

The Environmental Cost of Air Pollution Due to Stubble Burning: Evidence from New Delhi

Arti Agarwal^{*}, Nachiketa Tiwari

Department of Design, Indian Institute of Technology, Kanpur, India

ABSTRACT

Objective: Stubble burning in North India has been a major contributing factor to the growing menace of air pollution in the National Capital Region (NCR) of India for the last two decades. Though the health aspects of air pollution due to stubble burning have been extensively studied, its economic costs due to environmental damage have not been studied holistically. We attempt to estimate these costs using Instrumental Variable (IV) Analysis.

Methods: Using VIIRS Data from NASA, we count the number of field fires per day in Punjab and Haryana during the September-December harvesting season, called FIRECOUNT. We use FIRECOUNT as the IV to estimate the concentration of PM_{2.5} and PM₁₀ due to stubble burning. This is then regressed against the Gross State Domestic Product (GSDP) of New Delhi to identify the effect of increase in PM_{2.5} or PM₁₀ on the GSDP of New Delhi.

Results: We find that field fires in North India contribute significantly to PM_{2.5} and PM₁₀ concentration in New Delhi, and that an increase in PM_{2.5} by 100 percent is correlated with a decrease in the GSDP of New Delhi by approximately 1 percent.

Conclusion: This loss in GSDP can be in millions of Indian rupees and this externality has strong policy implications for lawmakers in New Delhi.

Keywords: Air pollution; Environmental economics; Environmental pollution; Stubble Burning; Economic costs

INTRODUCTION

Stubble burning, or the practice of burning agricultural residue like straw, leaves, and other vegetation after harvesting a crop, is common all across India. The burning of stubble releases toxic fumes into the air, which worsen the existing state of air pollution in India. Stubble burning is one of the top contributors to air pollution in India. Unlike wildfires and forest fires, stubble burning is a seasonal phenomenon which is undertaken by the farmers intentionally to get rid of the crop residue in a cheap, speedy way.

Although stubble burning is a problem prevalent across India, the National Capital Region (NCR), centred on New Delhi, the capital city of India, is uniquely positioned to face some of the worst air pollution in the country due to its geography, meteorology and surrounding terrain. New Delhi is surrounded

by the states of Punjab and Haryana which have huge expanses of agricultural land. Farmers in these states burn stubble after harvest, starting October, and the toxic fumes and particulate matter (PM_{2.5} and PM₁₀) released as a result of these field fires is carried by strong winds to the National Capital Region within 24 hrs, sometimes lesser. Low winds and higher humidity and fog in the National Capital Region (NCR) during the fall and winter trap the particulate matter and toxic gases close to the ground, leading to thick layers of smog for almost the entire winter season. It mixes with local air pollutants coming from industries and transportation as well. As per estimates, up to 50 percent of the particulate matter in New Delhi during October-December can be attributed to open field fires for stubble burning [1].

This smog is made up of gases like carbon monoxide, sulphur dioxide, nitrous and nitric oxides, ozone, PM_{2.5}, PM₁₀ and

Correspondence to: Arti Agarwal, Department of Design, Indian Institute of Technology, Kanpur, India, Tel: 9851132050; E-mail: arti21@iitk.ac.in

Received: 22-Jan-2023, Manuscript No. JPE-23-21508; **Editor assigned:** 24-Jan-2023, PreQC No. JPE-23-21508 (PQ); **Reviewed:** 07-Feb-2023, QC No. JPE-23-21508; **Revised:** 29-Apr-2024, Manuscript No. JPE-23-21508 (R); **Published:** 06-May-2024, DOI: 10.35248/2375-4397.24.12.400

Citation: Agarwal A, Tiwari N (2024) The Environmental Cost of Air Pollution Due to Stubble Burning: Evidence from New Delhi. J Pollut Eff Cont. 12:400.

Copyright: © 2024 Agarwal A. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

traces of aromatic gases like benzene and toluene. The air quality in the National Capital Region (NCR) is as high as 500 for most part of October-December. This is labelled as “hazardous” as per air quality guidelines and indicates serious illness like asthma, bronchitis and other lung problems as side effects [2]. People are advised to stay indoors. The citizens of the National Capital Region are held hostage to an air calamity every year, with no respite. A major political issue in North India, which involves the state governments of Punjab, Haryana, New Delhi and Uttar Pradesh, and the central government, the issue is as yet unresolved after more than a decade of policies and ostensible methods of redressal. The AQI in the winter of 2022 in New Delhi and its satellite towns remains as high as 500, often reaching 700. To put into perspective, the five days of “the great smog of London” in 1952 clocked an AQI close to 500 for five days, which was considered an environmental emergency, leading to the Clean Air Act of 1956 in the UK. Whereas in NCR, an AQI of 500 and above has been recorded every year for more than a decade, with no policy able to check this phenomenon.

This phenomenon is a catastrophic example of air pollution as an externality. The farmers and residents of Punjab and Haryana burn stubble to reduce their costs of disposal of agricultural waste, while at the same time endangering the health and wellness of millions of citizens of National Capital Region (NCR). The environmental cost of this contribution to air pollution *via* stubble burning is essential to establish the non-viability of this practice on health, environmental and economic grounds. Our study solves this precise problem. By pinning a cost to the damage caused by the stubble burning fires, we make a cogent, plausible case for implementation of methods to stop stubble burning. The cost-benefit analysis of any method that policymakers propose to curb stubble burning, must account for the entire economic cost of the phenomenon, rather than considering the costs purely due to health expenditure or loss of life. So far, no study has incorporated the externality of stubble burning as an environmental cost, and as such, many measures which could possibly help stop this practice have not been found to be cost-effective. While enough literature exists on the environmental costs of forest fires and wildfires, not enough literature is as yet documented on the environmental costs of open field fires due to stubble burning, especially as a heavy externality. Our study fills this lacuna in environmental and economic literature.

This paper estimates the economic cost of stubble burning in North-West India as externality for New Delhi, to add to the discussion on framing the right policies for addressing the challenge by state and central governments. We use Instrumental Variable (IV) analysis to estimate the economic cost to the city of New Delhi due to stubble burning alone, caused by air pollution. The instrumental variable is the daily count of number of field fires in the region of Punjab and Haryana, taken from NASA Active Fire Data [3]. The number of field fires is found to be correlated with the level of PM_{2.5} and PM₁₀ in New Delhi. Then the Gross State Domestic Product (GSDP) of New Delhi is regressed against the PM_{2.5} parameter to estimate the effect of pollution on the GSDP of New Delhi. Instead of including the entire National Capital Region (NCR),

only the city of New Delhi is taken in the IV analysis for the exclusion restriction to hold and because the city-level economic data is available only for the city of New Delhi.

The purpose of this study is to estimate economic costs due to air pollution from crop residue burning, and suggest policy design to not only address this challenge but nudge and incentivize innovation in this field. This study can be used further by researchers as a model to address similar problems of air pollution externalities in other regions around the world. We develop the theoretical and empirical framework to tackle the problem of air pollution externality. Since it's not possible to physically separate or demarcate between air pollutants based on their source, we take an alternative route. This study solves that problem by using an instrumental variable which is correlated with the pollutant variable, and does not affect the observed variable through any channel other than the pollutant variable.

The effects of air pollution on human health have been studied extensively. An estimated 1.67 million deaths were attributed to air pollution in India in 2019 [4], making up for 17.8 percent deaths in the country. High PM_{2.5} and PM₁₀ levels in Delhi have been found to be associated with higher levels of respiratory problems in children [5]. Both short-term and long-term exposure to high levels of PM_{2.5} and NO₂ were also found to be correlated with higher infection rates and mortality rates of COVID-19 [6]. As per a time-series analyses in Chennai, a 0.44 percent increase in daily all-cause mortality was seen per 10 pg/m³ increase in daily average PM₁₀ concentrations [7]. A study estimated the economic gains from lower levels of air pollution during lockdown in Delhi in 2020 to be USD 64.3 million, leading to lower damages due to morbidity [8]. Another study estimated the economic cost of air pollution due to health impacts in Agra city to be USD 254.52 million in 2014 and projected to be USD 570.12 million in 2020 [9].

We study existing literature on usage of econometric methods to evaluate economic costs and air pollution levels. Instrumental variable analysis [10] has been used to study atmospheric and aerosol chemistry. The economic cost of air pollution in Greece has been estimated using Cost-Of-Illness (COI) and Willingness-To-Pay (WTP) methods [11]. The economic cost of air pollution on health in Singapore was studied using damage function/dose response approach [12]. The relationship between economic growth, air pollution and life expectancy in Indonesia was studied using the Autoregressive Distributed Lag (ARDL) model [13]. The findings of all of these studies suggest that there is a negative correlation between changes in air pollution levels and change in economic output.

Discussions and debates around policies to control stubble burning have so far involved the cost-benefit analysis of mitigation [14]. Though this study elucidates how mitigation of pollution is a viable option, it has not factored in the effect of pollution externalities which directly affect economic output *via* multiple channels. Causes that led to the widespread adoption of paddy-burning in Haryana have been studied and identified in a study [15]. As per the study, limitations imposed by mechanized harvesting, falling groundwater levels, and high labour costs were some of the reasons why farmers use stubble burning to dispose of the agricultural residue, irrespective of

farm size or variety of the crop. The study suggests that better management practices are required, as opposed to mere institutional reforms.

One study estimated that the share of biomass burning in the pollutant levels in the megacity Delhi is between 1 percent and 58 percent [16], depending on the quantity of biomass burnt and meteorological factors like wind speed and thermal inversions. As per the National Clean Air Programme (NCAP) report by the Ministry of Forests, Environment and Climate Change (MOEFCC) government of India, open burning of agricultural residue in rural areas contributes about 7 percent to the total PM_{2.5} emissions in the country [17].

As per the Porter hypothesis, environmental regulations and policies that aim to reduce pollution levels may lead to technological innovations and thus further economic growth [18]. There is evidence to suggest that this is possible, and achievable through technological innovation. Best practices in paddy residue management have been identified [19,20], which are expected to not be an economic burden on the farmer and lead to reduction in stubble burning. The most viable methods to dispose of the paddy stubble have been pyrogenic conversion of rice straw [21], *in situ* management of the stubble through microbial biodegradation [22], and biochar production from the stubble [23] to name a few. These practices, if implemented *via* effective policies, can be a long-term solution to this problem.

Effect of air pollution on economic output

Air pollution can potentially impact economic output through multiple channels. Though the focus of the recent literature has been the empirical assessment of the impact of air pollution on the health of citizens, and the associated economic losses, it is important to also understand the theoretical underpinnings of this analysis. Our analysis looks at the economic impact of air pollution holistically, enumerating the various channels *via* which the economy is affected. Moreover, the impact of air pollution on the health as well is not merely due to life threatening diseases, but due to less serious but persistent illness and allergies as well, reducing the productivity of the labour force in multiple ways [10]. In a developing country like India, air pollution is not a mere environmental concern, but also a developmental concern as it affects the labour force participation and other factors of production. This understanding strengthens our argument of the need to mitigate severe air pollution due to crop residue burning.

Extensive margin of the labour workforce

Air pollution can contribute to reducing the extensive margin of the labour workforce *via* migration effects, mortality rates and birth rates. High levels of air pollution can contribute to higher level mortality rates and lower birth rates. It can also affect migration out of the region, due to high costs to health in the region. All of these effects lead to the lowering of the total number of workers available in the region [24].

Worker Productivity: High air pollution levels can lower the productivity of workers while they are at work. They may

experience discomfort, mild illness or allergies etc, which lead to less output by workers [25,26].

Worker Attendance: Air pollution can also affect the intensive margin of labour. Attendance of workers is affected by their health levels, which is affected by pollution levels. Higher levels of air pollution can potentially reduce worker attendance. Workers would be impelled to take more sick leave days, thus lowering the total time spent on work. This would lead to lower output by firms [27,28].

Capital Depreciation: Environmental effects like air pollution can cause short-term or long-term damage to natural as well as man-made capital. Machinery and equipment may face higher wear and tear with higher levels of pollution. Natural resources like tree cover, soil quality and water quality may also be compromised. Thus, air pollution can contribute to faster depreciation of all types of capital inputs, thus increasing the investment required for renewing the capital, and lowering economic output [29].

Direct Effects: The direct impact of air pollution on economic output is also important, though often ignored. High pollutant levels may directly interfere with certain industrial processes, reducing their efficiency and productivity, leading to lower outputs. The pollutants may react with certain substances and materials, making air pollution a significant factor in reduction of economic output.

Theoretical Model: To model the impact of air pollution on economic output Y , we build on the theoretical model given by Antoine Dechezlepretre, et al. [31] for evaluating the economic cost due to air pollution. We consider the output for the city of New Delhi, taking it as a closed economy. For the purpose of this analysis, taking New Delhi as a closed economy does not drastically alter our findings, since the practice of stubble burning does not alter the GSDP of New Delhi *via* any price or supply changes of goods "imported" into New Delhi from neighbouring states.

The overall output in a closed economy is given by $Y=F(K,L,P)$

Where:

Y is the total economic output,

K is the capital input,

L is the effective labour input,

P is pollution.

We take N as the number of representative households (which are assumed to supply labour inelastically).

Each representative household is endowed with a total time t for work. For simplicity, we ignore the effect of choosing leisure in this model and assume the total time available for work, t , to be the same for all households. Of this time t , the households spend time h for labour, and s on sick leave, when they are not capable of work. In this model, pollution is taken as exogenous to the model [17].

Thus $s=s(P)$ and $t=h+s(P)$.

Since P potentially affects the mortality rate, migration rate and birth rate of the population, thus affecting the total number of households available for work, we have $N=N(P)$.

The productivity of the workforce is defined as $A=A(P)$.

Thus, the effective labour available for work is

$$L(P)=N(P)A(P)[t-s(P)] \quad (1)$$

We use this relation for labour in the equation for economic output. Thus:

$$Y=Y(K,L(P), P) \quad (2)$$

Taking natural logarithms on both sides of the equation,

$$\log Y = \log(Y(K, L(P), P)) \quad (3)$$

Since logarithm is a monotonic transformation of the function Y, differentiation of Y on Left Hand Side (LHS) and Right Hand Side (RHS) of the equation w.r.t variable P will yield same results as differential of logY on LHS and RHS w.r.t P. Thus, the impact of a change in pollution on the economic output is given by differentiating the equation w.r.t P as follows,

$$\frac{d \log Y}{dP} = \frac{\partial \log Y(\cdot)}{\partial K} \frac{\partial K}{\partial P} + \frac{\partial \log Y(\cdot)}{\partial L} \frac{\partial L}{\partial P} + \frac{\partial \log Y(\cdot)}{\partial P} \quad (4)$$

Simplifying,

$$\frac{d \log Y}{dP} = \frac{\partial \log Y(\cdot)}{\partial K} \frac{\partial K}{\partial P} + \frac{L}{Y} \frac{\partial Y(\cdot)}{\partial L} \left[\frac{1}{L} \frac{\partial L}{\partial P} \right] + \frac{\partial \log Y(\cdot)}{\partial P} \quad (5)$$

The elasticity ψ of economic output with respect to effective labour is given by;

$$\psi = \frac{\frac{\partial Y(\cdot)}{Y}}{\frac{\partial L}{L}} \quad (6)$$

Thus, combining equations (1), (5), and (6) and simplifying, we get:

$$\frac{d \log Y}{dP} = \frac{\partial \log Y(\cdot)}{\partial K} \frac{\partial K}{\partial P} + \psi \left[\frac{\partial \log N}{\partial P} + \frac{\partial \log A}{\partial P} + \frac{\partial \log(t-s(P))}{\partial P} \right] + \left[\frac{\partial \log Y(\cdot)}{\partial P} \right] \quad (7)$$

The expression $\frac{d \log Y}{dP}$ on the LHS captures the total effect of pollution on economic output through the multiple channels discussed above, considered on the RHS. The first expression on RHS captures the effect of air pollution on capital, which in turn affects the economic output Y. The expression $\psi \frac{\partial \log N}{\partial P}$ measures the effect of pollution on the extensive margin of the labour force. The expression $\psi \frac{\partial \log A}{\partial P}$ measures the impact of pollution on worker productivity. The expression $\psi \frac{\partial \log(t-s(P))}{\partial P}$ measures the impact of pollution on worker attendance. The last term, $\frac{\partial \log Y(\cdot)}{\partial P}$ is the direct impact of pollution on production through other channels, like direct interference with a chemical process in an industry.

Through econometric analysis, we estimate the sign and magnitude of the expression on the LHS.

MATERIALS AND METHODS

Econometric Model

The objective is to evaluate the causal effect of air pollution due to stubble burning in Punjab and Haryana on the economic output of the national capital city of New Delhi.

In this study, we use regression analysis with instrumental variable analysis to account for confounding effects of other variables and reverse causality. The basic econometric model that can be used to characterize the relationship between economic output and air pollution levels due to stubble burning at time t can be stated as follows:

$$\log Y_t = \beta_0 + \beta_1 \log P_t + \beta_2 \gamma_t + \varepsilon_t \quad (8)$$

Where:

Y_t =Variable that denotes the economic output at time t,

β_0 =Captures the fixed effects of the city,

P_t =Pollution level at time t,

γ_t =Time-varying impacts on economic output in the region due other factors,

ε_t =Error term.

In this model, a few constraints need to be addressed. First is the reverse causality of increase or decrease in economic output on air pollution levels. In the past it has been observed that when pollution levels in New Delhi are very high, policy interventions cause certain businesses and schools to shut down or reduce the working hours [32] reducing pollution from industries. This further causes reduction in road traffic and air pollution from vehicular emissions as well. All of these factors contribute to lowering of air pollution levels due to ambient effects. Thus, there is a reverse causality between economic output and pollution due to local factors. This can be characterized as follows:

Economic activity $\uparrow \Rightarrow$ Air pollution levels \uparrow

Air pollution levels $\uparrow \Rightarrow$ Economic activity \downarrow

Economic activity $\downarrow \Rightarrow$ Air pollution levels \downarrow

Due to this cyclical causality, a simple regression with pollution level, in spite of other controlling variables, is unlikely to give an unbiased estimate of changes in economic output. Therefore, to measure the impact of air pollution due to stubble burning, which is not a local phenomenon, we need to use a different methodology.

Instrumental Variable Analysis

To address this constraint, we use instrumental variable analysis [33]. Since the phenomenon whose causal effect we wish to estimate is the burning of crop residue in North-West India, it is ideal to take the number of field fires at time t as an instrumental variable. We call this variable as FIRECOUNT. Open field fires in north-western regions of Punjab and Haryana

emit pollutants into the air, and the wind carries these pollutants to the national capital city of New Delhi. The number of fires at any given time t is a good indicator of the concentration levels of pollutants present in the air. The FIRECOUNT satisfies the requirements of an instrumental variable, namely:

- Exclusion restriction: It affects the dependent variable (economic output) only through the explanatory variable (pollution levels in New Delhi).
- It is correlated with the explanatory variable (pollution levels in New Delhi).
- It is not caused by the dependent variable (economic output of New Delhi), and is thus exogenous, and as good as randomly assigned.

FIRECOUNT, or the number of field fires in Punjab and Haryana, can be safely assumed to have no effect on the Gross State Domestic Product (GSDP) of New Delhi *via* any channel other than air pollution. The fires are in different states, and thus, are not correlated with or caused by economic activity in New Delhi. FIRECOUNT does have a direct correlation with the air pollution levels in New Delhi, as already found by numerous studies and satellite information. As visible in Figures 1 and 6, the variation of PM2.5 in New Delhi with time and the variation of FIRECOUNT in north-west India with time, both follow the same pattern every year, for all stations.

First stage: Since the effect of wind speed cannot be ignored on the concentration of pollutants in the air, wind speed is taken as a control variable in measuring the effect of field fires. Other control variables are meteorological parameters like Relative Humidity (RH), Wind Direction (WD) and Air Temperature (AT). Pollution levels are measured for PM2.5 or PM10 .

So, the first stage of the analysis (taking wind speed as a control variable) would be:

$$P_t = \alpha_0 + \alpha_1 FIRECOUNT_t + \alpha_2 WS_t + \theta_t \tag{9}$$

Where;

P_t =Pollution level at time t ,

α_0 =Captures fixed effects,

FIRECOUNT=Total number of field fires at time t ,

WS_t =Wind speed at time t (control variable),

θ_t =Error term at time t .

Second stage: A lot of variables affect the GSDP of a city: Economic and natural shocks, changes in world economic activity, changes in population demographics etc to name a few. It is important to control for these effects so as to be able to measure the effect of pollution levels on economic output without omitted variable bias. Since a lot of data of these variables is not available for the city of New Delhi, an index which is computed considering many such effects is instead used as a control variable-the subnational Human Development Index (HDI) for New Delhi by UNDP [34]. The subnational

HDI accounts for variation in changes in socio-economic structures with time γ_t . We use the Gross State Domestic Product (GSDP), GSDP per capita and Gross State Value Added (GSVA) of New Delhi as measures of economic output Y_t in the model. The second stage is thus given by:

$$Y_t = \beta_0 + \beta_1 P_t + \gamma_t + \epsilon_t \tag{10}$$

Where:

Y_t =Economic output,

β_0 =Captures fixed effects,

P_t =Pollution level at time t (as evaluated from the first stage regression),

Γ_t =Captures time-varying effects on economic output (control variable),

E_t =Error term.

Reduced form

We use “Two Stage Least-Squares (2SLS)” to compute the coefficient β_1 . To implement this, we first estimate the first stage, and then compute the predicted values of the regressor:

$$\hat{P}_t = \alpha_1 FIRECOUNT_t \tag{11}$$

Then we regress the economic output on the predicted values of pollutant levels: In a large sample, the coefficient thus got will be β_1 .

$$Y_t = \beta_0 + \beta_1 \hat{P}_t + \beta_2 \gamma_t + \epsilon_t \tag{12}$$

Due to a vast difference in the range of values in pollutant levels and GSDP, a log-likelihood estimation method is most appropriate for the analysis. Therefore, we instead use the log of GSDP and pollution levels as variables in the regression. Rewriting equation (12) to account for this change, we have:

$$\log(Y_t) = b_0 + b_1 \log(P_t) + b_2 \Gamma_t + \epsilon_t \tag{13}$$

The coefficient thus obtained, b_1 , gives the elasticity of economic output with respect to pollutant levels. A 1 percent increases in the levels of the pollutant causes a b_1 percent change in the economic output.

Air Pollution Data

For this study, air pollution data for nine years—from 2012 to 2021 was collected from fixed-site monitoring stations of the central and state pollution control boards in New Delhi. Data of 24 hr averages for PM2.5, PM10, CO, NO₂ and SO₂ were collected for five monitoring stations spread across the city:

Anand Vihar, ITO, Mandir Marg, Shadipur and R. K. Puram. For other monitoring stations, archival data for pollution levels was not available. Thus, these monitoring stations were selected. Since archival data of all the years for all the parameters was not available for every monitoring station, the years for which the data was available for each station was taken on a case-by-case basis. Each monitoring station records data independently of all other stations, and as such, the results from the analysis of the data of each station are independent of other stations.

A plot of the variation of pollution levels of different pollutants with time is seen in Figure 1. As stubble burning is a yearly phenomenon which takes place from September to December every year, the trend in increase in pollutant levels during this period is clearly visible in the plot. The gaps in the plot are the periods for which data for pollutant levels is not available. Time variation of PM2.5 levels for the different monitoring stations is seen in Figure 2. PM2.5 and PM10 concentrations were measured in $\mu\text{g}/\text{m}^3$, CO in mg/m^3 , NO_2 in $\mu\text{g}/\text{m}^3$ and SO_2 in mg/m^3 .

A correlation plot between the pollution parameters showed a high correlation between PM2.5 and PM10 values. Mild correlation was also observed between other pollutant values, as seen in Figure 2. PM2.5 was found to be the most correlated with FIRECOUNT values (Corr: 0.3), followed by PM10. Other parameters had low correlation values with FIRECOUNT, and as such, were not selected to be the endogenous variable. The data for other parameters was also found to be not as well documented as that for PM2.5 and PM10.

The relationship between FIRECOUNT and log (PM2.5) for station 1 can be seen in Figure 3. The relationship between FIRECOUNT and log (PM10) for station 4 can be seen in Figure 4.

Since PM2.5 and PM10 were found to be highly correlated, the two parameters are not used together in the regression, so as to avoid the multicollinearity effect. Instead, the two parameters are used as alternatives to measure the effect of fires on air pollutant concentrations.

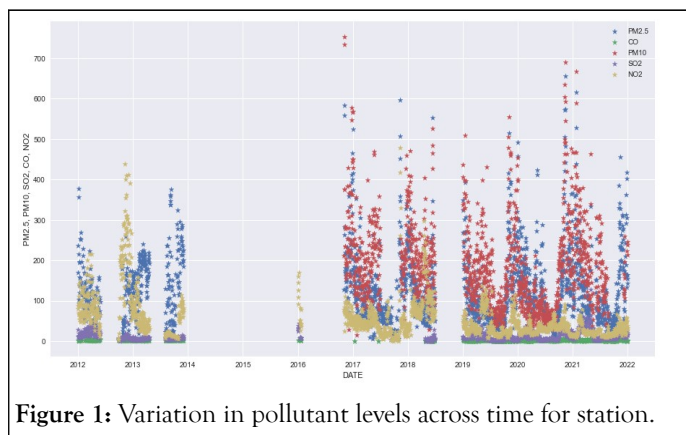


Figure 1: Variation in pollutant levels across time for station.

The objective of the study is to measure the effect due to stubble burning, and PM2.5 and PM10 both can be considered to be alternative variables for that effect. For some of the stations, PM10 data is missing for several months, and as such, PM2.5 is taken as the common variable for all regressions, and PM10 is used to verify the results, where the data is available.

Crop residue fire data

To get the crop residue fire data, Active Fire Data by Visible Infrared Imaging Radiometer Suite (VIIRS) taken by Earthdata by NASA was obtained for the years 2012-2021 [3]. The data is for the region of South Asia. We take the regions where crop residue burning takes place as the region lying between the latitudes and longitudes of 28.9 N and 34 N and 73 E to 77 E respectively. The data for this region was filtered from the larger dataset. Next, the number of fires for each day was counted for this filtered region. This yields a daily count of fires in the regions of Punjab and Haryana. Data for the months of September-December was further filtered out, since those are the months of interest when crop residue is burnt. This is the data that is used in the regression analysis.

The variable FIRECOUNT gives the daily count of number of open field fires thus obtained. The yearly repeating patterns and variations in the number of field fires per day can be in Figure 6.

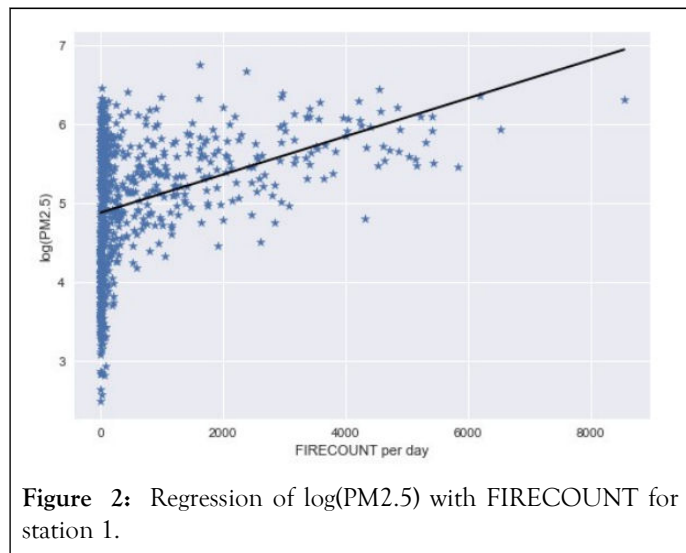


Figure 2: Regression of log(PM2.5) with FIRECOUNT for station 1.

Though FIRECOUNT does not reflect the intensity or extent of the field fires, the variable was found to be acceptable due to the averaging effect of small and big fires. Moreover, FIRECOUNT values vary from as low as 1 and 2 per day up to a maximum of 8000 fires per day. Due to this broad range of values, the peak season can be clearly identified in the data, usually in the months of October and November. Thus, FIRECOUNT was found to be a good variable to characterize the effect of stubble burning in the region.

Weather Data

Weather data for New Delhi is obtained from fixed-site monitoring stations of the central and state pollution control boards in New Delhi. Data for Relative Humidity (RH), Wind Speed (WS), Wind Direction (WD), Spectral Radiance (SR), Barometric Pressure (BP) and Air Temperature (AT) was compiled for 24 hr averages of values for the years (2012-2021).

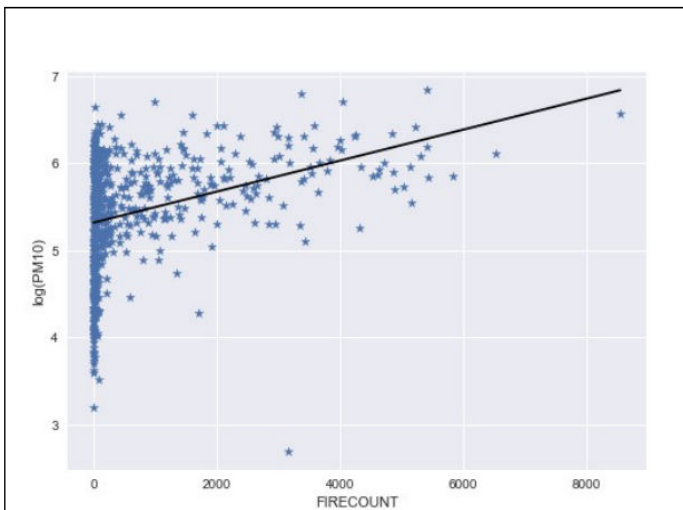


Figure 3: Regression of log (PM10) and FIRECOUNT for station 4.

2012-2021. Data for the months of September-December for every year was used for the regression analysis.

On regressing PM2.5 with each of the weather variables as control variables, not much difference was found on the coefficient of FIRECOUNT. Thus, the other variables are not significantly correlated with PM2.5. The main contributor to the changes in the coefficient was Wind Speed (WS), which was found to be negatively correlated with the pollution levels. This is expected since high wind speeds cause the winds to carry away pollutants from the capital city, thus lowering the pollutant levels in the air. When wind speeds are low, pollutants stagnate in the air above the ground, mixing with the fog and dust to form a thick layer of smog above the ground, which is not only harmful for health, but also reduces visibility on the ground.

Thus, Wind Speed (WS) was found to be the ideal control variable upon preliminary evaluation of the available data.

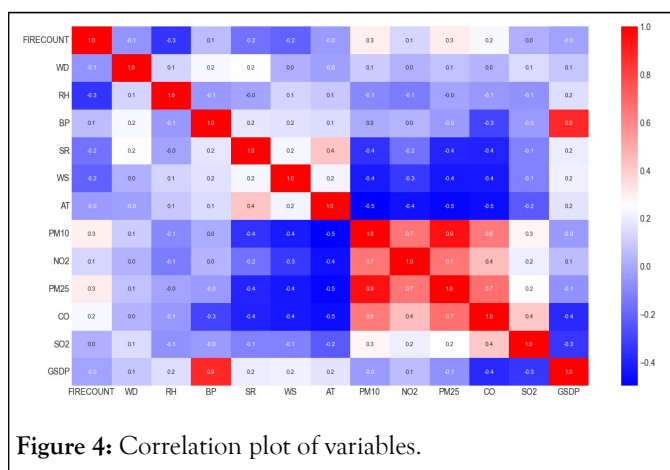


Figure 4: Correlation plot of variables.

Economic data

The economic data for New Delhi was available as yearly values. The GSDP, GSDP per capita and GSVA values were taken from the New Delhi government portal [35], reported in INR crores at current prices. To control for variations in the economic output due to other socio-economic factors, the subnational

Human Development Index (HDI) for New Delhi is used as a control variable. HDI for every year, from 2012 to 2019 is taken from the Global Data Lab website [34]. The subnational HDI was created by Global Data Lab using three sources: Statistical offices (including Eurostat, the statistical office of the European Union), the Area Database of the global data lab, and data from the HDI website of the Human Development Report Office of the United Nations Development Program. Thus, it can be considered to be a robust measure of socio-economic effects on the economy. It is reported as an index, with $0 \leq \text{HDI index} \leq 1$.

Since pollutant concentration values will not be negative or zero, and to account for the wide range of values in pollutant concentrations and economic output, logarithm of pollutant concentration was regressed against the logarithm of economic output in the second stage regression (Figure 5).

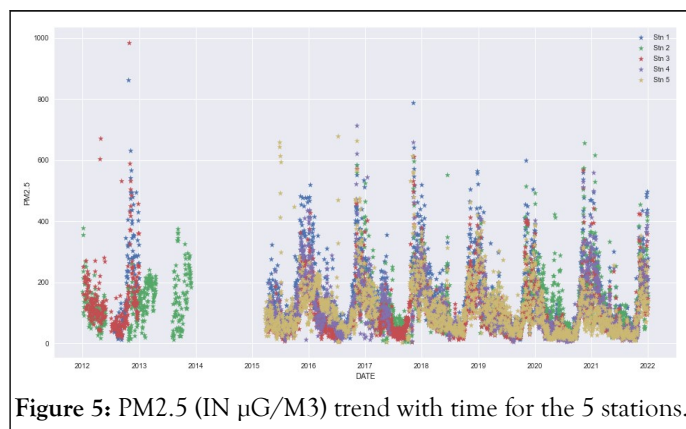


Figure 5: PM2.5 (IN µG/M3) trend with time for the 5 stations.

As expected, the FIRECOUNT data was found to have zero correlation with the economic data of GSDP, also seen in the correlation plot in Figure 4. This further validates our assumption that the exclusion restriction for the instrumental variable holds for this study.

Data Processing and Analysis in Python

The air pollution data and weather data obtained from central and state pollution control boards is 24 hr average data for the pollutant parameters. The crop residue burning FIRECOUNT data is counted for 24 hr periods, and is thus, a total count of the number of fires in the 24 hrs on that date. The weather data follows the pattern of the air pollution data, as it is from the same source. But the weather data is not available for every monitoring station in New Delhi. Thus, the weather data is taken from a monitoring station in central part of New Delhi, and assumed to be the same throughout Delhi, taking the effects of micro-climates to be negligible. This is a reasonable assumption since the variations in meteorological data across New Delhi are small.

Though we take PM2.5 and PM10 as the explanatory variables for pollution, the economic cost of air pollution due to crop residue burning can well be due to other pollutants as well, and the study estimates the cost due to all the pollutants (Figure 6).

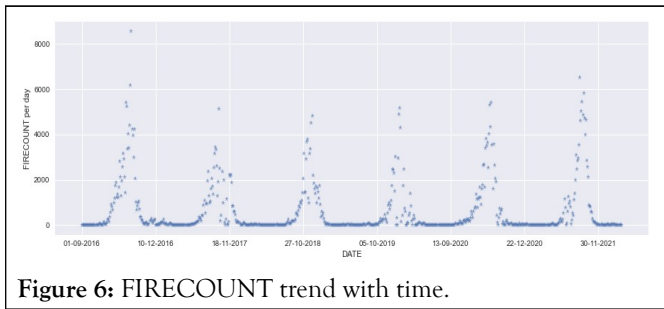


Figure 6: FIRECOUNT trend with time.

The possibility of a lag between field fires and rise in pollutant levels over New Delhi was considered. But it was found that the lag between lighting a fire in northwestern regions and the pollutants reaching the capital city is not more than 24 hours, based on the wind speed and other meteorological effects. Moreover, some of the fields are much closer to the city, whereas some of the fields are further way. On average, the fields can be considered to be within a 24 hr lag from lighting fires and rising of pollutant concentration levels in New Delhi. Thus, the effect of lag on the pollution data was taken to be negligible.

The economic data is reported by government and international organizations yearly. No data for GSDP or GSVa of Delhi could be found for any period shorter than one year. To merge this data with the rest of the data, the daily pollution, weather and FIRECOUNT data would have to be averaged yearly, or the GSDP can be taken as constant for every day in one year. Averaging the pollutant and other data would fail to capture the trends and variations in the pollutant levels with time. Averaging would remove all the variation of pollutant levels with increase and decrease in FIRECOUNT. Moreover, it would reduce the data to merely that of 9 years (six years for some of the monitoring stations for which data is not available for every year). This would not be sufficient to run a regression and arrive at results with any level of accuracy. Thus, the GSDP and HDI were instead taken as constant for every single day in a year, and the entire dataset for every monitoring station was regressed as discussed above.

We run two regressions (first stage and second stage) for each set of variables (FIRECOUNT, PM2.5, GSDP and control variables) for each of the 5 monitoring stations' data. We then check the results by replacing PM2.5 with PM10 and see if there is a remarkable difference in values. We then again check the results by replacing GSDP with GSVa and GSDP per capita.

And see if there is a remarkable difference in values reported. Thus, the reported results are a compilation of the results from 20 regressions, 4 each for the 5 monitoring stations.

RESULTS AND DISCUSSION

The results of the first stage regression show a consistent positive correlation between FIRECOUNT and PM2.5 levels, and a negative correlation between WS and PM2.5 levels. The same trend is observed for data across all the five monitoring stations. The values of the coefficients and standard errors for all the variables and intercept fall in the same range for all the five monitoring stations. The data of each of these monitoring stations is collected by sensors independent of other sensors, and the number of observations for each station is also different. Some of the stations are missing the data for certain years or months, and as such, the time period over which the data is available and used for each station is different. Nevertheless, the results for the regression fall within an acceptable range for all coefficients, and are statistically significant at the 99.9 percent level.

The OLS for PM10 and FIRECOUNT yields similar results, with a slightly wider variation in the coefficients (than those obtained in the regression with PM2.5), but still within acceptable levels. The overall trend of all the values for all the monitoring stations is the same, and all are statistically significant at the 99.9 percent level.

The output of the first stage OLS is in keeping with the theoretical understanding of the phenomenon. The coefficients of FIRECOUNT fall in a narrow range and are statistically significant, so the predicted values of PM2.5 and PM10 levels can be taken as the pollution levels due to crop residue burning, and plugged into the second stage OLS regressions.

The second stage OLS results are in alignment with our theoretical framework to assess the impact of pollution on economic output. The coefficient for pollutant levels is negative, implying that **the result of an increase in pollutant levels is a decrease in economic output**. The results for the second stage regression for log (PM2.5) and log (GSDP) are statistically significant at the 99.9 percent level (Table 1-7).

Table 1: First stage OLS for PM2.5 and FIRECOUNT.

First stage OLS for PM2.5 and FIRECOUNT					
	STN 1	STN 2	STN 3	STN 4	STN 5
Intercept	145.5626 (5.542) ^{***}	111.5125 (5.435) ^{***}	106.2443 (4.713) ^{***}	119.6057 (5.226) ^{***}	106.6421 (4.810) ^{***}
WS	-139.7994 (12.519) ^{***}	-103.4397 (11.500) ^{***}	-102.2080 (10.021) ^{***}	-118.2321 (11.113) ^{***}	-107.3008 (10.230) ^{***}
FIRECOUNT	0.0345 (0.004) ^{***}	0.0286 (0.003) ^{***}	0.0266 (0.003) ^{***}	0.0233 (0.003) ^{***}	0.0219 (0.003) ^{***}

No. of observations	753	646	702	643	666
---------------------	-----	-----	-----	-----	-----

Standard errors in parentheses. * p<. 1, ** p<.05, *** p<.01

The second stage OLS regression results yield a coefficient of approximately -0.01 for log (PM2.5) as regressor for log (GSDP). This implies that a 1 percent change in PM2.5 levels would cause the GSDP to decrease by 0.01 percent. Thus, a 100 percent increase in PM2.5 levels would cause the GSDP of New Delhi to decrease by 1 percent. During the peak stubble burning season, high values of PM2.5, exceeding 500 µg/m³ are often observed. Not only is this air toxic to breathe, this value of PM2.5 concentration is almost five times or five hundred percent of the PM2.5 values during the rest of the year (also visible in the yearly variation of pollutant levels in Figure 5). As per the relationship established from the OLS results, **this could**

cause up to 5 percent decrease in economic output of New Delhi, measured in GSDP at current prices. For context, 5 percent of the GSDP of New Delhi is approximately INR 39,915 crores (at current prices) for the year 2020-2021. This analysis is merely to illustrate how changes in air pollutant levels due to stubble burning can affect the economic output of New Delhi as per our findings. If the values of pollutants remain at high levels for long periods of time, they could potentially cause a higher decrease in GSDP in the coming years.

Table 2: First stage OLS for PM10 and FIRECOUNT.

First stage OLS for PM10 and FIRECOUNT					
	STN 1	STN 2	STN 3	STN 4	STN 5
Intercept	270.4192 (9.659)***	152.6973 (7.026)***	144.2370 (5.735)***	193.1273 (7.271)***	237.9926 (11.007)***
WS	-220.7986 (20.417)***	-115.3630 (14.519)***	-108.6652 (12.200)***	-156.9560 (15.363)***	-255.0205 (24.338)***
FIRECOUNT	0.0487 (0.006)***	0.0326 (0.004)***	0.0367 (0.003)***	0.0356 (0.004)***	0.0430 (0.006)***
No. of observations	624	473	702	646	244

Standard errors in parentheses * p<0.1, ** p<0.05, *** p<0.01

To check the robustness of the findings, we run the second stage regression again, changing the variables. Similar results are obtained by replacing PM2.5 with PM10 and GSDP with GSPA or GSDP per capita (Tables 3-7). In each regression, the coefficients of log (PM2.5) and log (PM10) have a negative sign and are in the same order of magnitude for all monitoring stations. The coefficient of HDI also follows the same sign and order of magnitude in each regression for all stations. The results are statistically significant. Data from each monitoring stations is independent of all other stations; as such, each

regression is an independent finding. Since the findings are consistent across all the variables, it is concluded that the results strongly suggest a negative impact of air pollution due to stubble burning on the economic output of New Delhi, and that the change in economic output corresponding to a one percent change in pollutant levels due to stubble burning is approximately -0.01 percent.

Table 3: Second stage OLS for log(PM25) and log(GSDP).

	STN 1	STN 2	STN 3	STN 4	STN 5
CONST	2.2618*** (0.1321)	0.5434** (0.2614)	1.2584*** (0.2328)	1.3102*** (0.2137)	1.2871*** (0.2197)
HDI	21.3598*** (0.1681)	19.0148*** (0.3494)	16.5808*** (0.3105)	16.5515*** (0.2836)	16.5825*** (0.2934)
log_PM25_fit	0.0112** (0.0044)	0.0092* (0.0047)	0.0086** (0.0037)	0.0129*** (0.0038)	0.0139*** (0.0035)
Rsquared	0.9725	0.8824	0.8661	0.9001	0.8879

R-squared Adj.	0.9724	0.8818	0.8655	0.8996	0.8873
R-squared	0.97	0.88	0.87	0.9	0.89
No. observations	509	402	457	399	419

Standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

Table 4: Second stage OLS for log(PM2.5) and log(GSDP per capita).

	STN 1	STN 2	STN 3	STN 4	STN 5
CONST	0.0478 (0.1091)	1.5168*** (0.2136)	2.9905*** (0.1901)	3.0328*** (0.1746)	3.0139*** (-0.1794)
HDI	17.3226*** (0.1389)	15.3188*** (0.2855)	13.3282*** (0.2536)	13.3042*** (0.2317)	13.3296*** (0.2397)
log_PM25_fit	-0.0092** (0.0036)	-0.0076* (0.0039)	-0.0070** (0.0030)	-0.0106*** (0.0031)	-0.0113*** (0.0029)
R-squared	0.9715	0.8794	0.8624	0.8972	0.8847
R-squared Adj.	0.9713	0.8788	0.8618	0.8966	0.8841
R-squared	0.97	0.88	0.86	0.9	0.88
No. observations	509	402	457	399	419

Standard errors in parentheses. * p<0.1, **p<0.05, ***p<0.01

Table 5: Second stage OLS for log(PM25) and log(GSVA).

	STN 1	STN 2	STN 3	STN 4	STN 5	
CONST	-2.5885***	-1.5914***	0.0296	0.0888	0.0621	
HDI	21.6298*** (0.1630)	(0.2848) 20.2625*** (0.3807)	(0.2663) (0.3551)	18.0664*** (0.2449) (0.3250)	18.0328*** (0.2515) (0.3359)	18.0686***
log_PM25_fit	-0.0123*** -0.0042	-0.0113** -0.0052	-0.0099** -0.0042	-0.0149*** -0.0043	-0.0159*** -0.004	
R-squared	0.9747	0.8778	0.8546	0.8908	0.8778	
R-squared Adj.	0.9746	0.8772	0.8539	0.8902	0.8772	
R-squared	0.97	0.88	0.85	0.89	0.88	
No. observations	509	402	457	399	419	

Standard errors in parentheses. * p<0.1, **p<0.05, ***p<0.01

Table 6: Second stage OLS for log(PM10) and log(GSDP).

	STN 1	STN 2	STN 3	STN 4
CONST	1.2066*** (0.2147)	0.7930*** (0.2411)	1.3176*** (0.2359)	1.3490*** (0.2094)
HDI	16.7654*** (0.2842)	17.2401*** (0.3198)	16.5653*** (0.3096)	16.5533*** (0.2760)

log_PM10_fit	-0.0200*** (0.0055)	-0.0167*** (0.0063)	-0.0167*** (0.0063)	-0.0187*** (0.0054)
R-squared	0.9029	0.8956	0.8666	0.9017
R-squared Adj.	0.9024	0.895	0.866	0.9013
R-squared	0.9	0.9	0.87	0.9
No. observations	382	344	458	405

Standard errors in parentheses. * p<0.1, ** p<0.05, ***p<0.01

Table 7: Second stage OLS for log(PM10) and log(GSVA).

	STN 1	STN 2	STN 3	STN 4
CONST	-0.0298	-0.1393	0.0976 (0.2698)	0.1334 (0.2399)
HDI	18.2786*** (0.3257)	18.8231*** (0.3668)	18.0485*** (0.3541)	18.0352*** (0.3161)
log_PM10_fit	-0.0230*** (0.0063)	-0.0832 (0.0071)	-0.0192*** (0.0072)	-0.0215*** (0.0062)
R-squared	0.8939	0.8861	0.8551	0.8926
R-squared Adj.	0.8934	0.8854	0.8545	0.892
R-squared	0.89	0.89	0.86	0.89
No. observations	382	344	458	405

Standard errors in parentheses. * p< .1, ** p<0.05, ***p<0.01

CONCLUSION

The results illustrate that crop residue burning in Punjab and Haryana are negatively correlated with the GSDP of New Delhi. Thus, controlling air pollution is urgently required, not just on moral, health or environmental preservation reasons, but for economic reasons as well. This requires a policy which incentivizes the wider adoption of existing solutions that eliminate the need to burn crop residue by giving better, or more economically viable alternatives to the farmers. Enriching the soil with biochar prepared from paddy stubble is one such alternative. Other methods can be better management of labour wages and availability in these states. The state governments can directly employ workers to cut the crop residue, and take the stubble from farmers in return for the labour. The biochar produced from this stubble can then be sold to farmers all over the country at subsidized prices. Policy designs which solve the causes of stubble burning are expected to be the most effective.

For a growing economy like India, ignoring the long-run and short-run economic impacts of environmental damage can prove to be a costly mistake a few decades down the line. Immediate policies and their implementation are required to solve this growing problem, which has been left unheeded for the last two decades.

COMPETING INTERESTS

No competing interests.

FUNDING STATEMENT

The authors received no specific funding for this work.

ACKNOWLEDGEMENTS

The authors thank Dr Alpana Agarwal, Dr Parth Parihar and Dr Nitish Mital for their valuable suggestions and inputs to the research article.

REFERENCES

1. CPCB. Central control room for air quality management-Delhi NCR. Government of India. 2017.
2. National Clean Air Programme. Technical report, ministry of forest, environment and climate change, government of India, 2019.
3. Abadie A, Cattaneo MD. Econometric methods for program evaluation. *Annu Rev Econ*. 2018;10:465-503.
4. Adhvaryu A, Kala N, Nyshadham A. Management and shocks to worker productivity. *J Political Econ*. 2022;130(1):1-47.
5. Ali N, Islam F. The effects of air pollution on COVID-19 infection and mortality. A review on recent evidence. *Front Public Health*. 2020;8:580057.

6. Dechezlepretre A, Rivers N, Stadler B. The economic cost of air pollution: Evidence from Europe. OECD Economics Department Working Papers, No. 1584, OECD Publishing, Paris. 2019;62.
7. Aragon FM, Miranda JJ, Oliva P. Particulate matter and labor supply: The role of caregiving and non-linearities. *J Environ Econ Manage.* 2017;86:295-309.
8. Balakrishnan K, Ganguli B, Ghosh S, Sankar S, Thanasekaraan V, Rayudu VN, et al. Part 1. Short-term effects of air pollution on mortality: Results from a time-series analysis in Chennai, India. *Res Rep Health Eff Inst.* 2011;1(157):7-44.
9. Bashir MF, Jiang B, Komal B, Bashir MA, Farooq TH, Iqbal N, Bashir M. Correlation between environmental pollution indicators and COVID-19 pandemic: A brief study in Californian context. *Environ Res.* 2020;187:109652.
10. Beig G, Sahu SK, Singh V, Tikle S, Sobhana SB, Gargeva P, et al. Objective evaluation of stubble emission of North India and quantifying its impact on air quality of Delhi. *Sci Total Environ.* 2020;709:136126.
11. Bherwani H, Nair M, Musugu K, Gautam S, Gupta A, Kapley A, Kumar R. Valuation of air pollution externalities: Comparative assessment of economic damage and emission reduction under COVID-19 lockdown. *Air Qual Atmos Health.* 2020;13:683-694.
12. Cusworth DH, Mickley LJ, Sulprizio MP, Liu T, Marlier ME, DeFries RS, et al. Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India. *Environ Res Lett.* 2018;13(4):044018.
13. Global Data Lab. Subnational HDI (v5.0). Institute for Management Research, Radboud University, 2013-2023.
14. Hansen-Lewis J. Does air pollution lower productivity. Evidence from Manufacturing in India. 2018;56.
15. Khare P, Deshmukh Y, Yadav V, Pandey V, Singh A, Verma K. Biochar production: A sustainable solution for crop residue burning and related environmental issues. *Environ Prog Sustain Energy.* 2021;40(2):e13529.
16. Gopi C, Khilnani, Tiwari P. Air pollution in India and related adverse respiratory health effects: Past, present, and future directions. *Curr Opin Pulm Med.* 2008;24(2):108-116.
17. Kumar A, Kushwaha KK, Singh S, Shivay YS, Meena MC, Nain L. Effect of paddy straw burning on soil microbial dynamics in sandy loam soil of Indo-Gangetic plains. *Environ Technol Innov.* 2019;16:100469.
18. Lohan SK, Jat HS, Yadav AK, Sidhu HS, Jat ML, Choudhary M, et al. Sharma PC. Burning issues of paddy residue management in north-west states of India. *Renew Sust Energ Rev.* 2018;81:693-706.
19. Maji K, Dikshit AK, Deshpande A. Assessment of city level human health impact and corresponding monetary cost burden due to air pollution in India taking Agra as a model city. *Aerosol Air Qual Res.* 2017;17(3):831-842.
20. Panteli M, Delipalla S. The economic cost of ill health due to air pollution: Evidence from Greece. *Int J Appl Bus Econ Res.* 2021.
21. Borgschulte M, Molitor D, Zou E. Smoked out: The effect of wildfire smoke on labor market outcomes. Technical report, technical report. Working Paper; 2016.
22. Ministry of Statistics and Programme Implementation. Economy of Delhi, 2021.
23. Mohr RD. Technical change, external economies, and the Porter hypothesis. *J Environ Econ Manage.* 2002;43(1):158-168.
24. NASA. Active Fire Data. Fire Information for Resource Management System, 2023.
25. Pandey A, Brauer M, Cropper ML, Balakrishnan K, Mathur P, Dey S, et al. Health and economic impact of air pollution in the states of India: The global burden of disease study 2019. *Lancet Planet Health.* 2021;5(1):e25-38.
26. Porter ME. Competitive advantage of nations: Creating and sustaining superior performance. Simon and Schuster; Free Press, 2011:896.
27. Quah E, Boon TL. The economic cost of particulate air pollution on health in Singapore. *J Asian Econ.* 2003;14(1):73-90.
28. Rajput P, Gupta T. Instrumental variable analysis in atmospheric and aerosol chemistry. *Front Environ Sci.* 2020;8:566136.
29. Sangwan V, Deswal S. In-situ management of paddy stubble through microbial biodegradation. In: *E3S Web of Conferences*, 10 March 2021, India, 241: 03001.
30. Sanjay, Swamy HM, Seidu M, Singh SB. Issues of paddy stubble burning in Haryana: Current perspective. *Paddy Water Environ.* 2021;19:55-69.
31. Sarkar S, Singh RP, Chauhan A. Increasing health threat to greater parts of India due to crop residue burning. *Lancet Planet Health.* 2018;2(8):e327-e328.
32. Siddique S, Ray MR, Lahiri T. Effects of air pollution on the respiratory health of children: A study in the capital city of India. *Air Qual Atmos Health.* 2011;4(2):95-102.
33. Solow RM. The economics of resources or the resources of economics. In *Classic papers in natural resource economics*, London: Palgrave Macmillan UK, 1974; 257-276.
34. The Lancet Planetary Health. Government indifference over air pollution crisis in Delhi. Technical Report 9, 12 2017.
35. Upadhyay A, Dey S, Chowdhury S, Goyal P. Expected health benefits from mitigation of emissions from major anthropogenic PM 2.5 sources in India: Statistics at state level. *Environ Pollut.* 2018;242:1817-1826.
36. Wang Z, Wei W, Zheng F. Effects of industrial air pollution on the technical efficiency of agricultural production: Evidence from China. *Environ Impact Assess Rev.* 2020;83:106407.