Air Pollution and Child Health in Urban India

Arkadipta Ghosh and Arnab Mukherji
Mathematica Policy Research — IIM Bangalore

Extended Abstract

1. **Introduction**:

Exposure to air pollution has been linked to poor child health outcomes in a range of studies that have looked at a variety of health measures (Bruce et. al. (2000), Chay and Greenstone (2003), Frankenberg et. al. (2004), Currie el. al. (2005) & Janke et. al. (2009)). For example, Jaychandran (2009) investigates the effect of outdoor air pollution due to the forest fires in Indonesia in 1997 on infants and reports that there was a 1% decline in the Indonesian cohort size due to these fires; Smith (2000) on the other hand looks at the impact of solid fuels used for cooking at home and suggests that as much as 4-5% of the national disease burden for India may be explained by indoor air pollution alone. A limitation in the current literature is the lack of objective measures of both indoor and outdoor air pollution in the same analysis. Additionally, several studies use air pollution proxies such as the occurrence of forest fires (Jayachandran 2009) or the type of cooking fuels used at home (Smith 2000), in the absence of more direct pollution measures. An alternative to using such proxies is to use air pollution data gathered from direct observation with expensive monitoring instruments.

In this paper, we combine directly measured ambient air pollution data for 2005-06, collected by the National Air Monitoring Program of the Ministry of Environment and Forests in India with household and child-level data from NFHS 3 (2005-6). This allows us to construct an analysis sample that not only has variation in the type of solid fuel used at home (our proxy for indoor air pollution, as in some of the previous studies), but also variation in the average level of air pollution that households are exposed to during the month of their interview. Hence, we are able to investigate the relative effects of both types of exposure to air pollution on child morbidity, as captured by the incidence of two common illnesses in children – cough and fever – in the week prior to the interview.

Apart from using both ambient and indoor pollution measures, we also explicitly tackle concerns on model specification and causality in this paper. With repeated observations from the same city, we are able to control for unobservable city fixed effects that may uniformly affect all children in the same city, for example, pollen content, traffic congestion, general quality of cooking fuel, proximity to a river, etc. We also address potential misspecification concerns by jointly modeling the probability of a child having a fever and that of having a cough, since they both pertain to the same underlying health status of the child and are likely to be determined by similar factors – both internal and external to the child. This is an important departure from the literature where these are usually studied as independent events.

We find that a rise in ambient air pollution significantly increases the likelihood of a child suffering from cough and fever in the past week. However, the type of cooking fuel used at home is not significantly related to child morbidity after accounting for ambient air pollution and other child- and household-level control variables. Thus, while bad air is bad for child health, we find that ambient air pollution appears to be a more significant determinant of the child health outcomes we study. This suggests that controlling city-wide air pollution should receive greater emphasis in urban planning, infrastructure development, and the formulation of urban development policies in general. We also find a significant correlation between the two child morbidity outcomes – fever cough, which suggests that models that do not explicitly account for this correlation are likely to be misspecified.

The plan of the paper is: Section 2 presents the background and data used, Section 3 presents the empirical strategy, Section 4 presents results, and Section 5 closes with a discussion of our findings.

2. Background and Data

That poor air quality leads to poor health outcomes for both adults and children is fairly well established in the literature. The mechanisms by which air quality affects health is usually thought to be through reduced pulmonary functioning leading to acute respiratory infections (Bruce et. al. 2000). To the best of our knowledge the focus in this literature has been on either ambient air pollution or on indoor air pollution, but not both. One of our contributions in this paper lies in examining the relative effects of both sorts of exposure on child health.

We use data from the city sample of the NFHS wave 3 that was collected over 2005 and 2006. For our analysis, we specifically use the child-recode datafile (IAKR51FL.DTA), and the health outcomes analyzed are incidence of fever and cough among children (born in the last 3 year) during the two weeks prior to the interview, as reported by the household respondent. Data on ambient air pollution is taken from the administrative records maintained by the Central Pollution Control Board (CPCB), Government of India, under its National Air Monitoring Program (NAMP). NAMP provides data on four key pollutants for the cities on which we have child-level data over the survey duration 2005-6.[1] We use average monthly levels of Nitrogen Dioxide (NO2), Respirable Suspended Particulate Matter (RSPM / $PM_{2.5}$) and Suspended Particulate Matter (SPM/PM_{10}) as our primary measures of ambient air pollution. We combine this with the NFHS dataset by calculating monthly averages for each source of

[1] CPCB is a statutory body under the Ministry of Environment and Forests. CPCB's primary responsibilities include the prevention, control and abatement of air and water pollution in India.

ambient air pollution and pairing it to each case's month of interview as reported in the NFHS survey (see Figure 1 and 2 to see monthly average, maximum, minimum and NAMP safe levels of different pollutants) . Data on indoor air pollution comes from the NFHS survey where we capture each household's indoor air quality using the type of cooking fuel used by the household. Cooking fuel is believed to be the most important source of indoor air pollution. From the NFHS data we know if households use electricity, LPG, natural gas, kerosene, coal, lignite, charcoal, wood, straw/ shrubs/grass, crop residue, or animal dung as the primary cooking fuel. We classify these into three categories, clean cooking fuel (i.e. electricity, LPG, natural gas), unclean fuel (i.e. kerosene, coal, lignite and charcoal) and unprocessed fuel (wood, straw, crop residue, and animal dung).

3. Estimation Strategy

Let child *i*, living in city *c*, in month *m* have a latent propensity to have fever and cough be captured by $\boldsymbol{h}^*_{\boldsymbol{imc}} = (h^*_{1imc}, h^*_{2imc})$ where the $h*$ represents unobserved latent propensities that are only partially observed. These latent propensities are related to a number of child-specific, household-specific, month-specific and city-specific effects; in these models we are specifically interested in the month- and city-specific ambient air pollution variables. Thus we have:

$$\begin{aligned} h^*_{1imc} &= \beta_1(inAP)_{imc} + \beta_2 f(outAP)_{mc} + \boldsymbol{x}_{ic}\boldsymbol{\beta} + \tau_c + \varepsilon_{1imc}; \\ h_{1imc} &= 1 \text{ if } h^*_{1imc} > 0 \\ h_{1imc} &= 0 \text{ otherwise} \end{aligned} \qquad (1)$$

$$\begin{aligned} h^*_{2imc} &= \gamma_1(inAP)_{imc} + \gamma_2 f(outAP)_{mc} + \boldsymbol{x}_{ic}\boldsymbol{\gamma} + \tau_c + \varepsilon_{2imc}; \\ h_{2imc} &= 1 \text{ if } h^*_{2imc} > 0 \\ h_{2imc} &= 0 \text{ otherwise} \end{aligned} \qquad (2)$$

However, neither h^*_{1imc} and h^*_{2imc} are observed; we only know if they are positive, and thus, we observe (h_{1imc}, h_{2imc}). If we assume that ε_{1imc} and ε_{2imc} are independently distributed N(0,1) then we can estimate two sets of independent probit regressions to estimate the underlying regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. However, this essentially assumes that these two health conditions are not correlated – i.e. the child's likelihood of having a fever is unrelated to having a cough. If the two are in fact correlated, as they are likely to be through unobserved child, household, or environmental attributes, then a bivariate probit model that assumes that the two errors are jointly distributed is more appropriate, i.e.

$$\varepsilon_{1imc}, \varepsilon_{2imc} \sim N(0,0,1,1,\theta) \qquad (3)$$

where θ is the correlation coefficient between the two error terms. This framework, therefore, allows us to explicitly account for the likely correlation between the two child morbidity measures that are affected by the level of air pollution. In equations (1) and (2), $inAP_{imc}$ are a set of dummy variables for the use of coarse bio-fuels for cooking at home as a proxy for indoor air pollution, while $outAP_{imc}$ measures alternative outdoor air pollution measures such as the average level of nitrogen dioxide, RSPM, or SPM.[2] Each of these measures is meant to capture a number of different aspects of the type of ambient air pollution that a child is exposed to. Apart from these measures, we also account for a number of child-specific variables such as sex, age, and health endowments (height and weight); as well as parental or household variables such as religion, education of the parent, social group, and wealth of the household. the τ_c are city-level fixed effects that account for city-specific unobserved attributes that could influence a child's health as well as the air pollution measures but doesn't vary either across households or over time (for example, latitude or geographical proximity to say hills, forests, rivers, or significant sources of pollution, such as a manufacturing hub).

4. Preliminary Results

Child health is known to be quite fragile and particularly so in developing countries. Even in our data, respondents report a high frequency of child morbidity in terms of the incidence of fever and cough that occurred in the two weeks prior to the interview (See Figure 3). Our preliminary estimates suggest that there is substantial correlation between the two health outcomes that is important to account for in the model. Second, in almost all the models, using alternative measures of outdoor air pollution, we find that the coefficients on the variables capturing indoor air pollution are not statistically significant, while those for ambient air pollution are large and statistically significant.[3] In all models, higher levels of NO_2 and RSPM lead to a greater likelihood of falling sick with either fever or a cough, but the incidence of cough is more responsive to ambient air pollution than fever. With a significant correlation coefficient of 0.8 between the error terms in the fever and cough equations, our results suggest that higher levels of outdoor air pollution leads to a greater likelihood of a child falling ill in general.

Simple omitted variable bias has been known to confound many estimates and here we present evidence to suggest that this may also be the case for the literature investigating the effect of indoor air pollution on child health without explicitly controlling for outdoor air pollution. If higher levels of ambient air pollution lead to worse health outcomes and are also correlated with the levels of indoor air pollution, then the estimated effect of indoor air quality on child health is likely to be an overestimate. Our findings suggest that greater emphasis needs to be placed on improving ambient air quality in general as part of urban planning and development, and policies targeted at city-wide reductions in air pollution can significantly improve child health.

---

[2] We use a number of alternative functional forms of these biweekly readings for our analysis – the mean, the standard deviation calculated over the month, the coefficient of variation, and finally, deviations in the monthly average levels from the permissible or safe levels NAPM defined by NAPM.

[3] We used alternative measures for the outdoor air pollution variables – range, mean, standard deviation, and the coefficient of variation, apart from deviation from permissible levels, and we always find ambient air pollution as being significantly associated with a greater likelihood of child morbidity.

References

Bruce, N., Perez-Padilla, R. and Rachel Albalak (2000). Indoor air pollution in developing countries: a major environmental and public health challenge. *Bulletin of the World Health Organization*. 78(9), 1078-1092

Currie J. and M. Neidell. (2005). Air Pollution and Infant Health: What can we learn from California's Recent Experience? *Quarterly Journal of Economics*. 120(3), 1003-1030

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

Duflo,E, M. G. and R. Hanna (2008). Indoor air pollution, health and economic well-being. *Surveys and Perspectives Integrating Environment and Society 1.1*(1).

Frankenberg, E., D. McKee, and D. Thomas (2004): "Health Consequences of Forest Fires in Indonesia," *Demography*, 42(1), 109–129.

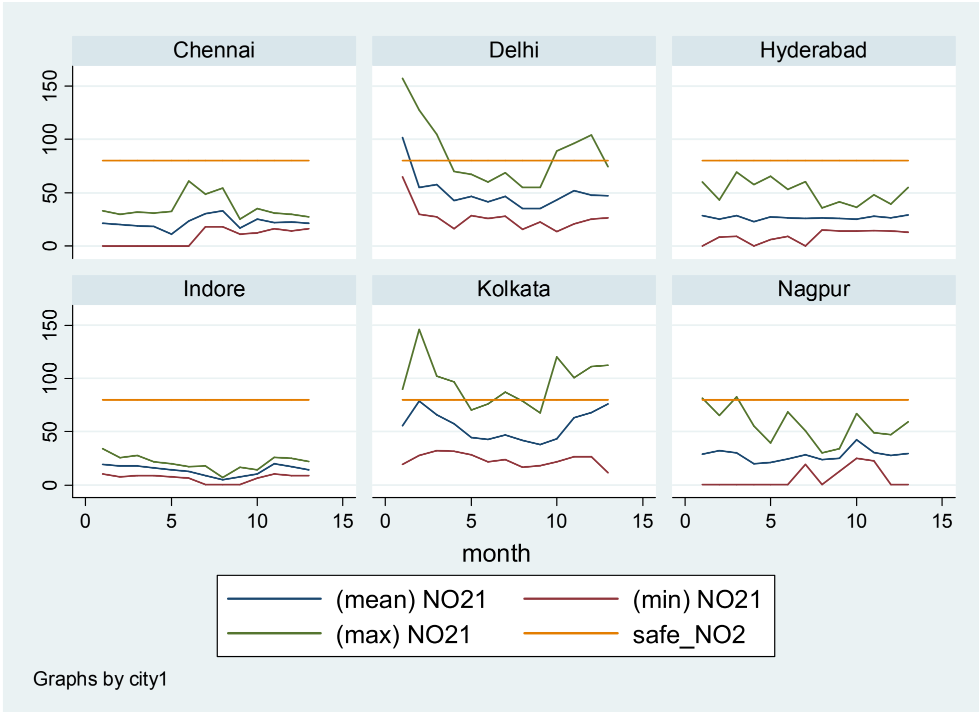
Janke, K., Propper, C. and John Henderson. (2009). Do current levels of air pollution kill? The impact of air pollution on population mortality in England. *Health Economics* 18(9), 1031-1055

Jayachandran, S. (2009). Air quality and early-life mortality during Indonesia's massive wildfires in 1997. *Journal of Human Resources 44*(4), 916–954.

Smith, K. R. (2000). National burden of disease in India from indoor air pollution. *Proceedings of the National Academy of Science 97*(24), 13286–13293.

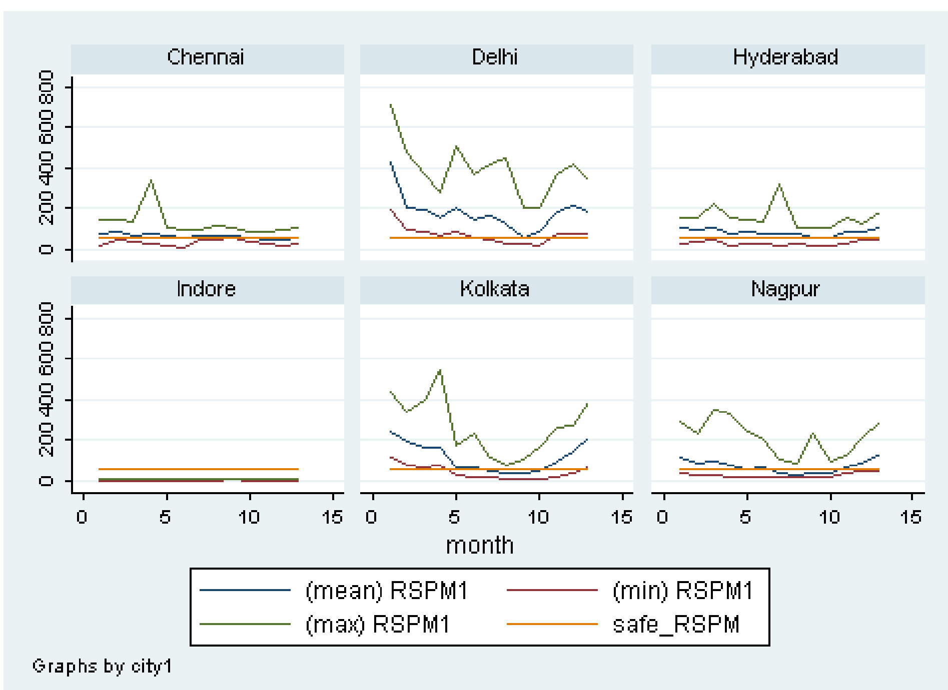
Figures and Tables

Figure 1 Monthly NO2 Distribution over NFHS 3 Interview Months across cities



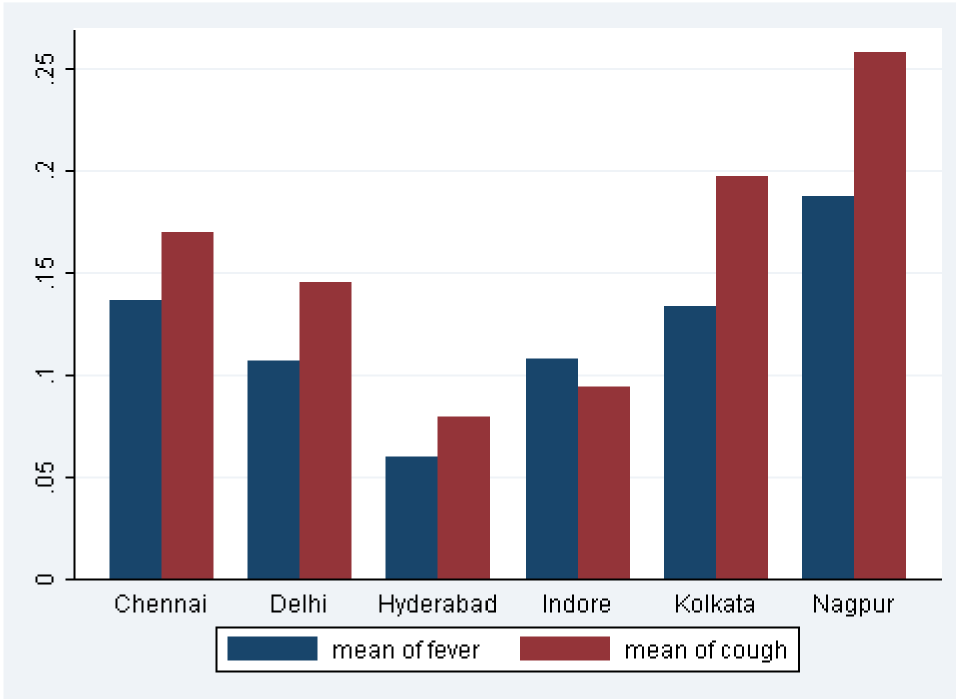
Source: Environmental Data Bank, MOEF, Govt. of India. Safe_NO2 is the level of NO2 that the MOEF advises as being safe. Each pollutant is measured in μm^3.

Figure 2 Monthly RSPM Distribution over NFHS 3 Interview Months Across Cities



Source: Environmental Data Bank, MOEF, Govt. of India. Safe_RSPM is the level of RSPM that is advised as being safe. Each pollutant is measured in μm^3.

Figure 3: City Level Prevalence of Child Health Outcomes



Source: NFHS 3 data