits a tendency to focus on mortality of infants rather than among all ages[3], as well as a tendency to focus on relatively large particles such as particulate matter less than 10 microns (PM_{10}) or total suspended particles (TSP).[4] In the scientific literature, as well as in U.S. EPA benefits assessments for air-quality rules, the evidence shows that the preponderance of lives saved as a result of tighter

[2] For a recent survey, see Graff Zivin and Neidell (2013).

[3] See, for example, Chay and Greenstone (2003), Currie and Neidell (2005), Currie et al. (2009), and Greenstone and Hanna (2014), where infant mortality is at issue, or Lleras-Muney (2010) where the impact on children's health outcomes, measured by children's hospitalizations, is estimated. Contrary cases that explore mortality among all ages include Anderson (2015), Heutel and Ruhm (2016).

[4] See, for example, Graff Zivin and Neidell (2013), Greenstone and Hanna (2014), and Heutel and Ruhm (2016).

air-quality standards are adults[5], and also that the most harmful fraction of particle pollution is $PM_{2.5}$.[6]

In the present paper, we apply empirical methods to data from India, aimed at identifying the causal relationship between $PM_{2.5}$ and what we shall call mortality, by which we mean mortality due to all causes among all ages. Our work owes a significant debt to that of Greenstone and Hanna (2014), the only study of air pollution and health based on Indian data of which we are aware. In the first part of this paper we employ a similar methodology to Greenstone and Hanna (2014), and more importantly, we use the dataset they made available along with their published work. We build on their framework, extending the analysis in two main directions. The first is to examine the effect of Indian environmental policy on (overall) mortality rather than only on infant mortality. This is made possible here due to a measure of mortality in Indian urban districts that we obtained from vital statistics records found in non-digitized government publications. Second, we employ a spatialized dataset of $PM_{2.5}$ concentrations for India that are generated from satellite observations and directly estimate the impact of $PM_{2.5}$ concentrations on mortality. This data source is relatively recent and will, we anticipate, allow economists and others to extend analyses of the threat of $PM_{2.5}$, the most harmful of air pollutants, into regions of the world where monitoring data are sparse or unavailable.

The paper contains three contributions, the first two related to the impact of policies to combat pollution and the third related to mortality data. First, using our new data on mortality, we establish the connection between clean-air regulations in India and (overall) mortality as opposed to only infant mortality, while also addressing endogeneity issues that may plague mortality that includes adults. We find that Indian air-quality regulations, implemented in the 1990s, were effective at reducing mortality by 19 percent. This result differs from that of Greenstone and Hanna (2014) for infant mortality, who find no significant impact.

Greenstone and Hanna (2014) show that the Indian air pollution regulations reduce both TSP and SO_2 significantly, and yet we still find high pollution levels in India. India's annual average $PM_{2.5}$ level in 2016 was 76 $\mu g/m^3$, as compared to 56 in China and 9 in the U.S (Based on Brauer et al., 2016). Hence the second contribution is to estimate directly the relationship, in 117 districts in India, between concentrations of $PM_{2.5}$ and mortality, and compare it with the effect of other pollution types. It is the availability of new satellite data on $PM_{2.5}$ that allows us to perform this estimation. The empirical approach we use is based on fixed effects, for both districts and time. A set of robustness checks confirms that higher levels of $PM_{2.5}$ do indeed have a positive effect on

[5] As an example, estimates in (U.S. EPA, 2012) show that reducing the U.S. National Ambient Air Quality Standards (NAAQS) for fine particles from 15 to 12 $\mu g/m^3$ saves as many as 1,000 adult lives in the U.S., but saves only one infant's life.

[6] See, for example, Figure 5 in Pope et al. (2002).

mortality rates. We find that SO_2 also has a positive impact on mortality, which becomes insignificant when district-specific quadratic trends are added to the specification. Moreover, we find no evidence that TSP and NO_2 have significant effects on mortality. This result suggests that the policies implemented under India's 1981 Air (Prevention and Control of Pollution) Act might achieve greater gains in health if they were revised to focus on $PM_{2.5}$ rather than TSP.

The third contribution is related to mortality data that we compiled from the Census of India. It is well known that the registered crude death rates from the Census underestimate the true mortality rate (Office of the Registrar General, 2010). In order to understand the extent of this measurement error, the Census statistically *estimates* the "actual" average mortality rate (annually) in each state of India using the Sample Registration System (SRS) survey, a representative sample of India. We use this state level SRS estimate to adjust the district level registered crude death rate (or what we will call "crude mortality rate") for urban areas. Our results using this novel measure of adjusted mortality rate, show a smaller impact from elevated $PM_{2.5}$ levels than do the unadjusted mortality rates. This result is possibly due to the fact that the adjusted mortality is not merely a scalar adjustment, in reality it follows a different trend compared to crude mortality rate (CMR) because the record keeping capabilities of some regions in India are worse than those in other regions.

The paper proceeds as follows: Section 3.2 discusses the environmental regulation history in India. Section 3.3 describes the data and its sources. Section 3.4 presents the empirical framework and the results for the estimates of the impact of regulations on pollution and mortality. Section 3.5 does the same for the effect of air pollution on mortality. Section 3.6 concludes.

3.2 History of Environmental Regulations in India

The Government of India passed the Air (Prevention and Control of Pollution) Act in 1981, which extended to the whole nation to prevent, control and abate air pollution. Despite having this legislation in place, it was not until 1984; when the toxic leak from a Union Carbide plant affecting the city of Bhopal in India, that comprehensive environmental protections came into effect.

The Bhopal leak is often considered the worst industrial disaster in the world, killing at least 3,800 people and causing many more thousand premature deaths (Broughton, 2005). It triggered a series of events that included reports from various organizations commenting on the lack of precautions in place, as well as public interest litigations filed in the Supreme Court of India. Eventually, these events led to India's Environment (Protection) Act (EA) in 1986, and the amendment of the Air Act in 1987. The EA empowered the central government to take steps to protect and improve the environment. More specifically, it issued comprehensive guidelines to lay down air quality stan-

dards, procedures and safeguards wherever needed, and take necessary steps towards prevention and abatement of air pollution.

Among the many environmental regulations that came into effect over this period, this paper specifically focuses on two main policies: the Supreme Court Action Plan (SCAP) and the catalytic converter (CC) regulations. They were implemented at various points in time in different cities that were considered to be critically polluted, which enables the quasi-experimental differences-in-difference design.

The SCAP was implemented in the city of Delhi in 1996, when in response to Supreme Court orders, Delhi's state government as well as the Central government developed comprehensive action plans to reduce air pollution. Based on a list of critically polluted cities identified by the Central Pollution Control Board (CPCB), the action plans were further expanded to other cities starting in 2003. The SCAPs consist of an array of policies that involve fuel regulation, building new roads, adopting cleaner technology and curbing industrial pollution. The Supreme Court mandated that all the cities that it identified had to develop action plans to reduce levels of respired suspended particulate matter. It included closing down certain factories, phasing out older vehicles, and implementing specific environmental standards. For instance, in the city of Delhi, all commercial passenger vehicles—buses, three-wheelers, and taxis—had to convert from diesel to compressed natural gas (CNG). This CNG action plan is said to have led to a significant improvement in Delhi's air quality (Narain and Krupnick, 2007).

The CC policy required that all vehicles should install a catalytic converter, an exhaust system device that reduces emissions of harmful pollutants. This regulation is aimed at reducing TSP, NO_2, and SO_2 levels. It was first proposed for Delhi by the Ministry of Environment and Forests (MoEF), and it is widely believed that the Supreme Court of India prodded the MoEF to do so (Narain and Bell, 2005). At the same time, the Petroleum Ministry banned the registration of new vehicles without a catalytic converter in all 4 major metropolitan cities, Delhi, Mumbai (previously, Bombay), Chennai (previously, Madras) and Kolkata (previously, Calcutta). This was then extended to more cities in a phased manner (see Table 3.1 for details). From 1999 onwards, India started following the standards based on European regulations on motor vehicle pollution, called the Bharat stage emission standards.

The Indian regulations, however, continue to target TSP/PM_{10} even after evidence showed more harmful effects of $PM_{2.5}$ relative to other pollutants, and even as the U.S. EPA promulgated its $PM_{2.5}$ rule as early as 1997. Only recently, in India's National Clean Air Program (NCAP) in 2014, was there a proposal to increase $PM_{2.5}$ monitoring across India.

3.3 Data and Sources

One of the main contributions of this paper is the compilation of extensive data for India, spanning 21 years. Using this dataset, which covers 117 districts in India over the years 1987-2007, we seek to understand the effect of regulations and ambient air pollution on mortality in urban India. This section describes in detail the data used in this paper, as well as their sources.

3.3.1 Regulation Data

Following several reports and litigations (Meagher, 1990), the Supreme Court of India instituted regulations to reduce pollution across India. It identified "critically polluted cities" and mandated that they take steps (in the form of SCAP and CC) to improve their air quality. It started with the capital city of Delhi in the mid-1990s, and other cities were phased in over the next several years.

Table 3.1 shows the policies in effect in the various Indian cities over the 21-year period. This dataset is a district-level panel for the years 1987-2007 across the urban areas of 117 districts in India. These data were collected by Greenstone and Hanna (2014) from the Department of Road Transport and Highways, the Ministry of Environment and Forests, and the State Pollution Control Boards (SPCB), set up across the states in India. The catalytic converter policy was first implemented in four major cities in India in 1995. In 1998, this policy was extended to the cities in 27 other districts. The SCAPs was first implemented in the city of Delhi and later rolled out through a 10-year period in phases across the various cities.

3.3.2 Mortality for All Ages and All Causes

This paper uses urban mortality data from the Census of India to understand the impact of a reduction in pollution on the mortality rate, which is a measure of public health. It focuses on urban mortality because the environmental regulations it studies have mainly targeted cities.

We integrate two systems that collect and estimate the mortality rate in India. The Civil Registration System (CRS) and SRS under the Census are the main sources of data for all vital events in India. The legal requirements of India mandate that every birth and death be recorded through the CRS. It is well known, however, that the data on registered deaths (i.e., CMR) from the CRS are under-reported (Mahapatra et al., 2007). In recognition of this measurement error, the Census uses SRS, a representative sample, to provide statistical *estimates* of annual mortality rates [7], at the state

[7]The Office of the Registrar General, Ministry of Home Affairs, publishes annual reports, "Vital Statistics of India

Table 3.1: Air Pollution Policy History

Year	Districts with TSP Monitors [1]	Districts with $PM_{2.5}$ Data [2]	SCAP [3]	CC Policy [4]
1988	21	*NA*	0	0
1989	28	*NA*	0	0
1990	40	*NA*	0	0
1991	43	*NA*	0	0
1992	52	*NA*	0	0
1993	60	*NA*	0	0
1994	51	*NA*	0	0
1995	36	*NA*	0	4
1996	63	*NA*	0	4
1997	68	*NA*	1	4
1998	52	117	1	31
1999	56	117	1	31
2000	52	117	1	31
2001	42	117	1	31
2002	50	117	1	31
2003	58	117	12	31
2004	61	117	17	31
2005	72	117	17	31
2006	87	117	17	31
2007	94	117	17	31

Sources: TSP (column 1) and regulation data (column 3 and 4) from Greenstone and Hanna (2014) as collected from the various pollution control boards across India; $PM_{2.5}$ data (column 2) from satellite observations based on van Donkelaar et al. (2016)
Notes:
(1) Column 3 is the number of districts that the SCAP was in place in that year and column 4 is the number of districts that Catalytic Converter policy was in place.

level through continuous enumeration and field verification of finalized deaths.

Using this estimated state-level mortality rate, we derive the urban registration percentage for each state—i.e., the percentage of deaths in urban areas reported in the CRS—which describes the extent to which the data are complete and represents all the deaths in the area. The Census has improved its recording of deaths; the urban registration percentage increased from 71% to 91% in our study period (Figure 3.1 shows that the two measures are converging).

Dividing the district-level urban CMR from the CRS by the urban registration percentage at the state level, we calculated what we call the urban adjusted mortality rates (AMR) at the district level. We used this AMR as one of the main dependent variable to measure the impact of regulations and air pollution on health. We also, however, ran corresponding regressions with the urban CMR as the dependent variable, which served as robustness checks. The AMR provides a more consistent and reliable measure of mortality rates across years and districts in India since it accounts for measure-

based on the Civil Registration System" with all the vital statistics for India.

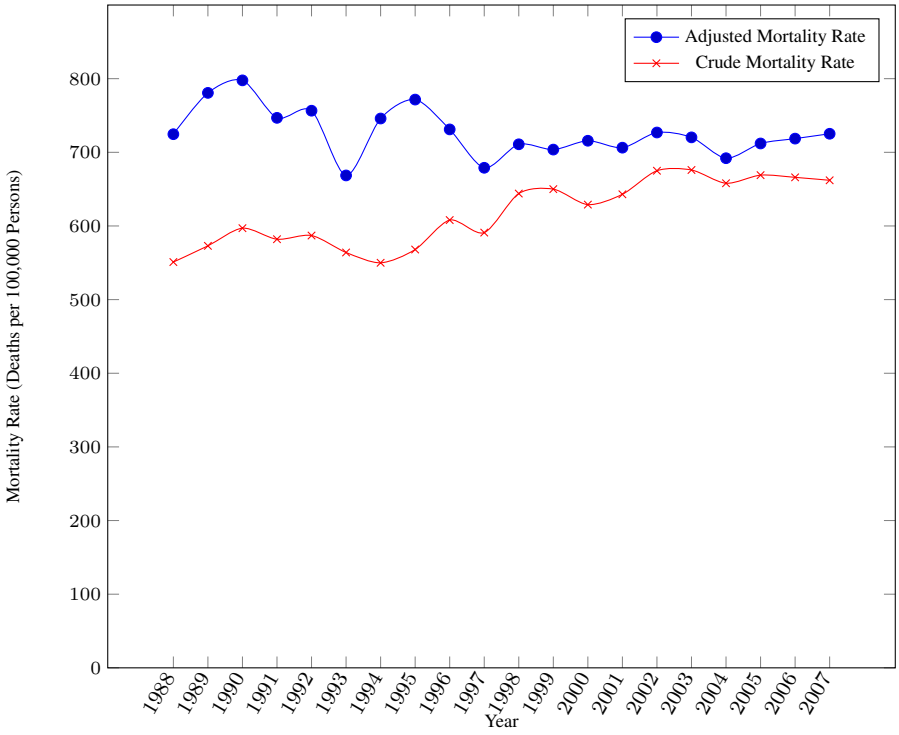


Figure 3.1: Adjusted Mortality versus Crude Mortality Rate

ment errors in the CMR, as is evident in Figure 3.1. Hence, the results from the regressions using urban AMR as the dependent variable are our preferred estimates.

Although the Indian government has made a concerted effort to collect better data over time (which is reflected in the increasing registration percentage in most parts of India), these percentages vary widely across the nation. The Central zone[8] has the highest mortality rate (adjusted for urban registration percentage), at 812 deaths per 100,000 persons, and also the lowest urban registration percentage, at 66%. The most complete data (as indicated by the registration rate) are in the Western zone[9], with a 98% urban registration percentage and a mortality rate of 632. The North-Eastern zone[10] has the lowest mortality rate, of 486, with a urban registration percentage of around 75%.

[8]The Central zone consists of the states, Madhya Pradesh and Uttar Pradesh

[9]The Western zone consists of the states, Daman and Diu, Goa, Gujarat, Maharashtra, Dadra & Nagar Haveli

[10]The North-Eastern zone consists of the states, Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland and Tripura

3.3.3 Pollution Data

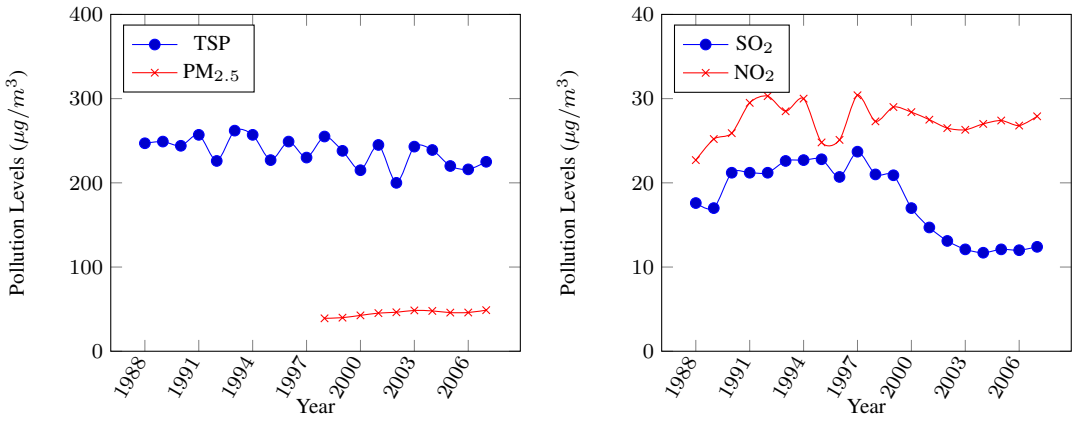
Particulate matter (PM) is a mixture of solid particles and liquid droplets found in the air. Total Suspended Particles (less than 100 micrometers in diameter), or TSP, includes dust, dirt, soot, or smoke, that is large or dark enough to be seen by the naked eye. Smaller particulate matter, such as PM_{10} and $PM_{2.5}$, i.e., particles less than 10 and 2.5 micrometers in size, are invisible to the naked eye and are also the most dangerous fraction of PM.

In our study, various measures of pollution were collected from different sources to obtain a robust estimate of the effect of pollution on mortality (see Figure 3.2). We used the pollution data for SO_2, NO_2 and TSP from Greenstone and Hanna (2014), while we obtained $PM_{2.5}$ data from the recently published satellite data from the Atmospheric Composition Analysis Group at Dalhousie University (van Donkelaar et al., 2016) to compare the effects of different types of pollution on mortality. This is inspired by papers, such as Pope et al. (2002), that have compared the impact of different pollution measures on public health, and found, using the U.S. data, that $PM_{2.5}$ is more dangerous to human health than TSP, SO_2 and NO_2 (among other pollutants). But so far few have measured the impact of fine particulate matter for India, where pollution levels are generally at higher levels (see Table 3.2).

The data for the analyses in this chapter end in 2007 because a subset of our data were obtained from Greenstone and Hanna (2014), which end in that year. Moreover, we could not extend our data beyond 2007 because the number of cities monitoring TSP declined dramatically after 2009, as monitoring data from MoEF and CPCB show. Many of these monitoring stations were closed or shifted, and by 2012, TSP was not monitored at all because this air quality measure was replaced by PM_{10} with data beginning to appear from 2008. In 2014, monitoring stations that collected data for $PM_{2.5}$ also started to appear in select areas.

Our innovation here is that we use data on $PM_{2.5}$ (albeit for a shorter time frame, 1998-2007) for India and compare its impact with that of other pollution measures. To create the district-level data on $PM_{2.5}$ levels, we collected the data that are based on van Donkelaar et al. (2016) for the cities included in Greenstone and Hanna (2014) and averaged them at the district level, using population weights. The data is for the years 1998 to 2007, spanning a shorter time period than the data for the other pollution variables and the rest of the data that cover the entire 21 years. We use this smaller dataset for the latter half of the paper, where we estimate the relationship between pollution levels and mortality in the fixed-effects design.

For the impact of other types of pollution, we use the data as collected by Greenstone and Hanna (2014) on TSP, SO_2 and NO_2 levels. While TSP is considered as a general indicator of pollution that arises from industrial and vehicular pollution as well as fossil fuel burning, SO_2 emissions



(a) TSP versus $PM_{2.5}$

(b) SO_2 versus NO_2

Notes:
(1) $PM_{2.5}$ levels are drawn from satellite data. (2) Pollution data for TSP, SO_2, and NO_2 from Greenstone and Hanna are drawn from the various pollution control boards across the country for the 21 years reported in the dataset. (3) The figures depict annual population weighted mean pollution levels. Annual means are taken across all monitors in a given district, and then averaged across all districts in a given year.

Figure 3.2: Trends Across Pollution Types

are mostly from thermal power generation, and NO_2 emissions are mostly attributable to vehicular pollution.

These pollution variables are monitored by the CPCB and the SPCBs. They have an extensive network of environmental monitoring stations across India that offers a rich dataset containing data on TSP, SO_2 and NO_2. The full dataset includes 572 monitors across 117 districts over the 21 years.

Panels (a) and (b) in Figure 3.2 show the trends of all the pollution variables. While average (population-weighted) $PM_{2.5}$ levels increased by 25% since 1998, TSP decreased by 12% (and by 9% since 1987). The levels of NO_2 and SO_2 went in opposite directions in those 2 decades, NO_2 increasing by 23% and SO_2 decreasing by 29%. The rising NO_2 levels might be because the number of vehicles in India has increased almost 3-fold in the last 21 years, and the falling SO_2 levels could be due to the many regulations imposed on hazardous industries.

The trends in Figure 3.2 (panel (a)) show that the levels of TSP behave somewhat differently than $PM_{2.5}$, with a correlation of 0.7. This is because TSP levels that include coarse particles can vary quite a bit as compared to $PM_{2.5}$: While TSP is largely affected by the temporal variability in dust levels and other ground material, variations in $PM_{2.5}$ levels are caused primarily by combustion processes and other secondary atmospheric chemistry (Hinds, 1999). Hence, the processes that give rise to TSP are largely unrelated to $PM_{2.5}$ emissions, leading to different trends.

Table 3.2: Summary Statistics

Measures	Urban Adjusted Mortality Rate	Urban Crude Mortality Rate	TSP	SO_2	NO_2	$PM_{2.5}$
Average	727	617	235	17	28	45
Standard Deviation	396	276	117	14	18	24
10th Percentile	422	331	88	5	10	22
90th Percentile	1,036	915	396	33	48	84
Number of Observations	1,948	1,952	1,086	1,158	1,189	1,124

Sources: Census of India, Dalhousie University, and Greenstone and Hanna (2014) as collected from the various pollution control boards across India
Notes:
The values of all (except $PM_{2.5}$) the above descriptive statistics are for the years 1987 to 2007. The values for $PM_{2.5}$ alone are for the years 1998 to 2007 because those are the years for which data exists. All the pollution measures are in $\mu g/m^3$. The mortality rate measures are in terms of the number of deaths per 100,000 persons. The averages are all population-weighted averages.

3.3.4 Demographic and Institutional Characteristics

To control for health and demographic information that could potentially affect mortality, we collected data from the Census of India, the National Family Health Survey and the Medical Council of India.

Literacy is associated with higher utilization of health services and better health (Nag, 1983). To control for this, we collected literacy rates for urban areas at the state level from the Census of India from 4 census years, 1981, 1991, 2001, and 2011. For the years in between, we interpolated the numbers to get the values for all 21 years in the dataset. Government data show that urban literacy has increased tremendously in India, increasing from an average of 62% in 1988 across the districts we are interested in, to 73% in 2007.

Since age is related to mortality, we also control for age in the model. The data for age is collected from (International, 2018), which compiles demographic data from the Census of India (among other countries of the world).

We have also included the number of doctors registered in the state per million persons as an indicator of access to healthcare. The state-wise distribution of the number of doctors registered is available with the Medical Council of India, a statutory body which mandates that all physicians register with them. The number of doctors registered has been increasing steadily, and almost doubled in proportion to the population from 1987 to 2007, increasing from 384 to 631 doctors per million people.

3.4 Impact of Policies to Combat Pollution: Framework and Results

In contrast to many of the epidemiological studies, the literature in economics is concerned about endogeneity and cumulative exposure that might lead to bias in the estimates of the linkage between pollution and human health. To estimate the causal impact in such situations, Autor (2003) and others came up with a differences-in-difference design, comparing the changes in the variables of interest (mortality and pollution levels in our paper) due to policy.

Greenstone and Hanna (2014) was the first papers to evaluate, using a differences-in-difference model, the impact of the Indian environmental policies (although estimating the impact on infants). Combining elements of the methodology used in Greenstone and Hanna (2014) with that of Autor (2003) and Glewwe and Todd (forthcoming), we estimate the impact of pollution-reducing policies on mortality in India. The policies were implemented in a phased manner across the cities in India, which makes it ideal for the differences-in-difference method.

3.4.1 Do Environmental Regulations Cause a Decline in Mortality?

The trends in nationwide mortality and the number of districts in India that adopted one or the other environmental regulation (i.e., the SCAP or the CC policy) show a negative correlation of -0.7 (see Figure 3.3). In other words, as more districts adopted environmental regulations, CMR and AMR declined. This correlation, however, does not imply causation. Since the policies have been adopted at different times in different districts, we can use the different mortality rates in adopting and non-adopting districts to identify the effect of the policies by using the differences-in-difference model (Autor, 2003).

Following the empirical framework described above, we attempt to estimate the effect of air pollution policies on mortality rate in the urban areas of India. Given that there are differences across the districts in mortality trends even before the policies were implemented, the econometric model that we follow in this paper is as follows:

$$\ln(\text{AMR}_{ct}) = \alpha + \theta \text{SCAP}_{ct} + \beta \text{CC}_{ct} + \delta X_{ct} + \mu_t + \gamma_c + \lambda_c \times t + \eta_c \times t^2 + \varepsilon_{ct} \tag{3.1}$$

where $\ln(\text{AMR}_{ct})$ represents the logarithm of adjusted mortality rates in district c and at time t. The policy dummies, SCAP and CC, are used to identify the effect on mortality. In addition to these policy variables, we also use district fixed effects, γ_c, that control for mean differences in mortality across the districts irrespective of the impact of policy, and year fixed effects, μ_t, that control for any national fluctuations over time that have the same effects on all districts. In order to include

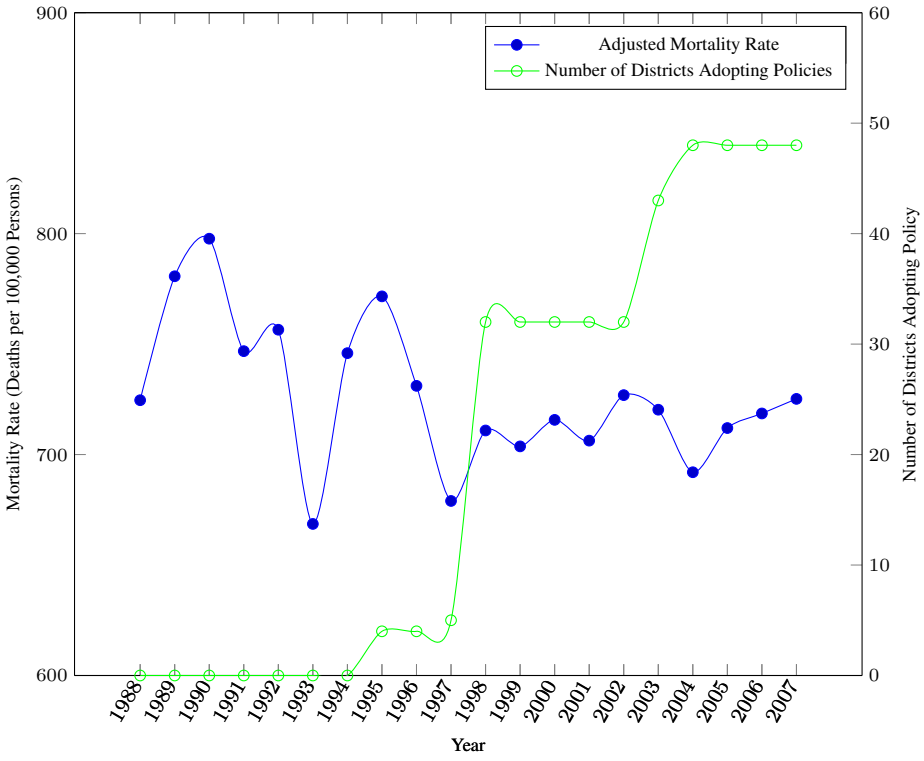


Figure 3.3: Adjusted Mortality versus Policy Adoption

Source: Census of India; Cities adopting from Greenstone and Hanna (2014)

systematic trends in the districts, we have also included district-specific linear and quadratic trends as represented by $\lambda_c \times t$ and $\eta_c \times t^2$. To verify that they make a difference to the model, we ran an F-test to see whether district trends are jointly zero, and we found that they are significantly different from zero at the 1% significance level. Following Greenstone and Hanna (2014) and many others, we have also included several control variables (X_{ct})— urban literacy rate, average age by state, and availability of doctors—to control for any omitted variable bias due to unobserved heterogeneity that is not captured by district and time fixed effects. All models use standard errors (SEs) clustered at the district level to correct for error correlation within the districts over time, given that there may be similarities within certain regions of the country.

The results of the model assume of course, that the policies that are instituted are independent of the residual term ε_{ct}, and that their effects do not spillover into non-adopting districts. However, given the nature of the legal and executive machinery in India, with many of these policies stemming from Supreme Court orders that came from public interest litigations, the timing of these policies may not be anticipated.

The coefficient estimates of the impacts of the policies will be unbiased only if the parallel trends assumption holds. By contrast, they could be biased if there are pre-existing trends in the districts that had policy intervention that differ from the pre-existing trends of the districts that did not. We run an event study analysis to formally test whether pre-trends in the Indian data are a concern. To do that, we estimate the following equation:

$$\ln(\text{AMR}_{ct}) = \alpha + \sum_{j=-5}^{5} \beta_j(\tau_t = j) + \delta X_{ct} + \mu_t + \gamma_c + \lambda_c \times t + \eta_c \times t^2 + \varepsilon_{ct} \tag{3.2}$$

where τ_t denotes the time relative to when SCAP was implemented such that $\tau_t = 0$ is the year when the policy was implemented in that district. The coefficients of the equation represent the changes in $\ln$(AMR) relative to the year, $\tau_t = -1$. We chose lags from -5 to +5 to illustrate how the pre-trends affect the outcome of interest.

We see that the coefficients from the estimation of Equation (3.2) are significant (Table D2 in the appendix) showing that there are indeed pre-trends. In contrast, we find that there are no pre-trends in the case of CC policy (see Table D3 in the appendix). These results show that our study is not likely to be biased for the CC policy, but there could be bias in the estimates of the SCAP. To further investigate whether that this result with regard to SCAP is an issue, we ran a joint significance test to see if all the pre-trends together are significantly different from zero. Our result show that we can reject the null that there are no pre-trends.

To account for pre-trends in SCAP, we modify the differences-in-difference model as described in Equation (3.1) to include separate fixed effects for districts that have the SCAP and CC policies in addition to the time trends for all the districts in the fixed effects, as recommended by Glewwe and Todd (forthcoming). We use this modified model to estimate the impact of the regulations on mortality rate (as is shown in Equation (3.3)).

$$\begin{aligned}\ln(\text{AMR}_{ct}) = \alpha + \theta\text{SCAP}_{ct} + \theta_1\text{SCAP}_{ct} \times t + \beta\text{CC}_{ct} + \beta_1\text{CC}_{ct} \times t + \delta X_{ct} + \mu_t + \kappa_t \times \text{I}_{SCAP,c} \\ + \tau_t \times \text{I}_{CC,c} + \gamma_c + \lambda_c \times t + \eta_c \times t^2 + \varepsilon_{ct}\end{aligned} \tag{3.3}$$

where SCAP_{ct} and CC_{ct} are the policy variables corresponding to SCAP and CC policies. They are the interactions between the districts that have ever had the policies with year that the policies went into effect as well as a dummy variable that has data ranging from -7 to 3 years before and after SCAP policy and -7 to 9 for CC policy. This is to ensure there are districts that have a minimum number of years pre-and post-implementation, and as recommended by Greenstone and Hanna (2014). The identification is based on different start times for different program districts. We also control for other variables, urban literacy rate, number of doctors per person, average age of a

person in the state, district and time trends.

The main coefficients of interest are θ and β. The indicator variables, $\mathrm{I}_{SCAP,c}$ and $\mathrm{I}_{CC,c}$, specify whether the district has ever had SCAP or CC respectively. These 2 variables do not depend on when the policies were enacted in the districts. The separate trends for the districts that ever had the SCAP policy or the CC policy are represented by $\kappa_t \times \mathrm{I}_{SCAP,c}$ and $\tau_t \times \mathrm{I}_{CC,c}$. These are separate effects in each year for those districts that were eventually treated and include the effects for years that precede the policy enactment as well as after. These SCAP and CC district-year fixed effects will account for the pre-trends that we find in the event study analysis described above. We include CC policy specific fixed effects in the model even though there are no pre-trends associated with this policy so that we have the most robust specification. In addition to the separate district-year fixed effects, we also have district-specific linear and quadratic trends, which takes care of the fact that different districts might have different trends (linear or quadratic). All of these fixed effects allows us to use this difference-in-differences setup even when the parallel trends assumption does not hold.

Table 3.3 shows different combinations of fixed effects when estimating the impact on AMR. We see that when there are no district-specific linear or quadratic trends, as in model [5], the effects of SCAP and CC policies are not significant. Including the district-specific linear and quadratic trends and district-year fixed effects for those districts that have ever had the policies changes the results. Results from model [6] in Table 3.3 show that the adoption of the CC by a district leads to a 19% reduction in AMR that is significant at 5% level. This effect declines as time goes on, with an increase in AMR by 4.4% every year from the second year of implementation. The impact of SCAP, however, is not significant.

Considering that the estimated model [6] from Table 3.3 accounts for non-parallel trends and policy district-specific time trends, we have decided to count these results as reflective of the policies' impacts. An F-test on the hypothesis that the district trends are jointly zero is rejected at 1% significance level, further justifying this model. An F-test also shows that the district year fixed effects are jointly significant at the 5% level.

To confirm that the model [6] in Table 3.3 is effective in negating the pre-trends we saw earlier, i.e., the coefficients in this model do indeed reflect the policies' impacts, we also run an event study analysis with the specification in model [6]. The results of the event study are displayed in the Tables D4 and D5, and both of the tables confirm the results we obtained in Table 3.3. The estimated coefficients on the lags for SCAP and CC show that there are no effects in the five years before these policies were implemented.

For the 5 years after the policies were implemented, we find that SCAP does not impact mor-

Table 3.3: Effect of Policy on Adjusted Mortality Rate

	[1]	[2]	[3]	[4]	[5]	[6]
Supreme Court Action Plan (SCAP)	0.028	0.081			0.040	0.084
	(0.065)	(0.057)			(0.065)	(0.055)
SCAP X Linear Trend	-0.040	-0.011			-0.044	-0.018
	(0.033)	(0.028)			(0.034)	(0.029)
Catalytic Converter Policy (CC)			-0.074	-0.221**	-0.079	-0.209**
			(0.058)	(0.095)	(0.058)	(0.092)
CC X Linear Trend			0.004	0.042*	0.006	0.043**
			(0.012)	(0.022)	(0.012)	(0.021)
Observations	1739	1739	1739	1739	1739	1739
Adjusted R^2	0.023	0.449	0.024	0.454	0.024	0.454
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	No	Yes	No	Yes	No	Yes
District-Specific Quadratic Trends	No	Yes	No	Yes	No	Yes
SCAP District-Year Fixed Effects	No	Yes	No	Yes	No	Yes
CC District-Year Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

tality rate whereas CC does impact the mortality rate, which is consistent with Table 3.3. The coefficients in Table D4 show that the effect for SCAP is significant only in the year of implementation at 10% level. Thereafter, in the four years after the first year of implementation, the effect is insignificant. When we did a joint F-test of the leads for SCAP, we found that the effect of SCAP is indeed insignificant, which is similar to what we see in model [6] in Table 3.3. The coefficients in Table D5 shows, however, that the implementation of CC reduces AMR sharply in initially, which then flattens out at the lower level of AMR for the four years after the first year of implementation. A joint F-test of the years post implementation of the CC policy shows that it has lead to a significant impact on AMR, which is consistent with the 19% decline in AMR as shown in Table 3.3.

As a robustness check, we run these models using the CMR (i.e., the unadjusted version of mortality) as the dependent variable. Estimates in Table D1 from the appendix show that the impact of policy on CMR is very similar, with a 21% decline in CMR. These results are unlike some earlier papers (Chay et al., 2003) that show that there is no significant impact on adult mortality.

Since the policies focused on cities that were most polluted, it is quite possible that gains are disproportionately bigger in these cities than others. For instance, the city of Delhi was the first city to experience these policies. In spite of using fixed effects and robust SEs, this heterogeneity in the effectiveness of policy could potentially affect the precision of the coefficient estimates. To test this,

we interacted the pollution levels before the policies were enacted with the policy dummies to see if this affected AMRs. From Table D6 in the appendix, we see that almost none of these interaction variables impact mortality.

Additionally, since Delhi is much more polluted than the rest of the districts of the country, we wanted to rule out that this outlier is biasing the coefficient estimates. Results from not including Delhi in the analysis (see Table D7 in the appendix) show that the coefficients are no different.

It seems from the foregoing results that the impact of the Indian environmental regulations are significant in reducing mortality in these districts. Specifically, results show that CC policy is effective in reducing mortality. Note that Greenstone and Hanna (2014) showed that the CC policy is the most effective policy when it comes to reducing pollution levels.

3.5 Do Mortality Rates Rise Due to Elevated Air Pollution?

One of the major justifications for environmental regulations in India is the effect that pollution exposure has on human health (National Clean Air Programme, 2015). Our results above (see the above Section 3.4.1) showed that the Indian regulations have significantly reduced mortality, and Greenstone and Hanna (2014) show that these regulations have been effective at reducing pollution levels. The Indian pollution levels, however, remain relatively high as compared to the other parts of the world (Brauer et al., 2016).

To design future policy that is more efficacious, its may be necessary to measure the fundamental relationship of different pollution types directly on mortality. When the Indian authorities first implemented the air quality standards, they aimed at reducing TSP and later PM_{10} levels, consistent with regulations across the world at that time. However, more recent epidemiological papers (Pope et al., 2002) demonstrate that $PM_{2.5}$ is the more dangerous form of ambient air pollution. Also, Economists such as Muller and Mendelsohn (2009) have shown that the highest damage in dollar terms is due to $PM_{2.5}$ among all the pollutants. Only recently, in 2014, did India's National Clean Air Program (NCAP) proposed to increase $PM_{2.5}$ monitoring across India in recognition of the harmful effects of $PM_{2.5}$ over other pollutants.

In 2015, the NCAP expressed the need to understand the relationship between various pollutants and public health specifically using Indian data (and not just extrapolating from international studies), in order to improve the database needed to strengthen efforts for cleaner air in the nation (National Clean Air Programme, 2015). Therefore, we ask in our paper, are elevated air pollution levels a cause for concern, leading to increased mortality? We estimate this impact for four different

pollutants: $PM_{2.5}$, TSP, NO_2 and SO_2. These causal relationships, if established, might encourage the more polluted cities to aim for cleaner air so that they could reduce pollution-related deaths.

We address this objective in our paper by using a fixed effects model to determine the impact of the different pollution measures directly on mortality. We tackle the many endogeneity issues that can hinder us from estimating the impact of pollution in these models, so that we can truly understand the magnitude of the effect of the pollutants.

The equation we use to understand this relationship is as follows:

$$\ln(\text{AMR}_{ct}) = \alpha + \delta \ln(Y_{ct}) + \beta X_{ct} + \mu_t + \gamma_c + \lambda_c \times t + \eta_c \times t^2 + \varepsilon_{ct} \tag{3.4}$$

We use the log-log specification to approximate the specification used by Krewski et al. (2009) and Burnett et al. (2018), both of whom showed that mortality increases at a decreasing rate with an increase in pollution. The treatment variable, $\ln(Y)$, is in this case any one of the measures of district-level ambient air pollution for that year. The district-level fixed effects are represented by γ_c and the time fixed effects by μ_t. Bertrand et al. (2004) showed us that district level differences-in-difference model are often affected by serial correlation that could bias the estimates of standard errors (SE) downwards. Hence, we have used SEs clustered at district level in all the models, which are robust to any arbitrary error auto-correlation of the error term within the districts.

Over the 21 years in the dataset, NO_2 increased by 23% (Figure 3.2), and $PM_{2.5}$ rose by 25% in the last 10 years. By contrast, TSP and SO_2 declined by 9% and 29%, respectively, in this 21 year period, with AMR increasing by 5%. The question then becomes: how are these changes in pollution concentrations related to the change in mortality rates? To investigate this relationship more thoroughly, we use the district-level urban AMR for the years 1987-2007, and the four pollution measures (TSP, $PM_{2.5}$, SO_2, and NO_2) in a fixed effects model.

It must be noted that air pollution concentrations depend on factors such as population density and economic activity in the area among others. To account for the omitted variable bias that could potentially affect mortality, the variable X represents the matrix of control variables. These variables include urban literacy rate of each state, the number of doctors registered per person within each state, and the average age of each state. All these variables vary across the time span.

Our aim in this section is twofold: One is to measure the effect of an increase in ambient air pollution on mortality after addressing issues of endogeneity, and the second is to investigate whether one type of pollution is more dangerous than the others in increasing mortality rates, i.e., comparing the impacts of $PM_{2.5}$, TSP, SO_2 and NO_2 on mortality rates.

Many papers based on U.S. data (e.g., Pope et al., 2002, Lepeule et al., 2012 and others) show

that an increase in $PM_{2.5}$ levels led to an increase in mortality. Pope et al. (2002)) went even further to show that $PM_{2.5}$ levels are more harmful than other types of pollution. Comparing studies by Krewski et al. (2009) and Lepeule et al. (2012) with Woodruff et al. (1997) shows that the impact of pollution on adults is manifold as compared to that of infants. Because these studies are all based on the U.S., it becomes all the more pertinent to ask whether elevated ambient air pollution levels (of different forms) lead to increased mortality rates in the case of India, which has much more highly polluted cities.

In estimating the coefficient δ, one must account for the fact that several factors other than environmental pollution, including average age, access to medical facilities and economic variables, can affect mortality. In addition, Y may not be entirely exogenous. Hence, identifying the effect of ambient air pollution could be complicated. Our identification strategy explaining how we dealt with the issues of reverse causality, unobserved heterogeneity, measurement error, violation of the Stable Unit Treatment Value Assumption, and omitted variable bias are described in detail in Section C.3 in the appendix.

3.5.1 Results

As per the identification strategy, we have estimated the model from Equation (3.4) for the four types of pollution measures that we have listed above, i.e., $PM_{2.5}$, TSP, SO_2, and NO_2.

Estimates of model [4] from Table 3.4 show that as $PM_{2.5}$ rises, mortality increases. This is after we have controlled for a variety of variables: urban literacy rate within each district, the density of doctors in the state, and the average age of the people in the states, in addition to the fixed effects. The estimates show that a 10% increase in the concentration of $PM_{2.5}$ leads to an increase of 2.54% in AMR.

By contrast, TSP does not have a significant impact on AMR (see Table D8). The impact of the two pollution measures, $PM_{2.5}$ and TSP, are displayed side-by-side to compare their different effects on mortality (in Table 3.5). Since the satellite $PM_{2.5}$ data is available for a shorter time span, we also run the models with TSP using the shorter dataset in order to get a true comparison of estimates. Estimates from models [2] and [3] in Table 3.5 show that the relationship between AMR and TSP is not significant, for either time period, indicating that TSP is not as dangerous as exposure to $PM_{2.5}$. The differences in the results could also be reflective of the fact the two types of pollution are collected from two different sources, where the TSP data is based entirely on monitor data across India that does not cover all locations at all times as satellite data do.

As a robustness check, we also ran the regressions for the impact of $PM_{2.5}$ and TSP on mortality

Table 3.4: Effect of Particulate Matter 2.5 Microns or Less ($PM_{2.5}$) Levels on Adjusted Mortality Rate

	[1]	**[2]**	**[3]**	**[4]**
$\ln(PM_{2.5})$	0.102	0.184*	0.244*	0.254*
	(0.098)	(0.109)	(0.136)	(0.146)
Observations	836	836	836	836
Adjusted R^2	0.019	0.031	0.183	0.409
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes
District-Specific Linear Trends	No	No	Yes	Yes
District-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

for the CMR measure. Estimates from Table D9 in the appendix show that increased $PM_{2.5}$ levels lead to a significant increase in mortality rate for both the adjusted and crude mortality measures, whereas the impact of TSP levels on mortality is insignificant.

In addition to TSP and $PM_{2.5}$, we also estimate the effects of SO_2 and NO_2 where we find that the impact of SO_2 is positive and significant at a 10% level with the specification in model [3] in Table D10. The impact of SO_2 on mortality is 0.4% for every 10% increase in SO_2 levels. This statistically significant result disappears when district fixed effects are interacted with quadratic time trends in the model, as shown in model [4] in this table. From Table D11 in the appendix, we see that the impact of NO_2 is insignificant with all the specifications in the table.

We also consider the possibility that including just one pollutant, while not including another in the same regression could potentially cause omitted variable bias. To check if this might be true, we estimate the impact of pairs of pollutants on mortality. The results of these models are shown in Table D12. In examining these results, we that while the impact of $PM_{2.5}$ is similar in magnitude and direction as in Model [4] of Table 3.4 all the estimates of the pollutants are not significant from zero. Hence, we have chosen to estimate each pollutant's effect separately.

The results in Table 3.5 show that the effect of $PM_{2.5}$ on mortality in India is positive and significant. These results are in line with other literature, for other regions of the world, which conclude that $PM_{2.5}$ is more dangerous and a greater risk to to human health than TSP (Pope et al., 2002, Dockery et al., 1993). Hence, $PM_{2.5}$ concentrations are a more reliable measure for the impacts on mortality and make for a more compelling target for abatement policies than TSP. These results are reinforced when the regressions are run with CMR as dependent variable (see Table D9).

This result is one of our main contributions, in that it highlights the importance of $PM_{2.5}$ and its

Table 3.5: Effect of $PM_{2.5}$ versus TSP on Mortality Rate

Dependent Variable:	ln(Urban AMR)		
	1998-2007 data		1988-2007 data
	[1]	**[2]**	**[3]**
ln($PM_{2.5}$)	0.254*		
	(0.146)		
ln(TSP)		0.027	0.010
		(0.022)	(0.024)
Observations	836	456	809
Adjusted R^2	0.409	0.374	0.566
District Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes
District-Specific Quadratic Trends	Yes	Yes	Yes

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Models [1] and [2] are run for only a subset of the data, i.e., for the years 1998-2007. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

impact on mortality. Much of the literature in economics focuses on the effect of pollution on infant mortality. We use infant mortality data from Greenstone and Hanna (2014) to estimate the impact of both TSP and $PM_{2.5}$. We find that the effect of both pollution measures on infants are insignificant (see Table D13 in the appendix). This result is consistent with Woodruff et al. (1997) that estimated that impact of pollution on infants is much smaller than that of mortality for all ages.

Given the magnitude of the effect of pollution on mortality for all ages as compared to just infants, we highlight two important aspects of our research: The level of $PM_{2.5}$ is more crucial for policy than TSP, and focusing on mortality for all ages rather than just infants is essential to public health.

3.6 Conclusion

Two essential elements concerning environmental regulations are discussed in this paper. First, we show the effect of the Indian regulations on AMR as the main variable of interest, unlike many papers in the economics literature, which focus on infant mortality as the outcome. In the second part of the paper, we analyze the impact of air pollution on AMR, the fundamental relationship that influences policy in the first place.

The effect of the CC policy on AMR was significant in reducing mortality by 19%, while the impact of SCAP was small and insignificant. The question then becomes whether the Indian regulations can become more effective, given that India is still highly polluted. We address this question by analyzing whether exposure to one type of air pollutant may be more risky than others to human health. In other words, are elevated levels of a particular pollution type cause increased mortality rates for India? The motivation of instituting policy regulations could stem from this relationship if particulate matter does indeed have a significant impact on health, e.g. mortality. Other studies have estimated a large number of pollution-related deaths. For instance, in their Global Burden of Disease paper, Lim et al. (2012), estimate that 3.2 million deaths are related to $PM_{2.5}$ pollution worldwide. From the WHO report, the top 20 most polluted cities in the world include 13 Indian cities. This investigation forms the core of the second part of the paper, which is made even more informative because of the availability of new satellite data on $PM_{2.5}$.

The results from the model, answering the above question, tell us that while some pollution types are particularly dangerous to human life, while others appear to be less dangerous. The estimated response of the 10% increase in fine particulate matter, or $PM_{2.5}$, is a 2.4% increase in AMR that is significant at a 10% level. We find, however, that an increase in TSP, SO_2, or NO_2 levels does not significantly impact mortality rates. These results are robust across different measurements of mortality rate, showing us that $PM_{2.5}$ is the more dangerous fraction of particles as compared to TSP and it does have a significant impact on health. Our results indicate that an increase in $PM_{2.5}$ from 60 to 70 $\mu g/m^3$ leads to a 4% increase in mortality. Comparing this result with estimates from Table 11 in Krewski et al. (2009), we find that for the same increase in pollution, the increase in mortality is lower at 1.2%. This comparison shows that estimating relative risk of mortality using Indian data can provide new and valuable information.

The results in this paper suggest that the Indian regulations may have to shift their focus from TSP to $PM_{2.5}$ in order to have a substantive impact on human health. To understand the magnitude of the health impact or the lives saved from pollution reduction, further research, in the form of longitudinal cohort studies, should be conducted to better understand this relationship for India. The extent of the impact on public health can then be used to answer this question of how policy standards should be set. The Integrated Exposure Response model (Burnett et al., 2014) tries to answer this question by integrating many models in order to identify and develop the shape of the mortality relative risk functions. This question as it relates to India is still open for further research.

Chapter 4

Pollution Attributed Mortality in India[1]

Abstract

Chronic exposure to ambient air pollution has been shown to increase the risk of mortality. Most epidemiological studies in this area have described this association for the U.S. and European settings, where they have established that the quality of air is significantly correlated with human health even with pollution levels lesser than 30 $\mu g/m^3$. Few papers,however, have studied this relationship and the shape of the concentration response curve for developing countries, where pollution levels are much higher than those of the developed nations. This shape will determine whether the air in India, which is much more polluted than that of the developed nations, affects the health of its population at a higher, lower, or the same rate as the U.S. centric impact. Our paper improves upon previous research by measuring the impact of pollution exposure, following a rigorous identification strategy, on human health by using a panel dataset that includes $PM_{2.5}$ and mortality rates for 119 districts in India spanning 17 years. We use this relationship to arrive at relative risk estimates of the burden of disease at higher pollution concentrations. We show that our estimates of the lives saved from reducing pollution are very similar to those of the Global Exposure Mortality Model.

Keywords: $PM_{2.5}$, concentration response, premature mortality

[1] co-authored with Jay S. Coggins and Andrew L. Goodkind

4.1 Introduction

The World Health Organization (WHO) named 13 cities in India among the 20 most polluted cities in the world in June 2015. India's annual average $PM_{2.5}$, that is particulate matter that is 2.5 microns or lower in size, for 2015 is 74 $\mu g/m^3$, which is an order of magnitude higher than compared to that of the U.S., where the annual average for that year was 9.5 $\mu g/m^3$.

Chronic exposure to ambient air pollution is proven to increase the risk of mortality due to various nonaccidental causes (Pope et al., 1995, Dominici et al., 2002, Krewski et al., 2009, Lepeule et al., 2012). Most of these studies have described this association for the U.S. and European settings, where they have established that the quality of air is significant to human health even with pollution levels lesser than 30 $\mu g/m^3$. Few papers have, however, studied this relationship for developing countries where pollution levels are much higher than those of the developed nations. Fewer still have focused on the shape of the concentration response curve [2], which characterizes the relationship between pollution and health. This shape will determine whether the air in India, which is much more polluted than that of the developed nations, affects the health of its population at a higher, lower or at the same rate as the U.S. baseline rate.

The prevailing understanding of the relationship between air quality and human health is that each unit change in pollution levels will lead to the same incremental change in the mortality rate regardless of the baseline concentration levels. Recent studies have, however, shown that the concentration response may not change in constant increments, and it may be concave[3].

The implication of this shape is that an aggressive policy reduction in pollution levels will lead to a greater reduction in premature deaths. This means that, as nations become cleaner, the reduction in health risk becomes higher. So, if this shape is accurate, it provides a rationale for policies to continue aiming for cleaner air even when pollution levels are low. Also, the large number of people facing high pollution levels in developing nations means that even small marginal benefits from initial improvements in air quality can lead to a large number of lives saved for such societies. This substantial population exposure is a compelling reason to initiate aggressive pollution policies, and supralinearity supports further cleaning up. A supralinear concentration response, therefore, means that an aggressive policy to reduce pollution is beneficial to a nation at all pollution levels rather than cleaning the air only when it is the dirtiest.

Previous research has calculated the burden of disease attributed to pollution in highly polluted countries by simply extrapolating the results from countries that experience lower concentrations

[2]The concentration response curve is defined as the risk of mortality or mortality rate for a range of pollution exposure levels.

[3]Epidemiologists often call this shape "supralinear"

(Cohen et al., 2004), which could lead to an unrealistically high or low impact on mortality. In the Integrated Exposure-Response (IER) model (Burnett et al., 2014), efforts were made towards consolidating models that estimated the relative risk of mortality [4] due to sources high $PM_{2.5}$ concentrations other than vehicular and industrial pollution, such as active smoking, and second-hand tobacco smoking, so that they could arrive at concentration responses for higher ranges of pollution. They mainly considered five non-accidental causes that led to higher risks of mortality for all age groups. These relative risk estimates were then used to derive impacts for many countries including India and China, which showed that the impact of pollution flattened out at very high levels. More recently, Burnett et al. (2018) estimated the shape of the concentration response using only outdoor air pollution exposure from 41 cohorts from 16 countries to establish the Global Exposure Mortality Model (GEMM), which included all non-communicable diseases and lower respiratory infections for people over 25 years of age. One of the cohorts is based on the impact of long-term outdoor $PM_{2.5}$ exposure on Chinese men (Yin et al., 2018), extending the range of pollution to 84 $\mu g/m^3$, which is much higher than any pollution level experienced in the U.S. or Europe. Once again, the shape of the concentration response curve from GEMM showed that the relative risk of mortality from pollution increased at a decreasing rate with every increase in $PM_{2.5}$ levels.

Other papers (Krewski et al., 2009) have also shown that the relationship between human health and pollution might be "supralinear", i.e., a given increase in concentration levels would lead to smaller increases in health risk as the air get dirtier. Conversely, as the air gets cleaner, net benefits of abatement increases, which is contrary to traditional theory in environmental policy that suggests that the net benefits are highest in the dirtiest places. This means that it would take an aggressive policy and a large reduction in pollution in dirtier places to see a decline in relative risk of mortality in the beginning. However, as the air gets cleaner, the net benefits increase faster, giving greater encouragement to cleaner places such as the US or Europe to keep on pursuing clean air polices.

Our paper is one of the first studies that estimates the shape of the concentration response curve by estimating the impact of pollution levels on mortality using actual Indian data for $PM_{2.5}$ and mortality rates across 119 districts in India for the years 1998 to 2015, following a rigorous identification strategy. Using a fixed effects model and the theoretical foundations from the Burnett et al. (2018), Burnett et al. (2014) and (Krewski et al., 2009), we arrive at the functional form of the relationship between $PM_{2.5}$ and the mortality rate including all causes and ages. This is unlike many of the previous studies that simply extrapolated model coefficients using U.S. or European data. We use this relationship to arrive at IER-like relative risk estimates of the burden of disease at higher pollution concentrations. Our main contribution to the literature is to estimate this concentration-response curve for the first time for India using actual mortality rate data from India. In our results,

[4] The relative risk of mortality is defined as the risk of mortality at a particular concentration level divided by the risk of mortality at a counterfactual concentration level. In this paper, as in Krewski et al. (2009), the counterfactual concentration level is 5.8 $\mu g/m^3$.

we find that the shape and impact of our models are very similar to that of the GEMM.

We compare our concentration response curve with that of Krewski et al. (2009), which has a much smaller $PM_{2.5}$ range of less than 30 $\mu g/m^3$ as compared to our range of 16 to 136 $\mu g/m^3$, and GEMM. We find that the relative risks of the log-log model from Krewski et al. (2009) is lower than those of our models whereas the log-linear model are much higher than those for our models. By contrast, the estimated relative risk of mortality from pollution exposure from GEMM is similar our models in both shape as well as impact, and the number of lives saved from a pollution reduction falls well within the 95% confidence interval of GEMM.

The remainder of this paper proceeds as follows: Section 4.2 describes the data and their sources. Section 4.3 discusses relative risk of mortality from previous studies. Section 4.4 presents the empirical framework that also includes the theoretical framework of pollution abatement in Section 4.4.1, the empirical methodology in this paper in Section 4.4.2, followed by the results in section 4.4.4. Section 4.5 concludes.

4.2 Data and Sources

Our paper uses the urban mortality rate derived from India's Vital Statistical System that is generated through its Civil Registration System (CRS) and the Sample Registration Survey (SRS), two chief sources of vital records in India. Registration of deaths is carried out under the provisions of the Registration of Births and Deaths (RBD) Act, 1969, which requires all deaths to be registered by the local Registrar appointed by the State Governments.

We collected data on the registered number of deaths from the CRS for 119 districts (out of a total of 466 districts spread across 25 Indian states and 7 territories as per the 1991 Census of India), focusing on urban areas, for the years 1998-2015. Here, a district is an administrative division within an Indian state or territory. Using the Census population figures for these districts in the urban areas, we calculated the crude death rate (CDR) or the all-cause mortality for all ages (ACM). Note that a few district definitions have changed over the years; and for the purposes of this research, we used the state and the district definition as per the 1991 Census.

Although the RBD Act of 1969 has been amended several times since it was enacted to improve the system of registration as well as to bring about uniformity across the country, there are still substantial differences between the actual number of deaths and registered deaths, with actual deaths being much higher than what is registered (see Figure 4.1). The measure that determines how well these vital events are recorded is the level of registration, defined as the percentage of registered

Table 4.1: Summary Statistics

Variable	Number of Observations [1]	Mean [2]	Standard Deviation [3]	10th Percentile [4]	Median [5]	90th Percentile [6]
Panel A: Outcome:						
Urban Adjusted Mortality Rate	1856	684	257	407	659	975
Urban Crude Death Rate	1856	646	237	378	642	929
Panel B: Pollution Measure:						
$PM_{2.5}$ ($\mu g/m^3$)	1822	49	27	25	40	96
Panel C: Control Variable:						
Urban Literacy Rate (%)	1853	79	6	70	81	85

Sources: The urban AMR are from CRS data that are adjusted using SRS estimates; the urban CDR are from CRS data from the Census in India; $PM_{2.5}$ data are from satellite observations based on van Donkelaar et al. (2016); literacy rates are from the Census in India
Note: The averages for Urban Adjusted Mortality Rate and Urban Crude Death Rate listed above are the number of deaths per 100,000 persons.

deaths relative to the deaths as estimated through the SRS for each state in India. The SRS uses a representative sample of India to estimate and project the state-level death rate, which reflects the actual death rate. We get the registration rate, which reflects the coverage of all deaths in the country, by dividing the CDR by the estimated death rate from SRS. Since the registration rate for India is available only at the state level, we use this state-level registration and the district-level CDR obtained from CRS to calculate the estimated district-level mortality rate by assuming that the performance of the district's urban areas in recording these vital events is the same for all districts within the state. This estimated urban mortality rate for each district that we derive from the above methodology is what we will from here on refer to as "adjusted mortality rate".

The urban adjusted mortality rate (AMR), which is an average of 756 deaths per 100,000 people for this sample, is over 20 percent higher than the urban CDR from CRS (see Table 4.1 for summary of variables). This difference has, however, declined over time (see figure 1) showing that the system has become better at recording these vital events. The AMR has increased 17 percent while the CDR has increased almost one-and-a-half times between 1998 and 2015, from which we can infer that the former statistic is more comprehensive than the one from the CRS. For these reasons, we will show the results from the model that uses AMR as the outcome variable.

The district-level data on $PM_{2.5}$ levels were collected from recently released data published by the Atmospheric Composition Analysis Group, based on van Donkelaar et al. (2016) for the 17-year time period. The average $PM_{2.5}$ for 2015 is 55 $\mu g/m^3$, a 23 percent increase from the 1998 average figure.

To estimate the causal impact of pollution on mortality one should control for the literacy rate, which affects the demand for better air quality as well better health. To do this, we collected literacy rates for urban areas at the state level from the Census of India from 3 census years, 1991, 2001,

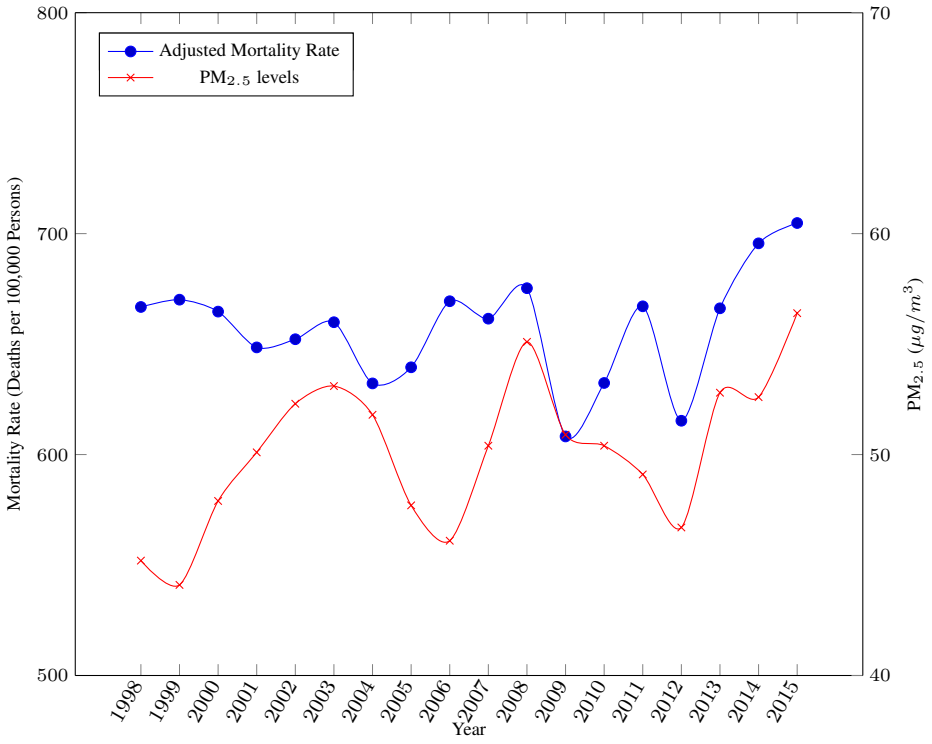


Figure 4.1: Adjusted Mortality versus Pollution Exposure

Note: Adjusted mortality rate for 13 districts were larger than 2000, which is unrealistically high. Hence, the data for these districts were removed from the analysis.

and 2011. For the years in between, we interpolated the numbers to get the values for all the 17 years. Government data show that the urban literacy rate has increased relatively rapidly in India, increasing from an average of 69 percent in 1998 across the districts we are interested in, to 79 percent in 2015.

4.3 Comparison of Relative Risks of All-Cause Mortality

Using the model we develop in this paper, we estimate the impact of $PM_{2.5}$ on human health as represented by one integrated statistic, i.e. AMR, that includes all diseases and all ages for 119 districts of India over a 17-year period. We use estimates from Burnett et al. (2018) and Krewski et al. (2009) as the basis for comparison to the relative risk calculated from our empirical model. In this section, we describe the methodology we use to make their estimates comparable to ours.

Previous studies estimate concentration responses based on North American and European datasets, with few of them focusing on developing countries that are highly polluted. To measure

the impact of high levels of pollution, therefore, many studies extrapolate from linear and log-linear models to estimate relative risks, often leading to unrealistically high relative risk estimates. To address this lacuna in measuring the impact of higher ranges of pollution, Apte (2015) used the Integrated Exposure Response (IER) model of Burnett et al. (2014) to estimate the impact of highly polluted areas on mortality. Burnett et al. (2014) estimate the relative risk for the five diseases that are attributable to pollution exposure to estimate the GBD. For adults that are 25 years or older, these diseases are ischemic heart disease (IHD), cerebrovascular disease (stroke), chronic obstructive pulmonary disease (COPD), lung cancer (LC), and for children under 5, acute respiratory lung infection (ALRI) is the risk factor attributed to particulate matter.

The dose-response figures built by Burnett et al. (2014) did not cover other possible diseases related to $PM_{2.5}$ exposure. Therefore, they were updated with outdoor air pollution levels with a higher range in the GEMM model (Burnett et al., 2018). Burnett et al. (2018) also extended their cohort analysis to include almost all non-accidental deaths beyond the 5 diseases mentioned above—non-communicable diseases (NCDs) and lower respiratory infections (LRIs)—to represent the total burden of disease due to $PM_{2.5}$ exposure. This model, called GEMM NCD+LRI, combines coefficients from multiple models while restricting the estimates at high ambient pollution levels so that they appear closer to reality. The resulting concentration response function indicates that the risk of mortality increases at a decreasing rate as pollution levels are high, confirming the supralinear shape of this function.

We use the hazard ratio of mortality from the GEMM NCD+LRI model as a basis for comparison to the relative risk of AMR from our model. We believe that risk from NCD and LRI together is a close approximation to that of AMR even though AMR will also include accidental deaths and other possible deaths that may not be included in NCD and LRI combined. This is especially since we are interested in the incremental lives saved from reducing $PM_{2.5}$ levels by certain amounts.

The GEMM model is represented by a log-linear model ($LL(z)$), which has the following form:

$$LL(z) = \frac{e^{\theta \ln \frac{z}{\alpha+1}}}{1 + e^{\frac{-(z-\mu)}{\nu}}} \text{ where } z = max(0, PM_{2.5} - 2.4\mu g/m^3) \tag{4.1}$$

where the theoretical minimum risk concentration level and is set at 2.4 $\mu g/m^3$, following the Burnett et al. (2018) framework. Since we use GEMM NCD+LRI as the basis of comparison, the parameter derived from estimating Equation (4.1) have the following values: (i) $\theta = 0.1231$; (ii) $\alpha = 0.6$; (iii) $\mu = 15.5$; and (iv) $\nu = 36.8$. The cohort analysis for GEMM NCD+LRI that we use to compare with our model also adds the Chinese male cohort so that larger ranges of pollution are included in the relative risk estimation.

The second point of comparison is the estimates in Krewski et al. (2009). They use the American Cancer Society (ACS) Cancer Prevention Study cohort using concentration levels and other covariates to derive the hazard function of mortality risk denoted by $\lambda(.)$. Their estimates are based on a Cox proportional hazard model with two main specifications, the log-log and the log-linear models.

The log-linear specification can be stated as follows (following the terminology in Goodkind et al. (2014):

$$\ln(\lambda^{lin}(X, PM_{2.5})) = \ln(\lambda_0) + X\beta^{lin} + PM_{2.5}\lambda^{lin} \tag{4.2}$$

The log-log specification can be stated as follows:

$$\ln(\lambda^{log}(X, PM_{2.5})) = \ln(\lambda_0) + X\beta^{log} + \ln(PM_{2.5})\lambda^{log} \tag{4.3}$$

where λ_0 is the baseline risk and X is the matrix of covariates that impact the risk of mortality or disease. The hazard ratio (HR), that is evaluated at two different pollution concentration levels, $PM'_{2.5}$ and $PM''_{2.5}$, and derived from Equations 4.2 and 4.3, can be defined for the 2 specification levels as:

$$\text{HR}^{lin} = \frac{\lambda^{lin}(X, PM'_{2.5})}{\lambda^{lin}(X, PM''_{2.5})} = \frac{\lambda_0 e^{X\beta^{lin} + PM'_{2.5}\gamma^{lin}}}{\lambda_0 e^{X\beta^{lin} + PM''_{2.5}\gamma^{lin}}} = e^{\gamma^{lin}(PM'_{2.5} - PM''_{2.5})} \tag{4.4}$$

$$\text{HR}^{log} = \frac{\lambda^{log}(X, PM'_{2.5})}{\lambda^{log}(X, PM''_{2.5})} = \frac{\lambda_0 e^{X\beta^{log}} {PM'_{2.5}}^{\gamma^{log}}}{\lambda_0 e^{X\beta^{log}} {PM''_{2.5}}^{\gamma^{log}}} = \left(\frac{PM'_{2.5}}{PM''_{2.5}}\right)^{\gamma^{log}} \tag{4.5}$$

Table 11 from Krewski et al. (2009) reports the ratio of hazard functions (i.e., the hazard ratio, HR) for two different pollution levels. The paper reports the HR values for changes in pollution concentrations of 10 $\mu g/m^3$. For the log-lin model, the estimated HR^{lin}=1.060 for a 10 $\mu g/m^3$ change in pollution exposure. By contrast, for the log-log model, the HR values are different based on the pollution level at the reference point. For instance, the estimated HR^{log}= 1.095 for a pollution change from 5 to 15 $\mu g/m^3$; and HR^{log}= 1.059 for a pollution change from 10 to 20 $\mu g/m^3$.

Using the above values for HR, in combination with equations 4.4 and 4.5, we can get values for γ (since this estimated value is not explicitly stated in Krewski et al. (2009)):

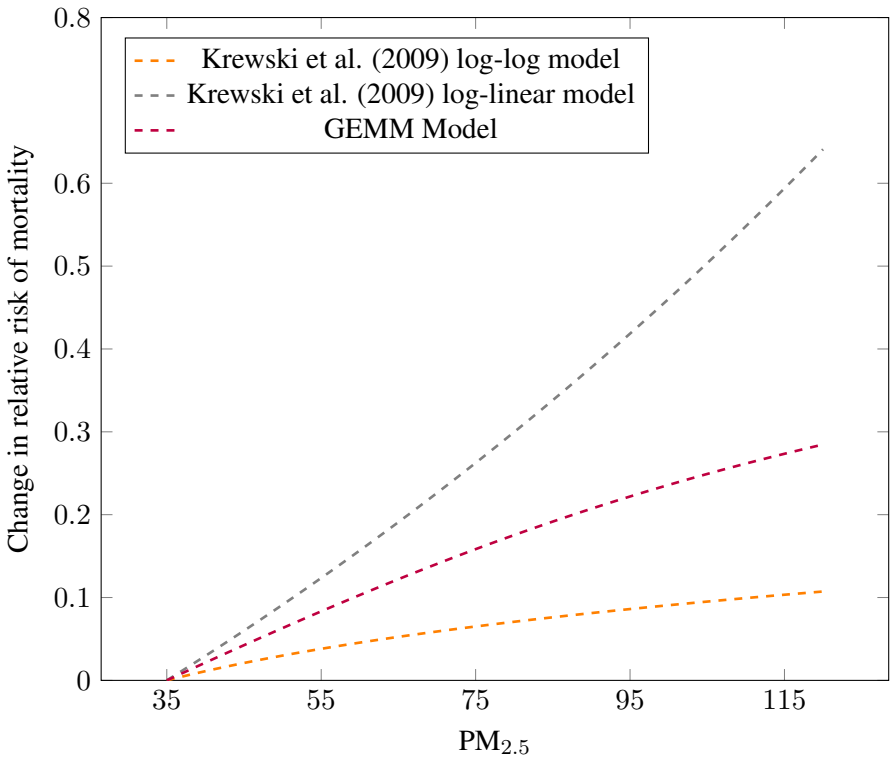


Figure 4.2: Marginal Relative Risk of Mortality by Model

$$\lambda^{lin} = \frac{\ln(1.060)}{10} = 0.005827 \tag{4.6}$$

$$\lambda^{log} = \frac{\ln(1.059)}{\ln 20 - \ln 10} = 0.082703 \tag{4.7}$$

The log-linear and log-log specifications of Krewski et al. (2009) represent two models that are very unlike each other and hence, represent the opposite ends of the spectrum. We see from Figure 4.2 that the GEMM function that is derived from Burnett et al. (2018), and the log-log form of Krewski et al. (2009) show signs of flattening out at the maximum concentration levels (of this graph) of 120 $\mu g/m^3$. The log-lin specification diverges away from the other two lines, with incremental relative risk rising at an increasing rate.

4.4 Empirical Framework

Recent studies (Burnett et al., 2018, Krewski et al., 2009) highlight the idea that highly polluted areas such as India would need a considerable decline in pollution levels to curtail pollution related

mortality substantially. The upside to this is that marginal cost of abatement at these high levels of pollution is comparatively low, enabling such countries to pursue an aggressive plan of pollution reduction. However, the size of population that is exposed to pollution, and hence the benefit from every controlled emission, underscores the importance of assessing the shape of the concentration curve for India. Our paper does exactly that, using mortality and satellite $PM_{2.5}$ data for 119 districts in India from 1998 to 2015 to determine the relationship between pollution exposure and AMR. While this association has been determined for cleaner areas such as the U.S. and Europe, or derived globally from a meta-analytic setup, no paper so far has empirically determined this relationship specifically for India, using Indian data for pollution and mortality.

4.4.1 Theoretical Underpinnings of Pollution Abatement

The theory of environmental policy takes into consideration abatement cost and damage functions in arriving at efficient allocation of emissions. The traditional approach assumes that the marginal abatement cost of reducing emissions is small in the beginning, and increases as emissions continue to decline. At the same time, marginal health damage reduces more at higher levels of emissions, with improvements declining as abatement increases, which is in contrast to the concave or "supra-linear" shape.

These assumptions on the curvature of these functions mean that the damage to health increases as emissions increase, with the marginal cost being small at high pollution levels. This perspective invites policy makers to focus on the dirtiest places first. In contrast, below a certain level, emissions are no longer dangerous to human health, and the cost of abatement is quite high.

Recent studies suggest a different reality. Goodkind et al. (2014) find that policy is more effective and has higher social net benefit and lower marginal cost as pollution declines, when the relationship between human health and pollution is assumed to be "supralinear". If this supralinear shape is indeed accurate, marginal health benefit from abatement increases as air gets cleaner, and also marginal cost is largest where the most abatement is needed. This strong concavity or "supra-linearity" suggests that the most polluted places in India will see the biggest benefits only after substantial improvement in the quality of air, which encourages nations to strive for cleaner air.

Even for empirical studies that analyze countries such as the U.S., with concentrations below 30 $\mu g/m^3$, there is some evidence of the supra-linear shape. For instance, Krewski et al. (2009) uses ACS cohort data to analyze the risk of $PM_{2.5}$ levels on mortality rate. They find that the risk of mortality declines with increased exposure, flattening out at higher exposures.

In their integrated exposure model, where they combine many models using different sources of

$PM_{2.5}$, Burnett et al. (2014) show that the concentration response is non-linear and relative risk of mortality levels off at higher exposure. Apte (2015), using high resolution pollution concentration data and the IER function established in Burnett et al. (2014), showed that small improvements in air quality in relatively clean areas (e.g., North America and Europe) can lead to substantial reduction in mortality attributed to pollution, with these mortality reductions declining in dirtier regions. This is due to the supralinear concentration response curve. Burnett et al. (2018) uses outdoor pollution exposure to also establish that there is a supralinear functional form in the concentration response curve.

Even with these implications from the supralinear shape, we know from the results section (Section 4.4.4) that the number of lives saved (as per the global burden of disease calculations in Burnett et al. (2014)) depends on the population exposed to these high levels of pollution, in addition to relative risk of mortality and the average mortality rate of the region. This means that even though we need substantial reduction in pollution to see the benefits in health, as per the shape of the curve, the high levels of population suggests that even small reductions in pollution will still lead to substantial number of lives saved.

The methodology highlighted below will enable this study to examine the shape of the concentration curve for India, describing the relative risk of mortality from long term pollution exposure for this 17-year period.

4.4.2 Methodology

The relationship between ambient air pollution and human health can chart the course of action for a nation's environmental policy [5]. We use panel data spanning 119 districts in India from 1998 to 2015 to estimate the concentration function for India, where these annual outdoor $PM_{2.5}$ levels range between 16 and 120 $\mu g/m^3$.

The equation of interest, and hence the framework we follow is:

$$AMR_{it} = \alpha + \beta X_{it} + \gamma D_{it} + \delta_i + \tau_t + \lambda_i \times t + \eta_i \times t^2 + \varepsilon_{it} \tag{4.8}$$

where AMR_{it} is mortality at the district level and D is the treatment variable or in this case any one of the measures of district-level ambient air pollution. δ represents the district-level fixed effects, τ is the vector of time fixed effects. In addition to district and year fixed effects, we have

[5] Sankar et al. (n.d.) estimates the relationship between pollution and mortality using 119 districts from 1998 to 2007, a subset of the data for this paper. This paper focuses on $PM_{2.5}$ and curvature of the response, while Sankar et al. (n.d.) also compares the impact on mortality across different pollution types, SO_2, NO_2, and TSP.

also included district-specific linear trends and district-specific quadratic trends that can control for any systematic trends in mortality within the districts irrespective of the impact of $PM_{2.5}$. X represents the matrix of control variables, which in this paper is represented by one variable, i.e., urban literacy rate. Bertrand, Duflo, Mullainathan 2001 showed us that district level difference-in-differences model were often affected by serial correlation that could bias the estimates of standard errors (SE) downwards. Hence, we have used SEs clustered at district level in all the models, which are robust to any arbitrary error correlation within the districts.

Note here that we identified a few districts that had over 2,000 mortality rate (i.e., deaths per 100,000 persons). We consider this as an improbable statistic. In these identified districts, we also find that data are not available for some of the years. Given that the data are either null or too high, we decided to remove these districts entirely from our analysis.

In order to identify the impact of pollution on mortality (i.e., γ), we follow a number of steps as part of our identification strategy. We address issues of reverse causality, unobserved heterogeneity, measurement error, violation of Stable Unit Treatment Value Assumption (SUTVA) and omitted variable bias that may confound the estimate of γ. This identification strategy is described in detail in Section E.5 of the Appendix.

4.4.3 Shape of the Concentration Response Curve

The main objective of the paper is to find the functional form of the relationship between AMR and $PM_{2.5}$ as it relates to India. In order to understand this relationship, we first use the box-cox transformation to inform us of the shape of this concentration response curve. We use the methodology as suggested by Box and Cox (1964), where the dependent variable and the main independent variable (here, $PM_{2.5}$) are transformed by using different powers or logarithms. Using maximum likelihood method, it then shows the transformation that works best for the data. For the data that we have for this paper, the transformation that fits the best is the log-log functional form. This suggests a "supralinear" shape that is line with other literature that studies this relationship (as mentioned earlier in this paper).

So far, we have used parametric methods to determine the shape of the relationship between mortality and $PM_{2.5}$. To allow for greater flexibility in the specification, and confirm what we saw from the Box Cox transformation, we use a semi-parametric approach to understand the real shape of the function representing this relationship. We use this approach as a robustness check to confirm whether the "supralinear" shape is correct. Specifically, we run a restricted cubic spline regression in the following form:

$$AMR_{it} = \alpha + \beta X_{it} + f(D_{it}) + \delta_i + \tau_t + \lambda_i \times t + \eta_i \times t^2 + \varepsilon_{it} \tag{4.9}$$

The only difference between the models as represented by Equations 4.8 and 4.9 is the way $PM_{2.5}$ enters into the relationship. The piecewise polynomial function, $f(.)$, or the spline, in Equation 4.9 allows the relationship to be nonlinear. While setting up the "correct" spline function, we have to choose the number of knots. Choosing the number of knots, however, is tricky because there is a trade-off between choosing fewer knots that do not have as much of the curse of dimensionality versus choosing more knots that offers greater flexibility in the specification. Especially since $PM_{2.5}$ is only one of several factors affecting mortality, we have chosen a model with 3 knots so that we can control for overfitting.

4.4.4 Results

Based on the methodology described above and Equation 4.9, we can see from Figure F1 that the spline regression yielded a curve that looked similar to the log-log functional form or the "supralinear" curve as suggested by much of the epidemiology literature. We found that it is more useful to display the results of this exercise using a graph rather than the coefficients of the regression since it is easier to understand the shape of curve that the non-parametric method recommends.

Based on the above recommendation, we examine the results from several models that follow a log-log or log-linear specification. Following this strategy in Equation 4.8 described above, we ran several specifications based on the log-log and the log-linear functional forms. We add the squared and cubed terms of concentration levels in some of the models to find if these curvatures describe the data better. The results of the model are displayed in 4.2.

Table 4.2: Relationship Between Adjusted Mortality Rate and $PM_{2.5}$

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
ln($PM_{2.5}$)	0.162760	0.193786*					
	(0.107042)	(0.108996)					
$PM_{2.5}$			0.002149*			0.006383	0.013380
			(0.001238)			(0.005174)	(0.010338)
$PM_{2.5}^2$				0.000008		-0.000023	-0.000130
				(0.000005)		(0.000024)	(0.000118)
$PM_{2.5}^3$					0.000000		0.000000
					(0.000000)		(0.000000)
Observations	1819	1819	1819	1819	1819	1819	1819
Adjusted R^2	0.374	0.453	0.451	0.451	0.450	0.452	0.452
AIC	-1443.1	-1815.2	-1811.3	-1806.9	-1810.1	-1809.6	-1808.8
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Quadratic Trends	No	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions are controlled for urban literacy rate, but the coefficients are not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

In line with previous work, our model finds that the log-log specification (Model [2]) offers the best fit in terms of adjusted R-squared. I also compare the models using AIC and find that the log-log specification has the best fit. Model [2] in Table 4.2 tells us that a 10 percent increase in $PM_{2.5}$ levels raises the mortality rate by 1.9 percent, which also happens to be significant at 10 percent level. In terms of relative risk (RR), we can express it as:

$$RR^{log} = \left(\frac{PM'_{2.5}}{PM''_{2.5}}\right)^{\gamma^{log}} = \left(\frac{PM'_{2.5}}{PM''_{2.5}}\right)^{0.194} = 1.019 \tag{4.10}$$

The fit for the log-linear specification is almost as good as the log-log version, in terms of adjusted R-squared. This impact, however, is not significant. From the coefficient in Model [3] in Table 4.2, we can estimate the relative risk in the following manner:

$$RR^{lin} = e^{\gamma^{lin}(PM'_{2.5} - PM''_{2.5})} \tag{4.11}$$

Using Equation 4.11, we can say that when $PM_{2.5}$ levels increase by 10 $\mu g/m^3$, relative risk increases to 1.02 or mortality risk increases by 2%.

When we added the squared term to this log-linear specification, the fit in Model [6] is similar to that of Model [3]. Also, the coefficient of the linear term is positive and the coefficient for the

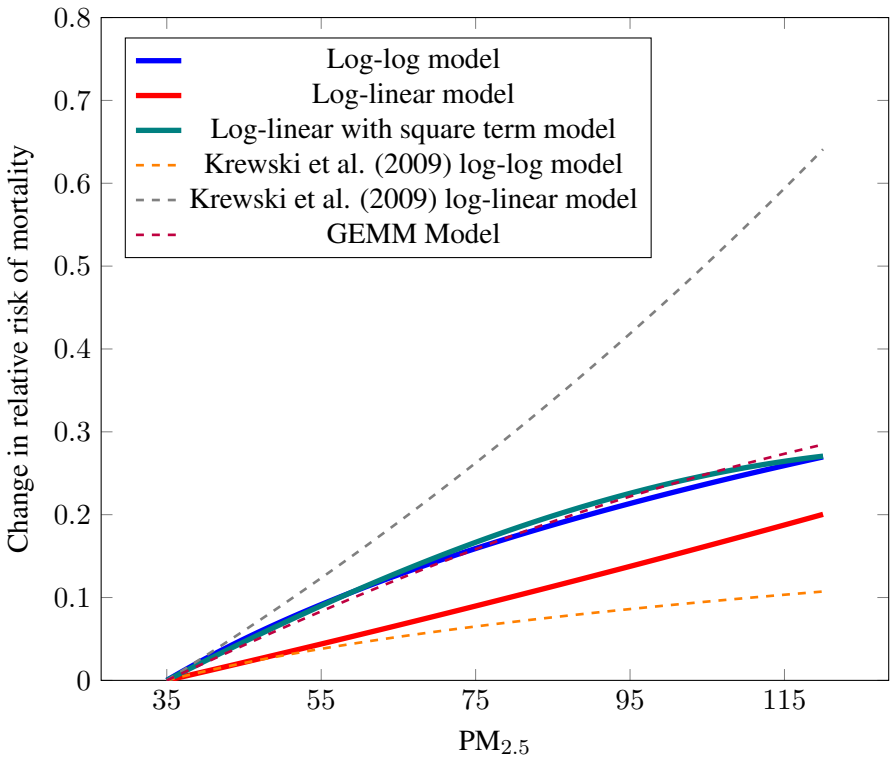


Figure 4.3: Marginal Relative Risk of Mortality by Model

squared term is negative, which means that the curve is showing signs of flattening out at higher pollution levels. Here, we can estimate RR^{sq} as follows:

$$RR^{sq} = e^{\gamma_1^{sq}(PM'_{2.5} - PM''_{2.5}) + \gamma_2^{sq}({PM'_{2.5}}^2 - {PM''_{2.5}}^2)} \tag{4.12}$$

Figure 4.3 consists of all the relative risk values using baseline $PM_{2.5}$ level of 35 $\mu g/m^3$. We chose 35 $\mu g/m^3$ to be the baseline level of pollution since this pollution level is WHO's recommended interim target. Model [7] from Table 4.2, with the cubed term, gives us relative risk values that are much higher than the confidence intervals of the GEMM and Krewski et al. (2009) models. The AIC as well as the adjusted R^2 show that the log-log specification show the best fit, and hence will be used to estimate the impact of reducing pollution levels on number of deaths.

Our paper adapts the methodology to estimate the Global Burden of Disease used in Burnett et al. (2014) to calculate the premature mortality for each district (M_i) from all causes attributable to ambient $PM_{2.5}$.

$$M_i = P_i \times \hat{I}_k \times (RR(C_i) - 1), \text{where} \hat{I}_k = \frac{I_k}{\bar{RR}_k} \tag{4.13}$$

where P_i is the district population; I_k is the annual average all-cause mortality rate including all ages for India as a whole; and $\bar{RR}_k$ across N districts for this sample is defined as follows:

$$\bar{RR}_k = \frac{\sum_{i=1}^{N} P_i \times RR(C_i)}{\sum_{i=1}^{N} P_i} \tag{4.14}$$

An aggressive policy to reduce pollution levels might look like the World Health Organization's Interim Target 1 (IT-1) where the guideline was to reduce $PM_{2.5}$ to 35 $\mu g/m^3$,s where the pollution was higher than that. A change in the mortality from such a policy can be estimated using the following version of Equation 4.13:

$$\Delta M_i = P_i \times \hat{I}_k \times (RR(C_i^*) - RR(C_i)) \tag{4.15}$$

Using the model described above, we estimated the number of lives saved if the districts listed in the data were to reduce their pollution levels to 35 $\mu g/m^3$, as per the IT-1 target. Here, we assume that there is no change in mortality for those districts that have $PM_{2.5}$ levels less than 35 $\mu g/m^3$.

If the $PM_{2.5}$ levels are reduced to 35 $\mu g/m^3$ for the districts in the sample for the data in 2015, the log-log specification tells us that there will 128,540 fewer premature deaths, with a confidence interval of [13324,-272,839]. The GEMM model in Burnett et al. (2018) estimates are very close our log-log model, with 127,085 lives saved when the pollution is reduced to standard of 35 $\mu g/m^3$. Krewski et al. (2009) log linear model estimates a much higher reduction, with 211,372 fewer deaths. In contrast, the log-log model in Krewski et al. (2009) predicts the smallest effect with 55,608 fewer premature deaths.

Our log-log and log-linear with square term models are well within the confidence interval of the GEMM. The 95% confidence interval estimates from GEMM is [-90,465, -163,696] as of 2015 if $PM_{2.5}$ exposure is reduced to 35 $\mu g/m^3$ across India. Our log-linear model, however, does not fall within this interval. This is the same case with the log-log as well as the log-linear model from Krewski et al. (2009).

4.5 Conclusion

Traditional theory suggests that marginal benefits from abatement get smaller as the air gets cleaner. Recent research, however, has shown that the relationship between air pollution and health could be supralinear, which means that countries with low baseline concentration levels could benefit from even lower pollution levels. By contrast, countries with dirtier air as in some parts of India will need a substantial reduction in air pollution to start seeing benefits in terms of a lower relative risk

in mortality from pollution.

Our paper is one of the first studies to show shape of this central relationship using data for India across 119 districts from 1998 to 2015. We find strong evidence that the impact of pollution on mortality increases, but flattens out at higher pollution levels. The impact of this result on policy is important. For countries such as the U.S., where the pollution levels are relatively low, this supralinearity provides the motivation to reduce pollution even when the air is clean. This is because marginal benefit of abatement increases even as pollution levels reduce.

Our results show that the shape of this curve is indeed concave or "supralinear" and is similar in impact to the GEMM model, as developed by the (Burnett et al., 2018). It must be noted, however, that for the same amount of marginal benefit from abatement the pollution reduction to be achieved in India has to be substantially higher. Although, from the model estimates, a large reduction has to be achieved before India can reap any of the benefits to health, the lives saved from reduced pollution also depends on the size of population exposed to the concentrations levels, baseline mortality rates, and relative risks. So, even though the marginal benefit might be low as air begins to get cleaner, the lives saved is still big enough for policy to pursue.

This paper shows that India can pursue policy to aggressively reduce its high concentration levels and see benefits. With the cost of energy (especially with renewables) declining fast, reduced abatement costs could potentially enable this process in the future. Further, a supralinear concentration curve tells us that even as the air gets cleaner policy could aim to keep tightening air quality standards.

Chapter 5

Conclusion

Overflowing landfills and exposure to ambient air pollution are serious problems in the world that adversely affects the ecosystem. Some of the tools that governments and communities have used to bring about pro-environment behavior are social capital and environmental regulations that set air quality standards. In my thesis, I examine the impact of these tools on better environmental quality, in terms of both less waste and greater air quality.

My first chapter examines the impact of social capital in two parts. Firstly, using a theoretical model of household utility, I show that the impact of an increase in social norm of recycling increases the recycling rate per ton of solid waste, while an increase in this social norm reduces solid waste per capita. Secondly, the two propositions stated above become the foundation of my empirical framework.

For the empirical part of the chapter, I first build a social capital index for all the 87 counties of MN for the years 1996 to 2013 using the methodology in Rupasingha et al. (2006). It uses 15 different variables that are classified into four major types of cohesiveness in a community. It combines total group organizations from the Census Business Patterns survey, non-profit organizations from the NCCS, the Census response rate, and voter turnout from the David Leip's Presidential Elections Atlas. The data for organizations are from secondary data sources, and hence do not have any of the usual measurement problems that data on factors like "trust" collected from opinions in surveys such as General Social Survey. In addition, the organizations that are considered in this index do not have members participating for a financial incentive, and can be viewed as people coming together to share information and forming social capital.

Using this social capital index, the aim of this chapter is to estimate its impact on recycling rate and waste per capita. In an instrumental variables regression setup, we see that an increase in social

capital leads to a significant decline of 183 tons in waste per capita, consistent with the theoretical model. By contrast, we see that an increase in social capital does not lead to a significant change in recycling rate. Some of this could be explained by the fact that an increase social capital leads to a significant decline in recycling per capita. I argue here that this response is driven by the impact social capital has on waste. When the social capital impacts in a way that people get conscious of how much they waste, the waste per capita declines. With it, recycling per capita also declines, and together they lead an insignificant impact on recycling rate. These results show that building social capital can help increase pro-environment behavior.

My second chapter examines another major global issue—ambient air pollution. I estimate the impact of environmental regulation on mortality in India. The Air Act of 1981 and the regulations that followed have tackled industrial and vehicular pollution by setting standards and targeting particulate matter in the last three decades. We examine this issue in two parts, where we take two decades of data for the regulations and investigate the impact of these policies on mortality in the first part, and estimate the impact of different pollution types directly on mortality in the second part.

In this chapter, we estimate the impact of regulations on mortality that includes all causes and all ages for the first time for India. In the economics literature, it is common to consider infant mortality as the outcome since it is easier to identify the effect. However, we also know from large studies like Krewski et al. (2009) and Woodruff et al. (1997) that the impact of adults is an order of magnitude bigger than that of infants. It is important to public health, therefore, to consider the impact of these policies on mortality.

The two major policies that we consider in this analysis are the Supreme Court Action Plan (SCAP) and Catalytic Converter (CC) policy. The SCAP includes many policies that set standards for various cities as well as industries, capping the amount of pollution they emit. The CC policy is more specific, where it directed people to retrofit their vehicles with the catalytic converter in order to filter out the polluted air.

In the first part of the chapter, using a difference-in-differences framework, and comprehensive data for 119 districts from 1987 to 2007, we find that the districts with catalytic converter policy have 19% lower mortality as compared to those that do not have this policy in place. Given that India's pollution levels are still very high at 74 $\mu g/m^3$ in 2015, we analyzed the impact of different pollution types directly on mortality in the second part of this chapter. In the policy statements outlining the SCAP and CC policies, the pollution type that was targeted until very recently was TSP, i.e., the particulate matter that is less than 100 microns in size.

In this second part, we show for the first time using actual data for India that $PM_{2.5}$ is more

dangerous than TSP to human health. In our estimates, we find that $PM_{2.5}$ leads to a significant increase in mortality, but an increase in TSP leads to a small and insignificant increase in mortality. In addition, we find that for a certain specification, an increase in SO_2 leads to a significant increase in mortality. This effect, however, becomes insignificant when all the fixed effects are included. We also find that the impact of NO_2 is insignificant. We conclude from this that $PM_{2.5}$ is more dangerous than TSP, but we cannot comment on the SO_2 and NO_2 since they are from different sources of combustion and not comparable to TSP and $PM_{2.5}$.

The results from the two parts of the chapter inform us that while the policies in India have been effective in reducing the mortality rate significantly, we might be able to get greater gains in human health if the policies were to shift its focus to $PM_{2.5}$. This result is also in line with recent policy statements that urge the pollution control boards to move away from monitoring PM_{10} and TSP, and monitor $PM_{2.5}$.

The third chapter of my thesis estimates the functional form of the relationship between mortality and $PM_{2.5}$ for India. It is the first paper that addresses the functional form of this relationship for India, and estimate the number of lives saved due to a reduction in pollution levels.

Using mortality and $PM_{2.5}$ data for 119 districts in India from 1998 to 2015, we estimate the shape of the concentration response curve. We examined this shape in three different ways. We first used a box-cox transformation that indicated that the shape took a log-log form. Second, we also ran a semi-parametric regression, where we considered a restricted cubic spline. The result from this estimation also showed a concave relationship, similar to the result from the box cox transformation.

As a third set of estimations, we ran fixed effects models with $\ln$(Mortality) as the outcome, following the results from the box-cox and spline regressions. As independent variables, we consider different permutations of $\ln(PM_{2.5})$, $PM_{2.5}$, and also the square and cubic terms. Based on the different regressions we ran, we found that the best fit (in terms of AIC) was the log-log, once again consistent with the box-cox regression.

Using the estimated fixed effects model with the log-log specification, and also the theoretical model from Goodkind et al. (2014), we estimated the number of lives saved from reducing pollution levels from the current levels to 35 $\mu g/m^3$, as per WHO's recommended target. The estimated shape informs us that pollution levels in places where the air is dirtier will have reduce the pollution more aggressively to see the same change mortality rates as compared to places with cleaner air. These estimates are based on the assumptions of marginal costs and damages at different pollution levels in the theoretical model. Based on these assumptions as well as population exposed to this pollution levels, we know that even though takes an substantial improvement in air quality to see a change in mortality, the number of lives saved is large enough to matter.

Bibliography

Abbott, Andrew, Shasikanta Nandeibam, and Lucy O'Shea, "Recycling: Social norms and warm-glow revisited," *Ecological Economics*, 2013, *90*, 10–18.

Anderson, Michael L., "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," *NBER Working Paper*, 2015, (21578).

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving," *The Economic Journal*, 1990, *100*, 464–477.

Angrist, Joshua David and Jörn-Steffen Pischke, *Mostly harmless econometrics:An empiricist's companion*, Princeton:Princeton University Press, 2009.

Apte, Joshua, "Addressing Global Mortality from Ambient PM2.5.," *Environmental Science and Technology*, 2015, *49*, 8057–8066.

Assessment, Millennium Ecosystem, "Environmental Degradation and Human Well-Being: Synthesis.," 2005.

Autor, David H., "Outsourcing at Will: The Contribution of Unjust Dismissal Doc-trine to the Growth of Employment Outsourcing.," *Journal of Labor Economics*, 2003, *21* (1), 1–42.

Azar, Ofer H., "What sustains social norms and how they evolve? The case of tipping," *Journal of Economic Behavior & Organization*, 2004, *54*, 49–64.

Beeler, Rob et al., "Effects of long-term exposure to air pollution on natural-cause mortality: An analysis of 22 European cohorts within the multicentre ESCAPE project," *Lancet*, 2014, *383*, 785–795.

Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky, "Lagged Explanatory Variables and the Estimation of Causal Effects," *Journal of Politics*, 2017, *79* (3), 949–963.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Bodvarsson, Örn B. and William A. Gibson, "An Economic Approach to Tips and Service Quality," *The Social Science Journal*, 1999, *36* (1), 137–147.

Box, G. E. P. and D. R. Cox, "An analysis of transformations," *Journal of Royal Statistical Society*, 1964, *B 26*, 211–252.

Brauer, Michael et al., "Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013.," *Environmental Science and Technology*, 2016, *50* (1), 79–88.

Broughton, Edward, "The Bhopal Disaster and Its Aftermath: A Review.," *Environmental Health*, 2005, *4* (6).

Burnett, Richard T. et al., "An integrated risk function for estimating the global burden of dis-ease attributable to ambient fine particulate matter exposure.," *Environmental Health Perspectives*, 2014, *122* (4), 397–403.

__ and __ , "Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter," *Proceedings of the National Academy of Sciences of the United States of America*, 2018, *115* (38), 9592–9597.

Chay, Kenneth Y. and Michael Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession.," *The Quarterly Journal of Economics*, 2003, *118* (3), 1121–1167.

__ , Carlos Dobkin, and Michael Greenstone, "The Clean Air Act of 1970 and Adult Mortality," *Journal of Risk and Uncertainty*, 2003, *27* (3), 279–300.

Clark, Andrew E., *Journal of Labor Economics*, 2003, *21* (2), 323–351.

__ and Andrew J. Oswald, "Comparison-concave utility and following behaviour in social and economic settings," *Journal of Public Economics*, 1998, *70*, 133–155.

Cohen, Aaron J. et al., "Urban Air Pollution," *Comparative Quantification of Health Risk*, 2004, pp. 1353–1434.

Cole, Ethan Cohen and Jason M. Fletcher, "Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis," *British Medical Journal*, 2009, *338* (7685), 28–31.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi, "Plausibly Exogenous," *Review of Economics and Statistics*, 2012, *94* (1), 260–272.

Conrad, Jon M. and Colin W. Clark, *Natural Resource Economics: Notes and Problems*, Cambridge University Press, 1987.

Currie, Janet and Matthew Neidell, "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?," *The Quarterly Journal of Economics*, 2005, *120* (3), 1003–1030.

__ , __ , and Johannes F Schmeider, "Air pollution and infant health: lessons from New Jersey," *Journal of Health Economics*, 2009, *28*, 688–703.

Dockery, Douglas W. et al., "An association between air pollution and mortality in six US cities.," *New England Journal of Medicine*, 1993, *329* (24), 1753–1759.

Dominici, Francesca et al., "Air Pollution and Mortality: Estimating Regional and National Dose-Response Relationships," *Journal of the American Statistical Association*, 2002, *97*, 100–111.

E.L.Glaeser, "The formation of social capital," *Canadian Journal of Public Policy*, 2001, *2*, 34–40.

Glewwe, Paul and Petra Todd, *Impact Evaluation in Developing Countries: Theory, Methods and Practice*, The World Bank, forthcoming.

Goodkind, Andrew L., Jay S. Coggins, and Julian D. Marshall, "A Spatial Model of Air Pollution: The Impact of the Concentration-Response Function," *Journal of the Association of Environmental and Resource Economists*, 2014, *1* (4), 451–479.

Graff Zivin, Joshua and Matthew Neidell, "Environment, Health, and Human Capital," *Journal of Economic Literature*, 2013, *51* (3), 689–739.

Greenstone, Michael and Rema Hanna, "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India.," *American Economic Review*, 2014, *104* (10), 3038–3072.

Heutel, Garth and Christopher J. Ruhm, "Air Pollution and Procyclical Mortality," *Journal of the Association of Environmental and Resource Economists*, 2016, *3* (3), 667–706.

Hinds, William C., *Aerosol Technology*, New York:John Wiley & Sons, 1999.

IHME, "GBD Compare Data Visualization," 2016, *Accessed 06/28/2018.*

International, IPUMS, "Integrated Public Use Microdata Series, International: Version 7.1 [dataset]," 2018.

Krewski, Daniel et al., "Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality.," *Special Report to the Health Effects Institute*, 2009.

Lepeule, Johanna et al., "Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009.," *Environmental Health Perspectives*, 2012, *120* (7), 965–970.

Lim, Stephen S. et al., "A Comparative Risk Assessment of Burden of Disease and Injury Attributable to 67 Risk Factors and Risk Factor Clusters in 21 Regions, 1990-2010: A Systematic Analysis for the Global Burden of Disease Study 2010," *Lancet*, 2012, *380* (9859), 2224–2260.

Lleras-Muney, Adriana, "The needs of the Army: using compulsory relocation in the military to estimate the effect of environmental pollutants on children's health," *Journal of Human Resources*, 2010, *35* (3), 549–590.

Mahapatra, Prasanta et al., "Civil registration systems and vital statistics: successes and missed opportunities.," *Lancet*, 2007, *370* (9599), 1653–1663.

Meagher, Patrick J., "Environmental Protection and Industries in Developing Countries: The Case of India Since Bhopal.," *Georgetown International Environmental Law Review*, 1990, *3* (1), 1–54.

Muller, Nicholas Z. and Robert Mendelsohn, "Efficient Pollution Regulation: Getting the Prices Right.," *American Economic Review*, 2009, *99* (5), 1714–1739.

Nag, Moni, "Impact of Social and Economic Development on Mortality: Comparative Study of Kerala and West Bengal.," *Economic and Political Weekly*, 1983, *18* (19/21), 877–900.

Narain, Urvashi and Alan Krupnick, "The Impact of Delhi's CNG Program on Air Quality.," *Resources for the Future Discussion Paper*, 2007.

__ **and Ruth Greenspan Bell**, "Who Changed Delhi's Air? The Roles of the Court and the Executive in Environmental Policymaking.," *Resources for the Future Discussion Paper*, 2005, pp. 05–48.

National Clean Air Programme, "National Clean Air Programme," 2015.

Nyborg, Karine et al., "Social norms as Solutions," *Science*, 2016, *354* (6308), 42–43.

Office of the Legislative Auditor, "Evaluation Report: Recycling and Waste Reduction," 2015.

Office of the Registrar General, "Vital Statistics of India Based on the Civil Registration System 2010," 2010.

Pope, C. Arden III et al., "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults," *American Journal of Respiratory and Critical Care Medicine*, 1995, *151* (3), 669–674.

__ **and** __ , "Lung Cancer, Cardiopulmonary Mortality, and Longterm Exposure to Fine Particulate Air Pollution," *The Journal of the American Medical Association*, 2002, *287* (9), 1132–1141.

Putnam, R.D., *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, Princeton, New Jersey, 1993.

Rupasingha, Anil, Stephan J. Goetz, and David Freshwater, "The production of social capital in US counties," *The Journal of Socio-Economics*, 2006, *35*, 83–101.

Sankar, Ashwini, Jay S. Coggins, and Andrew L. Goodkind, "Combating Particle Pollution: Estimating the Impact of Pollution Standards on Mortality in India." Unpublished.

UNEP, *United Nations Environment Programme Environmental Impact Assessment Training Resource Manual, Second Edition. 2002. Topic 7 – Mitigation and impact management*, UNEP Environmental Impact Assessment Training Resource Manual, 2002.

U.S. EPA, "Integrated Science Assessment for Particulate Matter (Final Report). EPA-600-R-08-139F.," 2009.

__ , "Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. EPA-452-R-12-005.," 2012.

__ , "Advancing Sustainable Materials Management: 2016 Recycling Economic Information (REI) Report," 2016.

van Donkelaar, Aaron et al., "Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors," *Environmental Science & Technology*, 2016, *50* (7), 3762–3772.

Woodruff, Tracey J., Jeanne Grillo, and Kenneth C. Schoendorf, "The Relationship between Selected Causes of Postneonatal Infant Mortality and Particulate Air Pollution in the United States," *Environmental Health Perspectives*, 1997, *105* (6), 608–612.

Yin, P et al., "Long-term Fine Particulate Matter Exposure and Nonaccidental and Cause-specific Mortality in a Large National Cohort of Chinese Men.," *Environmental Health Perspectives*, 2018, *125* (11).

A.1 Proofs Appendix

Proof of Proposition 1 (i) For those households that are of type $\theta > 0$, using Envelope theorem with 2.9 tells us the following:

$$\frac{dU^0((R_t^*, \theta)}{d\theta} = \frac{\partial U}{\partial \theta} - \mu^0 \frac{\partial g}{\partial \theta} \tag{A1}$$

where g is the function that represents the constraint function $X = \gamma w(1 - tR - \bar{l})$. From A1, we get:

$$\frac{dU^0}{d\theta} = -D(X^0(1 - R^0)) - \mu^0 \times 0 = -D(X^0(1 - R^0)) \leq 0 \tag{A2}$$

where μ^0, X^0 and R^0 are values of the lagrange multiplier, waste per capita and recycling rates at the equilibrium.

For those households who derive direct disutility from wasting (i.e., $\theta > 0$), Equation (A2) tells us that consumer utility decreases by $D(X^0(1 - R^0))$ with an increase in θ, in equilibrium.

Proof of Proposition 1 (ii) For all households, the change in utility due to an increase in social norm (using the Envelope theorem) moves in the following manner:

$$\frac{dU^0}{dR_t^*} = -\nu'(R^0 - R_t^*) \tag{A3}$$

At equilibrium, Equation (A3) tells us that an increase in the recycling social norm R_t^* increases or decreases utility based on whether the value of recycling rate at equilibrium is higher or lower than the prevailing social norm. Following assumption 3, we know that $\nu'(x) < 0$ for all $x > 0$. Hence, if $R^0 - R_t^* > 0$ or $R^0 > R_t^*$, $\nu'(R^0 - R_t^*) < 0$, which implies that utility at the optimal value increases with an increase in social norm. By contrast, if $R^0 < R_t^*$, utility at the optimal value decreases with an increase in social norm.

Proof of Proposition 2.

We combine Equations (2.7) and (2.8) to eliminate μ:

$$\gamma wt(B'(.) - \theta(1 - R)D'(.)) = \nu'(.) + \theta X D'(.) \tag{A4}$$

Totally differentiating Equation (A4), we get:

$$\begin{aligned} [\gamma wt(\theta D'(.) + \theta(1 - R)XD''(.)) - \nu''(.) + \theta X^2 D''(.)]dR \\ +[\gamma wt(B''(.) - \theta(1 - R)^2 D''(.)) - \theta D'(.) - \theta X(1 - R)D'']dX = -\nu'' dR_t^* \end{aligned} \tag{A5}$$

Totally differentiating Equation (2.4), we get:

$$dX = -\gamma wt dR \tag{A6}$$

(i) Combining Equations (A5) and (A6) to eliminate dX, we get:

$$\begin{aligned}[\gamma wt(\theta D'(.) + \theta(1-R)XD''(.)) - \nu''(.) + \theta X^2 D''(.)]dR \\ -[\gamma^2 w^2 t^2 (B''(.) - \theta(1-R)^2 D''(.)) - \gamma wt\theta D'(.) - \gamma wt\theta X(1-R)D'']dR = -\nu'' dR_t^*\end{aligned} \tag{A7}$$

Using assumptions 1 to 3, we know that the first line in Equation A7, i.e., $[\gamma wt\theta(D'(.)+(1-R)XD''(.))-\nu''(.)+\theta X^2 D''(.)]$ is positive. The phrase in second line $[\gamma^2 w^2 t^2 (B''(.) - \theta(1-R)^2 D''(.)) - \gamma wt\theta D'(.) - \gamma wt\theta X(1-R)D'']$ is however, negative. This implies that the left hand side as well as right hand side are positive, or:

$$\frac{dR}{dR_t^*} > 0 \tag{A8}$$

(ii) Using Equation (A6) to substitute for dR in Equation (A5), we get:

$$\frac{dX}{dR_t^*} < 0 \tag{A9}$$

B.2 Figures and Tables Appendix

Table B1: Ordinary Least Square Regression Results for Recycling Per Capita

	[1]	**[2]**	**[3]**	**[4]**
Social capital index	27.619*	23.231	-4.501	-53.077
	(16.004)	(16.282)	(27.765)	(39.311)
Persons with bachelor degree	-28.398**	-15.658**	13.817**	-1.087
	(10.985)	(7.362)	(5.785)	(13.206)
Cities that provide curbside recycling				38.261
				(27.341)
Observations	2107	2107	2107	1540
Adjusted R^2	0.211	0.119	0.412	0.548
County Fixed Effects	Yes	Yes	Yes	Yes
Linear Trend	No	Yes	No	No
Year Fixed Effects	Yes	No	No	No
County-Specific Linear Trends	No	No	Yes	Yes
County-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Recycling per capita per 1,000 persons; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county. Constants are not displayed. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B2: First Stage Regression Showing Impact of Unemployment

Dependent Variable:	Social Capital Index		
	[1]	**[2]**	**[3]**
Annual Average Unmeployment Rate	0.084***	0.053***	0.050***
	(0.015)	(0.009)	(0.006)
Cities that provide curbside recycling	0.077	0.003	0.018
	(0.113)	(0.053)	(0.078)
Persons with bachelor degree		-0.014	-0.023*
		(0.036)	(0.013)
Constant	0.397***	0.698	0.761***
	(0.131)	(0.581)	(0.184)
Observations	1595	1539	1539
Adjusted R^2	0.195	0.770	0.913
County Fixed Effects	Yes	Yes	Yes
Linear Trend	Yes	No	No
County-Specific Linear Trends	No	Yes	Yes
County-Specific Quadratic Trends	No	No	Yes

Notes: Clustered Huber-White SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B3: Endogeneity Checks for the Impact of Social Capital (NCCS) on Recycling and Waste Generation

Dependent Variable:	Recycling rate			Recycling per capita			MSW per capita		
	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**	**[7]**	**[8]**	**[9]**
Charitable Org. Per 10000 Persons	-0.029	-0.154*	-0.090	-1.999	-4.675	-4.560**	-4.930**	-6.250***	-5.888***
	(0.073)	(0.079)	(0.074)	(2.305)	(3.036)	(2.182)	(2.502)	(2.365)	(2.124)
Cities that provide curbside recycling	1.430	-0.969	2.235	34.036	16.593	28.973	-4.206	-56.389**	-32.835**
	(2.537)	(1.897)	(1.725)	(27.945)	(23.177)	(28.280)	(30.235)	(22.765)	(14.026)
Persons with bachelor degree		-0.542*	-0.092		9.593	9.467*		23.145*	18.652
		(0.300)	(0.452)		(6.908)	(5.631)		(12.260)	(14.653)
Observations	1394	1345	1345	1595	1539	1539	1394	1345	1345
Adjusted R^2	0.015	0.368	0.529	0.001	0.172	0.499	-0.164	0.233	0.494
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear Trend	Yes	No	No	Yes	No	No	Yes	No	No
County-Specific Linear Trends	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
County-Specific Quadratic Trends	No	No	Yes	No	No	Yes	No	No	Yes
F-Statistic (Instrumental Variable)	61.60	116.08	270.76	31.36	41.44	123.96	61.60	116.08	270.76

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county.
Unemployment rate is used as the instrument for number of non-profit organizations (used here as another indicator of social capital).
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table B4: Placebo and Falsification Tests

Dependent Variable:	Waste Tires/MSW Rate		Recycling rate		MSW per capita	
	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**
Social capital index	-0.020	0.036				
	(0.035)	(0.062)				
Waste Tires/MSW Rate			0.364	0.285	4.303	-3.760
			(0.458)	(0.364)	(15.408)	(9.098)
Cities that provide curbside recycling	-0.148	-0.034	1.580	1.212	71.225	-27.022
	(0.137)	(0.084)	(2.406)	(1.287)	(54.802)	(22.531)
Persons with bachelor degree	-0.005	-0.003	-0.745**	0.195	6.359	38.082***
	(0.009)	(0.012)	(0.326)	(0.479)	(7.151)	(11.766)
Observations	1344	1344	1344	1344	1344	1344
Adjusted R^2	0.005	0.132	0.200	0.625	0.004	0.639
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Linear Trend	Yes	No	Yes	No	Yes	No
County-Specific Linear Trends	No	Yes	No	Yes	No	Yes
County-Specific Quadratic Trends	No	Yes	No	Yes	No	Yes

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each county. Constants are not displayed. $^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

C.3 Identification Strategy Appendix

This section discusses all the identification issues that may arise in the model developed in Section 3.5.

There are five broad sources of endogeneity that could thwart any efforts made to identify γ in Equation (3.4). They are (i) reverse causality, (ii) unobserved heterogeneity, (iii) measurement error, (iv) violation of the Stable Unit Treatment Value Assumption, and (v) Omitted Variable Bias. We describe all these issues in detail below and also list the methods we use to address these issues.

C.3.1 Reverse Causality

We are claiming that growing air pollution leads to increased mortality. But it is quite possible that concerns about higher mortality levels due to polluted air could raise the demand for cleaner air leading to younger and healthier people perhaps migrating to cleaner locations, leaving behind the older and the sicker people, and muddling the impact of pollution on mortality. Controlling for the district and year specific fixed effects are effective in accounting for people behaving such clusters.

Census data, however, seems to indicate that dirty air is not the primary reason why people in India migrate within the country. As per the 2001 census, over 90% of migration within India is due to employment, marriage, education, business or familial reasons. Even if air pollution is a reason for some people in India to migrate, we believe it is small enough that reverse causality from mortality to ambient air pollution will not bias the results one way or the other. Further, to test if people move in anticipation of pollution related health problems, we ran a Granger causality test (Table D14) and show that $PM_{2.5}$ at different leads and lags do not affect mortality.

One could also argue that cumulative exposure (Pope et al., 2002, Lepeule et al., 2012) and medical history (Dockery et al., 1993) could affect the relationship between pollution and mortality, potentially exaggerating the effect. We dealt with this issue by introducing lags into the model (in addition to the control variables already present) to see if they affect mortality. The results from Lepeule et al. (2012) suggest that including "lag 1 (i.e., exposure during the previous year) to be the best fit exposure window for all-cause mortality". We included two lags for $PM_{2.5}$ as well as TSP to see their impact. Table D15 from the appendix shows that almost all the lag terms are insignificant, the only exception being model [2] when the current $PM_{2.5}$ and first lag of $PM_{2.5}$ were included together in the equation. It turns out that the coefficient on the first lag in this equation is negative and significant. This is a common phenomenon, harvesting, where higher mortality in one year leaves a population with lower mortality in the next year.

C.3.2 Unobserved heterogeneity

Using district-specific and year-specific fixed effects, we aim to remove most of the unobserved heterogeneity in the district-year observations. While district and year-specific fixed effects control for most unobserved

heterogeneity, there could be systematic district and year specific trends. To control for these, we use district-specific linear and quadratic trends. In addition, we tried zone level fixed effects but, they did not make any difference to the fit of the regression or the coefficient of interest. It should also be noted here that the inclusion of district-specific linear time trends obviates the need for separate linear time trend and zone-specific linear time trend variables.

C.3.3 Measurement error

In India, as with many other low-income countries, registration of deaths is incomplete (Mahapatra et al., 2007). The SRS estimates of deaths rates have historically been used to get a sense of how complete the civil registration system is, in the country. Although still not perfect, the SRS shows that this system has improved from a modest 56% average registration rate in 1987 by over 20 percentage points in the last 2 decades.

This vital statistic, however, is available only at the state level. Using this, we can derive the registration rate at the state level. For a more disaggregated number, we still must rely on the Census. Therefore, to deal with this issue of completeness, this paper uses both measures as the outcome variable: the district level mortality rate as measured by the Census (CMR) as well as the mortality rates from Census that are adjusted for by the SRS registration rate (AMR).

If there is still systematic under-reporting of this statistic, the estimation of γ is likely to be biased. Despite this being a threat to identification, it must be noted that under-reporting will cause the estimate of γ to be lower than the actual. In other words, since we expect this coefficient to be positive, $\hat{\gamma}$ will biased towards zero.

In addition to measurement error in the outcome variable, there is also possibly under-reporting in the independent variables, i.e., TSP, SO_2, and NO_2 levels as recorded by the Pollution Control Boards across India. This error on the right-hand-side variables could lead to attenuation bias where the slope, γ, could be biased towards zero. In this case, we have collected data for $PM_{2.5}$ levels from GIS that does not have under-reporting issues. The regressions using $PM_{2.5}$ levels will then act as a robustness check for the relationship.

To sum up, we can say that reverse causality is not an issue in this case for India. Unobserved heterogeneity is largely controlled for using all the fixed effects. In the case of measurement error, the registration rate measures the extent of the error. Any further systematic error would most likely will bias the coefficient of interest downwards, thus underestimating the effect of pollution on mortality.

C.3.4 Violation of Stable Unit Treatment Value Assumption (SUTVA)

Another confounding factor in identifying the effect of air pollution on mortality is the violation of SUTVA, i.e., the effect of ambient air pollution in a certain district-year might affect the mortality rate in another district-year.

There are different ways in which this violation might have an effect. Within a district, air pollution in one year might affect mortality the following year. Or, pollution in a district in a certain year might affect a person from the neighboring district that same year or even the following year. For instance, several people live in the neighboring areas surrounding the district of Delhi and can be affected by the dirty Delhi air that might travel to their area. This could potentially affect their health that year or the following year.

The fixed effects imposed in these models can take care of many of the issues listed above. Also, the tests for cumulative exposure (Table D15) show us that lags in pollution levels do not affect mortality. To deal with any other leftover violations of the SUTVA, more structure must be imposed on the model. That would, along with all the fixed effects, introduce a lot more stress on the data. We understand that this might be a limitation of this paper, but since the smaller dataset (with $PM_{2.5}$ levels) has only 770 rows, we choose not to include any more variables into this model.

C.3.5 Omitted Variable Bias

Even after accounting for unobserved heterogeneity and other endogeneity issues through a battery of fixed effects, there could be bias due omitted variables. It must be reiterated that this study is based on data for the urban areas within the districts of India. So, all the omitted variables are related to urban and not rural areas. To address this, we considered a number of control variables in Equation (3.4). In our initial runs of the models in Equation (3.4), we included income as a an indicator of economic activity, which however was not significant in any of the model specifications. Hence, we removed this variable from our subsequent regression runs. We include literacy rate to indicate the area's economic activity as well as the role it might play in enhancing health conditions, similar to Pope et al. (1995) as well as Pope et al. (2002). But just as in these papers, we found no significance in the impact of these variables. We also include age as a control variable (as in Pope et al., 1995), although we do not divide the data set into age-specific cohorts, and we find significance. In addition, to see the impact that the availability of medical services might have on mortality we included the number of doctors per capita. However, this variable was not significant.

We also considered including temperature, but did not include this variable in our models. Since Pope et al. (1995) specifically mentions that high, mean, and low temperatures are not correlated with particulate matter or sulfate levels, and had little effect on the estimated impact of pollution on mortality, we did not include this in our model.

D.4 Figures and Tables Appendix

Table D1: Effect of Policy on Crude Mortality Rate

	[1]	[2]	[3]	[4]	[5]	[6]
Supreme Court Action Plan (SCAP)	-0.007	0.083			0.001	0.094
	(0.083)	(0.081)			(0.080)	(0.079)
SCAP X Linear Trend	-0.026	-0.019			-0.033	-0.030
	(0.034)	(0.028)			(0.035)	(0.029)
Catalytic Converter Policy (CC)			-0.102	-0.251*	-0.107	-0.240*
			(0.070)	(0.133)	(0.071)	(0.128)
CC X Linear Trend			0.011	0.058	0.013	0.060*
			(0.013)	(0.035)	(0.014)	(0.035)
Observations	1743	1743	1743	1743	1743	1743
Adjusted R^2	0.090	0.490	0.092	0.498	0.092	0.498
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	No	Yes	No	Yes	No	Yes
District-Specific Quadratic Trends	No	Yes	No	Yes	No	Yes
SCAP District-Year Fixed Effects	No	Yes	No	Yes	No	Yes
CC District-Year Fixed Effects	No	Yes	No	Yes	No	Yes

Notes: Dependent variable: Log of Urban CMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**} p < 0.05, ^{*} p < 0.1$

Table D2: Event Study Analysis: Effect of SCAP

Dependent Variable:	ln(Urban AMR)
$SCAP_{-5}$	0.023
	(0.076)
$SCAP_{-4}$	0.136**
	(0.057)
$SCAP_{-3}$	0.157***
	(0.058)
$SCAP_{-2}$	0.032
	(0.060)
$SCAP_{0}$	0.064
	(0.044)
$SCAP_{+1}$	0.014
	(0.046)
$SCAP_{+2}$	-0.022
	(0.084)
$SCAP_{+3}$	-0.058
	(0.126)
$SCAP_{+4}$	-0.149
	(0.200)
$SCAP_{+5}$	-0.037
	(0.089)
Observations	1739
Adjusted R^2	0.448

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. The control variables include district fixed effects, year fixed effects, district-specific linear and quadratic trends, urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D3: Event Study Analysis: Effect of CC policy

Dependent Variable:	ln(Urban AMR)
CC_{-5}	-0.007
	(0.055)
CC_{-4}	-0.033
	(0.055)
CC_{-3}	-0.040
	(0.052)
CC_{-2}	0.002
	(0.072)
CC_{0}	-0.058
	(0.060)
CC_{+1}	-0.126
	(0.089)
CC_{+2}	-0.049
	(0.087)
CC_{+3}	-0.053
	(0.087)
CC_{+4}	-0.005
	(0.051)
CC_{+5}	0.016
	(0.035)
Observations	1739
Adjusted R^2	0.446

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. The control variables include district fixed effects, year fixed effects, district-specific linear and quadratic trends, urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D4: Event Study Analysis: Effect of SCAP After Accounting For Pre-Trends

Dependent Variable:	ln(Urban AMR)
$SCAP_{-5}$	0.035
	(0.143)
$SCAP_{-4}$	0.080
	(0.126)
$SCAP_{-3}$	0.085
	(0.148)
$SCAP_{-2}$	-0.029
	(0.079)
$SCAP_{0}$	0.075*
	(0.042)
$SCAP_{+1}$	0.051
	(0.048)
$SCAP_{+2}$	0.047
	(0.084)
$SCAP_{+3}$	0.015
	(0.132)
$SCAP_{+4}$	-0.092
	(0.211)
$SCAP_{+5}$	-0.037
	(0.107)
Observations	1739
Adjusted R^2	0.447

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. The control variables include SCAP-districts-year trends and CC-districts-year trends, district fixed effects, year fixed effects, district-specific linear and quadratic trends, urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D5: Event Study Analysis: Effect of CC policy After Accounting For Pre-Trends

Dependent Variable:	ln(Urban AMR)
CC_{-5}	0.071
	(0.084)
CC_{-4}	0.045
	(0.036)
CC_{-3}	0.113
	(0.095)
CC_{-2}	0.010
	(0.132)
CC_{0}	0.001
	(0.106)
CC_{+1}	-0.265
	(0.161)
CC_{+2}	-0.246*
	(0.126)
CC_{+3}	-0.139
	(0.103)
CC_{+4}	-0.173
	(0.115)
CC_{+5}	-0.179**
	(0.087)
Observations	1739
Adjusted R^2	0.456

Notes: Clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. The control variables include SCAP-districts-year trends and CC-districts-year trends, district fixed effects, year fixed effects, district-specific linear and quadratic trends, urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D6: Heterogeneity Checks for Pollution Levels on Mortality

	[1]	**[2]**	**[3]**	**[4]**	**[5]**
Supreme Court Action Plan (SCAP)	0.067	0.046	0.034	0.040	0.034
	(0.059)	(0.085)	(0.103)	(0.097)	(0.103)
SCAP X TSP_{-1}		0.000	0.000		-0.000
		(0.000)	(0.001)		(0.001)
SCAP X TSP_{-2}			0.000		
			(0.001)		
SCAP X Average[TSP_{-1}, TSP_{-2}]				0.000	0.000
				(0.000)	(0.002)
Catalytic Converter Policy (CC)	-0.273**	-0.384*	-0.412**	-0.428**	-0.412**
	(0.119)	(0.224)	(0.205)	(0.199)	(0.205)
CC X TSP_{-1}		0.000	-0.000		-0.001
		(0.001)	(0.001)		(0.003)
CC X TSP_{-2}			0.001		
			(0.001)		
CC X Average[TSP_{-1}, TSP_{-2}]				0.001	0.002
				(0.001)	(0.002)
Observations	1739	1739	1739	1739	1739
Adjusted R^2	0.454	0.453	0.453	0.454	0.453
District Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes	Yes	Yes
District-Specific Quadratic Trends	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. $TSP_{-\tau}$ is the TSP level τ periods before the start of a particular policy for each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D7: Heterogeneity Checks for Pollution Levels on Mortality **sans** Delhi

	[1]	**[2]**	**[3]**	**[4]**	**[5]**
Supreme Court Action Plan (SCAP)	0.089	0.075	0.062	0.068	0.062
	(0.077)	(0.094)	(0.112)	(0.106)	(0.112)
SCAP X TSP_{-1}		0.000	0.000		-0.000
		(0.000)	(0.001)		(0.001)
SCAP X TSP_{-2}			0.000		
			(0.001)		
SCAP X Average[TSP_{-1}, TSP_{-2}]				0.000	0.000
				(0.001)	(0.002)
Catalytic Converter Policy (CC)	-0.289**	-0.401*	-0.431**	-0.446**	-0.431**
	(0.131)	(0.233)	(0.215)	(0.209)	(0.215)
CC X TSP_{-1}		0.000	-0.000		-0.001
		(0.001)	(0.001)		(0.003)
CC X TSP_{-2}			0.001		
			(0.001)		
CC X Average[TSP_{-1}, TSP_{-2}]				0.001	0.002
				(0.001)	(0.002)
Observations	1726	1726	1726	1726	1726
Adjusted R^2	0.454	0.454	0.454	0.454	0.454
District Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes	Yes	Yes
District-Specific Quadratic Trends	Yes	Yes	Yes	Yes	Yes

[1] Delhi is one of the most polluted cities in India, and where the policies got implemented first.

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. $TSP_{-\tau}$ is the TSP level τ periods before the start of a particular policy for each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D8: Effect of TSP levels on Mortality Rate

	[1]	**[2]**	**[3]**	**[4]**
ln(TSP)	0.004	0.006	0.019	0.010
	(0.025)	(0.031)	(0.029)	(0.024)
Observations	809	809	809	809
Adjusted R^2	0.030	0.030	0.471	0.566
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes
District-Specific Linear Trends	No	No	Yes	Yes
District-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D9: Effect of $PM_{2.5}$ versus TSP on Crude Mortality Rate

Dependent Variable:	ln(Urban CMR)		
	1998-2007 data		1988-2007 data
	[1]	**[2]**	**[3]**
ln($PM_{2.5}$)	0.273*		
	(0.158)		
ln(TSP)		-0.003	-0.010
		(0.026)	(0.022)
Observations	836	456	810
Adjusted R^2	0.503	0.519	0.575
District Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes
District-Specific Quadratic Trends	Yes	Yes	Yes

Notes: Dependent variable: Log of urban CMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Models [1] and [2] are run for only a subset of the data, i.e., for the years 1998-2007. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D10: Effect of SO_2 Levels on Adjusted Mortality Rate

	[1]	**[2]**	**[3]**	**[4]**
ln(SO_2)	0.017	0.029	0.043*	0.028
	(0.024)	(0.026)	(0.022)	(0.020)
Observations	894	894	894	894
Adjusted R^2	0.024	0.028	0.368	0.471
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes
District-Specific Linear Trends	No	No	Yes	Yes
District-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D11: Effect of NO_2 Levels on Adjusted Mortality Rate

	[1]	**[2]**	**[3]**	**[4]**
ln(NO_2)	-0.010	-0.007	0.002	0.024
	(0.038)	(0.037)	(0.026)	(0.035)
Observations	915	915	915	915
Adjusted R^2	0.023	0.024	0.360	0.465
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	No	Yes	Yes	Yes
District-Specific Linear Trends	No	No	Yes	Yes
District-Specific Quadratic Trends	No	No	No	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table.The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D12: Effect of the Combination of Different Pollutants on Adjusted Mortality Rate

	[1]	**[2]**	**[3]**	**[4]**
Ln ($PM_{2.5}$)	0.243	0.206		0.256
	(0.209)	(0.214)		(0.223)
Ln (NO_2)	-0.084		0.014	-0.097
	(0.054)		(0.039)	(0.061)
Ln (SO_2)		-0.017	0.025	0.019
		(0.052)	(0.023)	(0.055)
Observations	560	538	886	531
Adjusted R^2	0.349	0.344	0.472	0.361

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions in this table have district and year fixed effects, and district-specific linear and quadratic trends. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D13: Impact of Pollution on Infant Mortality

	[1]	[2]	[3]	[4]
$\ln(PM_{2.5})$	0.338	0.098		
	(0.722)	(0.644)		
ln(TSP)			-0.056	-0.070
			(0.103)	(0.100)
Constant	-30.922	-336.264	-0.299	32.427
	(121.782)	(3884.116)	(16.136)	(17.075)
Observations	330	330	411	411
Adjusted R^2	0.343	0.542	0.295	0.564
District Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes	Yes
District-Specific Quadratic Trends	No	Yes	No	Yes

Notes: Dependent variable: Log of infant mortality; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Supreme Court Action Plan is used as the instrument for TSP levels. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D14: Granger Causality Testing the Impact of the Anticipation of Pollution Levels

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
$\ln(PM_{2.5t-3})$				-0.031				
				(1.134)				
$\ln(PM_{2.5t-2})$			-0.077	-0.010				
			(0.224)	(2.525)				
$\ln(PM_{2.5t-1})$		-0.222	-0.406	-0.460				
		(0.164)	(0.345)	(3.700)				
$\ln(PM_{2.5})$	0.254*	0.123	0.022	0.162				
	(0.146)	(0.227)	(0.353)	(4.386)				
$\ln(PM_{2.5t+1})$		0.099	0.041	0.177				
		(0.158)	(0.285)	(3.537)				
$\ln(PM_{2.5t+2})$			0.016	0.150				
			(0.237)	(2.673)				
$\ln(PM_{2.5t+3})$				-0.033				
				(1.444)				
$\ln(TSP_{t-3})$								0.012
								(0.102)
$\ln(TSP_{t-2})$							-0.076	-0.134
							(0.046)	(0.183)
$\ln(TSP_{t-1})$						0.022	-0.015	-0.048
						(0.036)	(0.049)	(0.137)
$\ln(TSP)$					0.010	0.023	0.038	0.041
					(0.024)	(0.031)	(0.039)	(0.100)
$\ln(TSP_{t+1})$						0.037	0.044	0.105*
						(0.031)	(0.035)	(0.054)
$\ln(TSP_{t+2})$							0.080**	0.144***
							(0.035)	(0.041)
$\ln(TSP_{t+3})$								0.103**
								(0.048)
Observations	836	657	478	322	809	528	337	215
Adjusted R^2	0.409	0.340	0.310	0.178	0.566	0.541	0.557	0.542

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions have district, year, district-specific linear and quadratic fixed effects. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table D15: Effect of Cumulative Pollution Exposure on Mortality

	[1]	**[2]**	**[3]**	**[4]**	**[5]**	**[6]**
$\ln(PM_{2.5})$	0.254*	0.165	0.303**			
	(0.146)	(0.160)	(0.141)			
$\ln(PM_{2.5t-1})$		-0.224*	-0.115			
		(0.132)	(0.156)			
$\ln(PM_{2.5t-2})$			-0.019			
			(0.153)			
ln(TSP)				0.010	0.011	0.012
				(0.024)	(0.027)	(0.030)
$\ln(TSP_{t-1})$					0.012	-0.012
					(0.030)	(0.032)
$\ln(TSP_{t-2})$						-0.027
						(0.046)
Observations	836	754	672	809	666	537
Adjusted R^2	0.409	0.383	0.404	0.566	0.564	0.589
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Linear Trends	Yes	Yes	Yes	Yes	Yes	Yes
District-Specific Quadratic Trends	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable: Log of urban AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. Additional control variables include urban literacy rate, number of doctors per 100,000 persons, and state's average age. They are included in all the regressions above, but not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

E.5 Identification Strategy Appendix

This section discusses all the identification issues that may arise in the model developed in Section 4.4.

While this relationship is key in informing policy, it is fraught with many issues related to endogeneity. These issues fall under three broad categories of endogeneity, that we elucidate in detail in our working paper, Sankar et al. 2018. These issues: (i) reverse causality, (ii) unobserved heterogeneity, and (iii) measurement error could potentially foil our efforts in extracting the causal impact, γ. We will go through them briefly in this section.

E.5.1 Reverse Causality

High pollution in certain areas could lead to sorting, where younger and healthier people move perhaps leaving behind older and sicker people. This type of endogeneity would then bias the impact of pollution upwards, exaggerating the effect of pollution on health.

Our paper deals with this issue partly by using extensive fixed effects, controlling for district and year fixed effects, and also linear and quadratic trends within districts. Furthermore, Indian Census data indicates that dirty air is not the primary reason why people in India migrate within the country. As per the 2001 census, over 90 percent of migration within India is due to employment, marriage, education, business or familial reasons. Even if air pollution is a reason for some people in India to migrate, we believe it is small enough that reverse causality from mortality to ambient air pollution will not bias the results one way or the other.

A related question that could hinder causality is if there is cumulative exposure and chronic illness, which could potentially magnify the impact of pollution. To test if this will bias the coefficient, we introduced lags into the model to see pollution up to three years ago (following Lepeule et al. 2012) can perhaps affect mortality in the current year. Table F1 in the Appendix shows that none of the lag terms are significant showing that this is not a concern for this sample of aggregate level annual data for India.

In establishing causality, we also want to know if the cause comes strictly before the effect. If this is indeed true, then any lag or lead terms of pollution should have no effect on mortality in the current year (Angrist and Pischke, 2009). The Tables F2 and F3 in the Appendix show that the lag and lead terms do not have any effect on the mortality. This is true for both the log-log as well as the log-linear models. In a similar vein, we have also included the district-specific linear and quadratic trends which allows each district to have its own trend. Models [2], [3] and [4] from Tables F2 and F3 (below) show that these terms do not make any difference to the impact of pollution on mortality.

E.5.2 Unobserved heterogeneity

We use district-specific and year-specific fixed effects to remove most of the unobserved heterogeneity in the district-year observations. There could also be systematic district and year specific trends, and to account for these effects we use linear and quadratic trends.

E.5.3 Measurement error

It is well known that CDR from the CRS is underreported due to reasons that are enumerated in the data and sources section. Because of these reasons, we use the registration rate in the urban areas, as estimated by the SRS, to account for the completeness of the data. The registration rate is then used to adjust the CDR and create AMR in order to make it closer to the actual death rate in that district. If this AMR is still lower than the actual mortality rate, then this will lead to biasing the impact of pollution on mortality downwards, which means the estimate of γ, $\hat{\gamma}$, will be the lower bound. This AMR is not expected to have any measurement errors. If they are errors in measurement, they would be random and would not affect the estimated coefficients.

We do not expect any systematic measurement errors with respect to $PM_{2.5}$ data since they are estimated using satellite based data as well as calibrated using surface $PM_{2.5}$.

E.5.4 Violation of SUTVA

Another confounding factor in identifying the effect of air pollution on mortality is the violation of SUTVA, i.e., the effect of ambient air pollution in a certain district-year might affect the mortality rate in another district-year.

There are different ways in which this violation might have an effect. Within a district, air pollution in one year might affect mortality the following year. Some of this has been studied when we tested for cumulative exposure or granger causality, as shown in Tables F2 and F3. Furthermore, pollution in a district in a certain year might also affect a person from the neighboring district that same year or even the following year. For instance, several people live in the neighboring areas surrounding the district of Delhi and can be affected by the dirty Delhi air that might travel to their area. This phenomenon could potentially affect their health in the current time period or the next. However, since the data are annual, there is lesser chance that pollution in the current year would affect health in the following year.

This effect has been mostly tackled by the battery of fixed effects that controls for any year effects affecting the whole nation or any district specific effects and trends. However, even if this is still an issue, we will need more structure that must be imposed on the model in order to deal with it. That would, along with all the fixed effects, introduce a lot more stress on the data. We understand that this is a limitation of this paper, but since the dataset has only 2,142 rows, we choose not include any more variables into this model.

E.5.5 Omitted Variable Bias

Even after accounting for unobserved heterogeneity and other endogeneity issues through a battery of fixed effects, there could bias due omitted variables. It must be reiterated that this study is based on data for the urban areas within the districts of India. So, all the omitted variables are related to urban and not rural areas. To address this, we considered a number of control variables in Equation (4.4). In our initial runs of the models in Equation (4.4), we included income as a an indicator of economic activity, which however was not significant in any of the model specifications. Hence, we removed this variable from our subsequent regression runs. We include literacy rate to indicate the area's economic activity as well as the role it might play in enhancing health conditions, similar to Pope et al. (1995) as well as Pope et al. (2002). But just as in these papers, we found no significance in the impact of these variables.

We also considered including temperature, but did not include this variable in our models. Since Pope et al. (1995) specifically mentions that high, mean, and low temperatures are not correlated with particulate matter or sulfate levels, and had little effect on the estimated impact of pollution on mortality, we did not include this in our model.

F.6 Figures and Tables Appendix

Figure F1: Robustness Check for the Concentration Response Curve Using Non-Parametric Regression

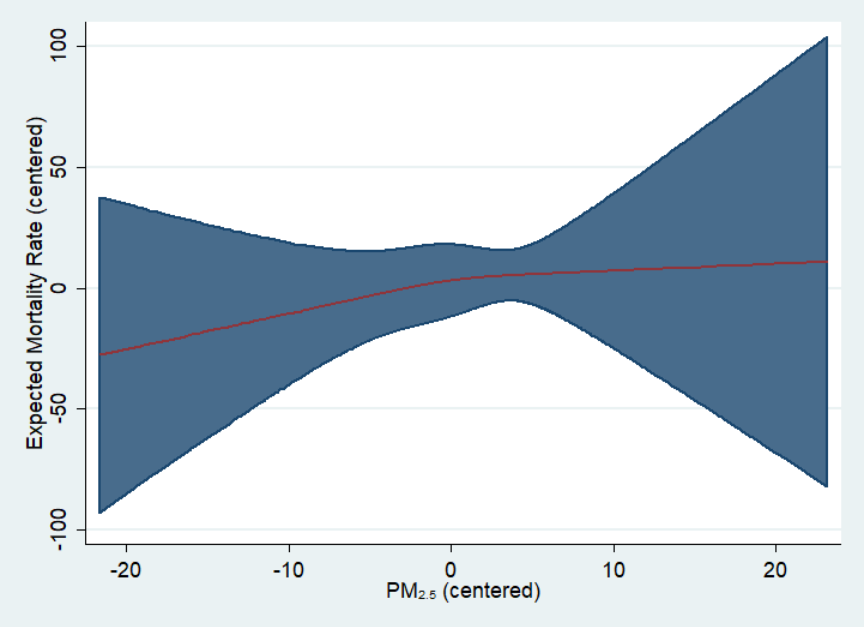


Table F1: Cumulative Exposure of $PM_{2.5}$ Levels on Mortality

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
ln($PM_{2.5}$)	0.194*	0.171*	0.130	0.117	0.141	0.155	0.148
	(0.109)	(0.099)	(0.099)	(0.104)	(0.126)	(0.142)	(0.124)
ln($PM_{2.5t-1}$)		0.112	0.099		0.088		
		(0.092)	(0.102)		(0.101)		
ln($PM_{2.5t-2}$)			-0.111	-0.122	-0.126		
			(0.084)	(0.081)	(0.093)		
ln($PM_{2.5t-3}$)					-0.023	-0.024	
					(0.072)	(0.080)	
ln(($PM_{2.5t-1}$+$PM_{2.5t-2}$+$PM_{2.5t-3}$)/3)							-0.076
							(0.167)
Observations	1819	1726	1633	1633	1540	1540	1540
Adjusted R^2	0.453	0.466	0.468	0.468	0.466	0.465	0.465

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions are controlled for urban literacy rate, district, year, district-specific linear and quadratic fixed effects, but the coefficients are not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table F2: Granger Causality Test for Impact of $PM_{2.5}$ Levels on Mortality (Log-Log Specification)

	[1]	**[2]**	**[3]**	**[4]**
$\ln(PM_{2.5t-3})$				0.175
				(0.256)
$\ln(PM_{2.5t-2})$			-0.060	0.021
			(0.259)	(0.488)
$\ln(PM_{2.5t-1})$		0.053	-0.072	0.028
		(0.136)	(0.318)	(0.610)
$\ln(PM_{2.5})$	0.194*	0.144	-0.006	0.136
	(0.109)	(0.151)	(0.279)	(0.496)
$\ln(PM_{2.5t+1})$		0.179	0.015	0.122
		(0.155)	(0.315)	(0.615)
$\ln(PM_{2.5t+2})$		-0.100	-0.245	-0.128
		(0.183)	(0.347)	(0.587)
$\ln(PM_{2.5t+3})$		-0.002	-0.118	-0.122
		(0.142)	(0.255)	(0.525)
$\ln(PM_{2.5t+4})$		0.038	-0.113	-0.183
		(0.140)	(0.269)	(0.458)
$\ln(PM_{2.5t+5})$		0.039	-0.076	-0.121
		(0.164)	(0.256)	(0.284)
Observations	1819	1176	1083	990
Adjusted R^2	0.453	0.457	0.494	0.571

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions are controlled for urban literacy rate, district, year, district-specific linear and quadratic fixed effects, but the coefficients are not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table F3: Granger Causality Test for Impact of $PM_{2.5}$ Levels on Mortality (Log-Linear Specification)

	[1]	[2]	[3]	[4]
$PM_{2.5t-3}$				0.007
				(0.004)
$PM_{2.5t-2}$			-0.000	0.006
			(0.004)	(0.007)
$PM_{2.5t-1}$		-0.000	-0.001	0.007
		(0.002)	(0.004)	(0.009)
$PM_{2.5}$	0.002*	-0.000	-0.001	0.009
	(0.001)	(0.002)	(0.005)	(0.010)
$PM_{2.5t+1}$		0.002	0.001	0.010
		(0.002)	(0.005)	(0.011)
$PM_{2.5t+2}$		-0.001	-0.003	0.005
		(0.003)	(0.006)	(0.010)
$PM_{2.5t+3}$		-0.001	-0.002	0.003
		(0.003)	(0.004)	(0.009)
$PM_{2.5t+4}$		-0.000	-0.002	0.000
		(0.002)	(0.004)	(0.007)
$PM_{2.5t+5}$		0.001	0.000	0.000
		(0.002)	(0.004)	(0.005)
Observations	1819	1176	1083	990
Adjusted R^2	0.451	0.456	0.492	0.572

Notes: Dependent variable: Log of AMR; clustered SEs are included in parenthesis to allow for arbitrary correlation of residuals within each district. All regressions are controlled for urban literacy rate, district, year, district-specific linear and quadratic fixed effects, but the coefficients are not displayed in the table. The regressions are all population weighted. $^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$