The Air Pollution-linked Productivity Impacts of a Groundwater Conservation Policy in India

Anshuman Tiwari*

10 November, 2022

Abstract

Alarming rates of groundwater aquifer depletion in North India are linked to water-intensive rice cultivation based on cheap electricity for water pumps. In this second-best setting where optimal marginal pricing of groundwater is not possible, the northwestern states of Punjab and Haryana with the highest groundwater depletion rates instead instituted laws in 2009 intended to foster reliance on rain-fed irrigation by mandating a delay in rice crop transplantation to coincide with monsoon arrival. At the same time, rice crop residue burning in these two states contributes to high particulate matter levels over North India. In this paper, I use satellite data on monthly fires and a difference-in-differences framework to document that the groundwater laws shifted more than half of all agricultural fires into early winter, when meteorological conditions favor longer suspension of particulate matter over North India. I then quantify the consequences of this increased air pollution on Indian GDP by estimating two further elasticities. First, I develop a novel instrument for PM2.5 that summarizes the exposure of a given location to all upwind fires, showing that 10% higher district exposure to November fires increases annual PM2.5 by 0.3%, and that 4% of within-district annual variation in PM2.5 can be explained by exposure to November fires. Second, I estimate the effect of higher PM2.5 levels on GDP with data on Indian districts between 2007-2013 using district and year fixed effects combined with a first differences approach that is more efficient for non-stationary data, and with the fire exposure instrument to tackle residual reverse causality. With this approach, I find estimates that a 10% increase in PM2.5 reduces GDP by 1.8%, with a 95% interval of [-0.4%, -3.17%]. With these two elasticities and the structure of the instrument, I estimate that the groundwater laws decrease yearly Indian GDP by 0.125% due to the increase in November fire-driven air pollution.

*London School of Economics (email: a.tiwari@lse.ac.uk). I am grateful to Gharad Bryan, Robin Burgess and Sefi Roth for their guidance. Financial support from the Grantham Research Foundation and LSE STICERD is gratefully acknowlegded. All errors are mine.

1 Introduction

The relationship between economic growth and the environment at various levels of development is poorly understood. The Environmental Kuznets Curve hypothesized an inverted-U shape, with economic growth worsening environmental outcomes at low levels of GDP per capita, but improving these outcomes at high levels (Grossman and Krueger 1995). Recent evidence has disputed this characterization, instead focusing on uncovering the causal factors behind various environmental outcomes at different stages of development[1] (Jayachandran 2022; Stern 2017). This paper documents that the protection of groundwater resources in India substantially increased air pollution, resulting in substantial cost to economic growth.

Aquifer depletion due to over-exploitation of groundwater is a well-documented and pressing problem in India (World Bank 2021), Aquifer depletion has documented economic costs today (Blakeslee et al. 2020) with adaptation to long-term water loss uncertain (Hagerty 2021). In order to combat particularly acute groundwater depletion, the states of Punjab and Haryana in India passed groundwater conservation laws that inadvertently increased concentrations of particulate matter less than 2.5 microns in diameter (PM2.5) locally and across inter-state boundaries.[2] India has the highest average PM2.5 concentrations in the world at 7 times the WHO standards (Greenstone 2021). The economics literature documents wide-ranging impacts of this type of air pollution, including on human health and mortality (Schlenker and Walker 2016; Deryugina et al. 2019) and labor productivity (Graff Zivin and Neidell 2012; Chang et al. 2019; Fu et al. 2021; Adhvaryu et al. 2022; Borgschulte et al. 2022) among others. This paper quantifies the short-term consequences for economic growth of increases in PM2.5 driven by the groundwater conservation laws, and summarized in the Gross Domestic Product (GDP). Thus, decisions to protect local groundwater resources in the interest of local long-term growth caused inter-state air pollution externalities with wide-ranging and immediate economic costs.

The groundwater depletion problem in Punjab and Haryana has its roots in the worldwide Green Revolution of the 1960s that allowed poor countries such as India to grow sufficient food and avoid regular famine events (Pingali 2012). The set of institutions that took root during that period in Punjab and Haryana led to the cultivation of water-intensive rice crop in these states that had not cultivated any rice before. Farmers were incentivized to pump out groundwater to irrigate paddy fields.[3] By the early 2000s, the state governments had realized the problem; but rather than

[1] For a sample of urban areas across the world, Jayachandran (2022) documents that greenhouse gas emissions keep increasing with GDP per capita, lead pollution displays the EKC inverted-U pattern, air pollution (particulate matter concentrations) displays a linear and correlation while Ozone does not display any correlation at all.

[2] Other pollutants such as Ozone and PM10 may also be correlated with increases in PM2.5.

[3] Rice is synonymous with paddy, with interchangeable references to paddy or rice fields.

incentivizing a shift away from the rice crop or instituting a marginal price for groundwater[4], these states decided to force farmers to push back the dates of rice transplantation[5] from mid-May to mid-June, hoping that the arrival of the monsoon by late June would reduce dependence on groundwater to grow rice.

Even before these groundwater conservation laws were passed in 2009, farmers in Punjab and Haryana would set fire to their fields after the rice harvest in October in order to prepare the same fields for planting the staple wheat crop. These fires clear out leftover residue after rice harvest that come in the way of planting wheat seeds, and have become popular as the cheapest method to get rid of this residue. Although agricultural fires are nominally illegal, enforcement is rare, with an average expected fine in Haryana of 0.75 USD while the average marginal cost to clear the field without fires is 50 USD (Behrer 2019; Lohan et al. 2018). By forcing a shift in the transplantation dates to early June in order to conserve groundwater, these states also shifted rice cultivation dates from October into November. Since any delays in planting of wheat crops can reduce yields (McDonald et al. 2022), farmers had further incentives to utilize fires to get rid of the rice residue after a later harvest due to the laws. These factors had the unintended consequence of shifting the peak of agricultural fires from October into November, when the onset of winter brings lower wind speeds and temperatures that slow the dispersion of particulate matter (Vallero 2014).

I utilize a two-way fixed effects design with information on the timing of the groundwater laws to document that the laws increased November fire count by 54% and the measure of biomass burnt by 72% in districts of Punjab and Haryana. Small anticipation effects imply that these may be slight underestimates. At the same time, October fire count and measure of biomass burnt decreased by 58% and 57% respectively. Since winter fire activity is predominantly concentrated in these two months, these results document the shift in monthly fire patterns toward early winter. Next, I develop a novel method to capture the effect of the increase in November fire activity on annual PM2.5 levels. November fire exposure strongly affects annual PM2.5 levels across India, with changes in fire exposure explaining 4.2% of annual deviations in within-district PM2.5, compared to 16% explained by local weather.

Next, I analyze the effect of PM2.5 levels on district GDP using newly available panel data for 530 Indian districts between 2007 and 2013. In order to account for the non-stationary nature of the GDP data[6], I utilize district-specific time trends as well as a first differences approach. The latter performs better with strongly non-stationary data series and is commonly used in macroeconomic

[4] The marginal price of groundwater extraction continues to be close to zero, since electricity for pumps is subsidized with flat tariffs rather than marginal pricing, and any outstanding farm electricity bills are rarely paid back to state distribution companies.

[5] This process of moving seedlings grown in nurseries into fields reduces weed removal and produces higher yields. More details at http://www.knowledgebank.irri.org/training/fact-sheets/crop-establishment/manual-transplanting

[6] Average yearly district GDP growth rate in this period was 7%.

analysis of GDP data (Wooldridge 2010). Identification of the causal effect of PM2.5 on district GDP relies on yearly deviations in PM2.5 being plausibly exogenous, controlling for year and district fixed effects as well as time trends. However, causality may yet run from GDP to PM2.5, with larger yearly deviations in pollution systematically being a result of higher economic activity in that district in the given year. To tackle this endogeneity concern, I instrument for PM2.5 using the novel variable linking all upwind fires to local PM2.5 concentrations. With this instrument, I show that a 1% increase in PM2.5 levels reduces GDP by 0.18% in Indian districts.

With these causal estimates in hand, I calculate the effect of the increase in November fires due to the groundwater laws on net GDP in India. Districts that are closer to and downwind of districts in Punjab and Haryana see a larger increase in particulate matter concentrations, and therefore reductions in GDP. I calculate an estimate of yearly net GDP losses using the three estimated elasticities: the increase in November fire strength due to the laws, the increase in downwind PM2.5 due to higher November fire exposure, and the reduction in GDP from an increase in PM2.5 levels. This leads to an estimated yearly loss of 0.125% of national GDP due to the groundwater laws. For comparison, the share of yearly expenditure as a percentage of GDP on the National Rural Employment Guarantee Scheme (NREGS), a flagship welfare program, is about 0.25%. The loss figure of 0.125% is also an underestimate of the overall economic costs associated with the increased pllution due to these laws, since it does not account for increased infant and old-age mortality, as well unaccounted expenditures on health, lost schooling years etc. that are not monetized into GDP.

This paper contributes to the literature on how institutions affect environmental outcomes in developing countries. More stringent regulation (Burgess et al. 2019), use of technology (Assunção et al. 2022) and higher resource allocation to monitoring and enforcement (Duflo et al. 2018) can improve environmental outcomes in developing countries. Weak state capacity impedes the implementation of regulations on the books that prohibit the use of fires in agriculture in Punjab and Haryana, leading to large economic costs.

Another explanation may be that while the groundwater externality is localized to the two states, the air pollution externality is an inter-state phenomenon. Lipscomb and Mobarak (2017) show that decentralization of regulatory authority in Brazilian municipalities leads to larger water pollution externalities across border. Kahn et al. (2015) document that providing promotion incentives to reduce some water pollutants to local officials reduces their externality on downstream neighbors. While the states of Punjab and Haryana are able to successfully implement one set of laws intended to conserve local groundwater, they are unable (or unwilling) to implement regulations on fires, which cause downwind externalities on top of local ones. This result suggests that designing regulatory institutions for environmental protection at the appropriate level is important in determining the outcomes of regulation.

Finally, this paper also relates to the literature on second-best institutions in developing countries. Rodrik (2008) argue that institutional design in developing countries with multiple distortions should not insist on the first-best, since the desired outcome can be achieved at lower cost through second-best practices. The textbook, first-best solution to the groundwater externality in Punjab and Haryana is marginal pricing of groundwater. But this is very unlikely to occur given the political power of farmer lobbies in these two states. Unfortunately, the second-best solution to utilize the monsoon rains for rice cultivation backfired by exacerbating the air pollution externality.

The rest of the paper is structured as follows. Section 2 describes the data while section 3 presents the context and motivates the QSE model by estimating the extent of pollution externalities from agricultural fires. Section 4 presents the model of general equilibrium while section 5 describes the estimation of the parameters governing equilibrium. Section 6 conducts counterfactuals and section 7 concludes.

2 Background

India is the largest user of groundwater in the world; but with almost 20% of the world's population, it only has about 4% of the world's freshwater resources (World Bank 2021). The resulting overuse of groundwater to meet population needs has caused rapid aquifer depletion and led to an urgent environmental crisis, particularly affecting the alluvial plains of North-Western India. This section first discusses the factors behind this depletion in the North-Western states of Punjab and Haryana, leading to the passage of a set of groundwater conservation laws in 2009. I then discuss how these laws unintentionally pushed agricultural fires into early winter, when their impact on air pollution is exacerbated due to meteorological conditions.

2.0.1 Groundwater conservation laws in Punjab and Haryana

Until the advent of the so-called Green Revolution of the 1960s that raised agricultural productivity dramatically across India and much of the poor world (Pingali 2012), North-Western India was a primarily wheat-growing region with little rice consumption or production locally. One of the institutional innovations of the Green Revolution in these states was the provision of large subsidies for tubewells and borewells. Individual farmers could now access shallow groundwater to irrigate fields even if they did not have access to the large, pre-existing canals systems built by the colonial British empire. Over time, modern pumps running on electricity were combined with practically zero tariffs to farmers so that they could irrigate their fields at minimum cost.

This newfound access to groundwater allowed farmers to diversify their crop portfolio by allowing the cultivation of the highly water-intensive rice crop during the "Kharif" or monsoon season (June-October). The wheat crop is cultivated during the "Rabi" or winter season, when the lower temperatures and plenty of sunshine provide perfect weather conditions for growth (Kataki et al. 2001); planting happens in early winter and harvest in early spring.

The state of Punjab contributed less than 1% of India's rice in 1961; by the late 1990s this figure was up to 10%; absolute rice output across India rose from 11 million tonnes to 75 million tonnes in this period, underlining the massive increase in rice cultivation in Punjab (Subramanian 2017). Similar trends in rice cultivation were seen in Haryana. This fundamental change in the cropping patterns of the region exacerbated the depletion of groundwater resources, since the paddy fields were flooded primarily using groundwater, pumped out before the annual monsoon reached Punjab and Haryana. Taken together, the unregulated exploitation of groundwater had led to an acute water crisis by the early 2000s, although concerns about excessive extraction almost 1.5 times the natural recharge rate had been expressed by agricultural scientists and government committees going back to the 1980s (Singh 2009).

Despite the alarm expressed by various stakeholders, the state governments largely ignored the problem until the early 2000s. When asked about these concerns, the then-Chief minister of Punjab, Prakash Singh Badal, is quoted in the media as saying, *"The problem is not as acute as is being projected. It is a theoretical evaluation and there is no truth in it"* (Down To Earth 1999). The political economy of both states, but particularly of Punjab, centers around medium and large sized farmers who receive a range of state subsidies that incentivizes rice and wheat cultivation. Apart from the Green Revolution era technological subsidies for higher-yielding seeds, fertilizers and pesticides, tubewells and electric pumps, provision of cheap electricity is also important in explaining groundwater levels (Ryan and Sudarshan 2020). Often these dues are not paid to state distribution companies at all, resulting in lack of investment in the power grid (World Bank 2018). But, most importantly, the procurement of wheat and rice crop by the state governments of Punjab and Haryana at so-called Minimum Support Prices that distort market signals (Parikh et al. 2003) precludes farmers from switching to other crops with higher price and yield risks.

The practice of transplantation of rice before mid-June was thought to be particularly cause too much reliance on groundwater (Singh 2009). In response, sections of the state bureaucracy had made efforts starting in the early 2000s to shift the transplantation of rice closer to the monsoon, since this was thought to ease the strain on groundwater use. The two governments took executive action through ordinances in 2008[7] to extend the practice of delaying rice transplantation state-wide. Given the generally favorable response to this ordinance, the legislatures of Punjab and

[7] These do not have the same power in Indian law as a statute and cannot be renewed beyond a few months.

Haryana separately ratified the Preservation of Subsoil Water Acts of 2009 (“laws” from now on) in an effort to conserve groundwater.

These laws prohibited early transplanting of rice before the monsoon in an attempt to reduce groundwater usage for irrigation. Much of the rice transplantation would occur in the peak of summer during May when evapotranspiration (water loss from plants as well as soils and water bodies) is high. These laws specified June 10 as the earliest transplantation date, and it was shifted further to June 20 later[8]. When planting rice in May, farmers were solely dependent on groundwater reserves for rice growth; moving transplantation to June allowed rice growth to depend more on monsoon rainfall. This was expected to lead to a lower rate of groundwater extraction.[9]

2.0.2 Increase in agricultural fires due to shifting of rice crop harvest

The primary use of fires in Indian agriculture today is to clear the field of leftover residue from harvesting a crop, before sowing and planting the next crop (Shyamsundar et al. 2019); this differs from slash-and-burn agriculture that is practiced in parts of Africa and Indonesia (Andini et al. 2018). Figure 1 shows that agricultural fires are concentrated in the states of Punjab and Haryana, which are also characterized by a Rice-Wheat crop system.

In the Rice-Wheat system of Punjab and Haryana, fires are used to clear rice residue before the planting of the wheat crop on the same land, since rice residue comes in the way of planting wheat. This practice dates back at least to the 1990s. The earliest observations of fires from the NASA FIRMS database (described in the next section) starting in 2002 clearly demonstrate that North-western India already had a disproportionate share of fires in Indian agriculture.

The delay in rice transplantation due to the laws also pushed back harvest dates. The resulting delay in rice harvest from mid October to late October and November meant that farmers had fewer days between rice harvest and wheat plantation. Any delays in wheat plantation beyond the first two weeks of winter reduces yields substantially (McDonald et al. 2022). Therefore, the law had the unintended consequence of increasing the intensity of agricultural fires in November, when slower winds and lower temperatures tend to worsen downwind air quality.

[8] The Indian Met Department (IMD) sets out July 1 as the expected date of onset of the monsoon in North-Western India. Details here

[9] Groundwater recharge is typically a slow-moving process that takes place over a longer period than the period of study here. I plan to conduct an assessment of the change in groundwater levels to the present day due the policy in the future

3 Data and Measurement

3.0.1 Agricultural fires

The burning of residue from crop harvest is referred to as agricultural fires. To analyse the impact of the groundwater conservation laws on the monthly pattern of fires, the ideal data would include precise location of each individual fire set for the purpose of burning crop residue. But there are no representative ground-level observations of this phenomenon. To overcome this challenge, I utilize the Fire Information for Resource Management System (FIRMS) product from the National Aeronautics and Space Administration (NASA) agency of the United States that is widely used to identify terrestrial fires. This product provides information on daily fires detected at latitude/longitude level across the world and has been recently used to analyze agricultural fires in the economics literature (Behrer 2019).

FIRMS provides a few related products: Near-Real Time (NRT) fires using the MODIS instrument aboard Terra and Aqua satellites, standard product from the same instrument but with a 2-3 month lag and another NRT product using the VIIRS instrument from the Suomi-NPP and NOAA-20 satellites. The main difference between the first two and the third is the resolution of the data. MODIS products are at 1 km resolution and are available from 2000 (more reliable from 2002 when Aqua satellite was launched) whereas VIIRS products are at 375 m but only available from 2012. The primary analysis utilizes the MODIS standard product which differs from the NRT data in that corrections are made to the imprecise location of the Aqua satellite in the NRT data. Imagery data from Aqua and Terra satellites is available at least four times daily for each pixel on Earth and is processed using a NASA algorithm isolate a ground-level fire signal from other signals such as solar flares.

I combine this data with information on land use from the European Space Agency Climate Change Initiative's land cover map (version 2.07).[10] This allows the subset of fires that is found on agricultural land to be separated from natural forest fires since this paper is interested in agricultural fires. I aggregate and resample the land cover data which is at a resolution of 300 m to the fire data grid (at 1 km resolution), with an indicator for agricultural land use as the main output from this process. All fires are then masked based on this indicator variable to find the subset of agricultural fires.

[10] Data is available at https://cds.climate.copernicus.eu/cdsapp/#!/dataset/satellite-land-cover

3.0.2 Air quality

An important consideration for air quality data is complete geographical coverage. Whereas ground-level monitoring station coverage in India is extremely sparse (Greenstone and Hanna 2014), satellite imagery-based products provide complete coverage. Secondly, ground-level observations may be susceptible to manipulation (Greenstone et al. 2022; Ghanem and Zhang 2014). Therefore, the main source of data on air quality is Hammer et al. (2020), a gridded reanalysis product of global surface PM2.5 concentrations at a resolution of 0.01° that should be much less susceptible to such manipulation. This product combines satellite imagery data on Aerosol Optical Depth with state-of-the-art chemical transport models, and calibrates the output to global ground-based observations. It is easy to aggregate the gridded product to the necessary resolution for analysis at pixel, city or district level. Forthcoming sections will detail the aggregation procedure for each analysis.

3.0.3 GDP data

To estimate the impacts of PM2.5 on GDP, we would like data at the most granular sub-national level possible. While satellite-based data on PM2.5 levels are available at a 1x1 km grid, output data are rarely available at such sub-national scales. Fortunately, GDP measures at the district level in India between 2007-2013 have recently been compiled by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in their District Level Database (DLD).[11] I clean and combine these data with other district-level data using district identifiers from ICRISAT and the Census of India, 2011.

3.0.4 Meteorological data

Hourly wind data are used to construct exposure to agricultural fires for every origin-destination pixel pair. Details of the methodology follow in the section 3 below. The source of these wind data is the European Center for Medium Range Weather Forecasting (ECMWF) ERA5 family of global gridded reanalysis datasets.[12]. Reanalysis data combine ground-level observations and satellite data with Chemical Transport Models that represent physical and chemical processes in the atmosphere to produce reliable and complete coverage for the world. Since ground-level observations are particularly sparse in developing countries these reanalysis data are widely used in the literature on climate and air pollution in Economics (Auffhammer et al. 2013) Hourly wind speed and direction data are taken from the ERA5-Land hourly dataset which is available at a resolution of 0.1°.

[11] http://data.icrisat.org/dld/src/crops.html
[12] Data available at https://cds.climate.copernicus.eu

These are combined with daily agricultural fires at the pixel level to construct the fire exposure variable. Apart from being used to quantify the contribution of distant residue burning on local air pollution, fire exposure also is an instrument for pollution at the city and district level in estimation of certain elasticities. Finally, I also construct temporal averages for weather variables including rainfall, temperature and relative humidity from this dataset to be used as controls in the regression analysis.

3.0.5 Constructing instrument for air quality using agricultural fires

Since the laws push fires in Punjab and Haryana into November when there used to be very few fires earlier, I only consider the month of November in constructing this instrument. I capture the contribution of daily agricultural fires F_{ot} from 1x1 degree origin pixel o on air quality in destination pixel d, at a distance of $dist_{od}$[13], in the month of November, as follows

$$\omega_{od} = \left(\sum_{t \in Nov} \frac{windfrac_{odt} * F_{ot}}{dist_{od}} \right) \tag{1}$$

$windfrac_{odt}$ is the daily average fraction of time that the wind at o blows towards d on day t. In order to calculate $windfrac_{odt}$, I start by assigning each hourly wind observation in o on day t into one of 36 bins of 10 degree span each, based on the wind direction that hour (true north is 0 degree as in the figure). I then construct the wind speed-weighted fraction of time the wind was blowing in each of these 36 bins by aggregating hourly observations for day t. $windfrac_{odt}$ is then calculated by summing up wind fraction for the bins which are in the direction of d from o as shown in figure 3.

I construct this instrument for various distances between origin and destination districts.[14] These are increased sequentially so that the distance that maximizes power in predicting PM2.5 can be selected (Details in the next section of this estimation). Yearly variation in the instrument is driven by two factors: (i) changes in the temporal distribution of fires at origin and (ii) changes in the daily wind patterns at origin during November.

Table 1 summarizes the main variables used in the analysis.

[13] The distance elasticity is assumed to be -1 but will be estimated using NLS

[14] Distances are calculated using district centroids.

4 Research Design

4.0.1 Effect of policy on local fires

I utilize a difference-in-differences research design with fixed effects to test how the groundwater conservation laws shifted the monthly pattern of fires. The outcome variables in each district-month-year period from 2002 to 2020 are the count of fires and the total strength of these fires as measured by the fire radiative power. These are aggregated to the district-level to reflect the administrative unit at which state policy is implemented in India. I estimate a Poisson fixed effects model to recover the coefficient of interest, assuming the standard exponential link function (Behrer 2019; Ranson 2014) for the count or measure of biomass burnt F_{dmy} in district *d*, month *m* and year *y*. The conditional expectation function given regressors $\mathbf{X}_{dmy}$ is as follows

$$\mathbf{E}[F_{dmy}|\mathbf{X}_{dmy}] = exp(\sum_{m\in[1,12]} \delta_m \, D_{dmy} + \alpha_d + \tau_{sm} * Y_y) \tag{2}$$

where the RHS inside the exponential function contains $\mathbf{X}_{dmy}$. Since the laws came into force at the state-level in 2009, the treatment indicator D_{dmy} turns on for district-months in Punjab and Haryana in and after 2009. District fixed effects D_d control for unobserved determinants of fires that do not change over time. Comparison of fires within state-by-month cells (τ_{sy}) flexibly controls for other within-state determinants of fire seasonality such as different crop calendars, crop mixes etc. that do not change over time. Year fixed effect Y_y controls for any common trends across the country (such as the country-wide increase in fires driven by the Mahatma Gandhi National Employment Guarantee Scheme or NREGS documented by Behrer (2019)).

The count nature of the data and the nontrivial presence of zeros in the count data motivate the use of a Poisson model. A log transform of F_{dmy} would create bias in a linear model estimation whereas an inverse hyperbolic sine transform makes the interpretation of the elasticity slightly more complicated (Bellemare and Wichman 2020). Further, the Poisson FE model only requires that the conditional expectation function be specified correctly for consistent estimation of the parameters (Wooldridge 2010). It produces unbiased estimates of the coefficients even if the fire data do not match the Poisson distributional assumptions (Wooldridge 1999a, 1999b; Lin and Wooldridge 2019). This is not true of other models that are used to handle count data such as negative binomial (Blackburn 2015). I estimate this model using quasi-maximum likelihood method through the *fixest* package in R (Berge et al. 2022).[15]

[15] The Poisson model can be used with non-integer data such as the measure of biomass burnt as well, and the strengths of the Poisson over other model when the data have nontrivial presence of zeros also holds (Wooldridge 2010)

Taking log of (2) yields the following

$$log(\mathbf{E}[F_{dmy}|\mathbf{X}_d my]) = \sum_{m \in [1,12]} \delta_m \, D_{dmy} + \alpha_d + \tau_{my} \tag{3}$$

Therefore the coefficients of interest δ_m give the monthly elasticity of fire count to the policy. As with any difference-in-differences design, the main identifying assumption for the δ_ms is that trends in monthly fires would be similar between treatment and control districts in the absence of the policy change. I discuss this assumption in more details in the results section. Standard errors are clustered two ways at the district and state-by-year level to account for the district-level autocorrelation as well as implementation at the state-by-year level.

4.0.2 Effect of PM2.5 on GDP

This section describes the estimation of the causal impact of higher PM2.5 levels on district GDP in India. I build up to an instrumental variables strategy for PM2.5 that allows the quantification of impact of the groundwater laws on downwind GDP. Before describing this IV strategy, equation (4) presents an OLS regression model of the effect of PM2.5 on district GDP.

$$log(GDP_{dy}) = \beta \, log(PM_{dy}) + \gamma \, Weather_{dy} + g_d * t + \alpha_d + Y_y + \epsilon_{dy} \tag{4}$$

$$\begin{aligned} Weather_{dy} = \{&Temp_{dy}, \; Temp\_squared_{dy}, \; Rain_{dy}, \; Rain\_squared_{dy}, \\ &Relative\_Humidity_{dy}, \; Surface\_Pressure_{dy}, \; Windspeed_{dy}\} \end{aligned}$$

The quantity of interest β is the percentage reduction in GDP for a 1% increase in PM2.5 levels. This model contains district and year fixed effects D_d and Y_y respectively, which control for fixed factors that raise productivity or increase pollution as well as account for any common macroeconomic shocks. Weather variables such as temperature, rainfall, humidity and wind speed are known to affect PM2.5 (Dechezleprêtre et al. 2019; Bondy et al. 2020). Therefore, I control for yearly average weather that could determine the level of pollution from given emissions.

The main residual concern with identification of β in this model is that deviations of GDP and PM can be jointly determined. Higher economic activity in a given year can itself cause an increase in PM that year by increasing emissions. At the same time, higher deviation in PM can stunt GDP growth that year through channels such as increased worker morbidity and lower labor productivity

I adopt three approaches to tackle this endogeneity issue. Figure 4 shows that GDP exhibits strong growth in this period and therefore is not stationary; between 2007-2013, average Indian GDP growth rate was 7%. First, I fit a district-specific linear time trend $g_d * t$ in GDP. The time trend will capture district-specific factors that cause constant GDP growth, leaving only deviations from the trend line in the outcome. This approach can also help reduce omitted variables bias (OVB) that jointly determines both GDP and PM2.5 (eg. demand shocks that affect certain districts). Such OVB can cause the causal chain to run from GDP to PM2.5, leading to reverse causality that biases the estimate upwards, since an increase in economic activity increases PM2.5 levels.

Secondly, I also conduct analysis using first differences (FD) that is the preferred over fixed effects to deal with non-stationary, autocorrelated data series in both outcome and explanatory variables. Further, an FD specification that also includes a fixed effect allows for a district-specific linear growth rate g_d in the outcome. The FD approach is commonly used in the macroeconomic literature to deal with serial correlation in aggregated GDP data. Equation (5) specifies the regression framework for the FD model.

$$\Delta log(GDP_{dy}) = \beta\, \Delta log(PM_{dy}) + \gamma\, \Delta Weather_{dy} + g_d + \Delta Y_y + \Delta\epsilon_{dy} \tag{5}$$

This specification examines how the growth rate of PM2.5 affects the growth rate of GDP, controlling for year-on-year changes in weather and common macroeconomic conditions. The district fixed effect g_d captures the constant growth rate of GDP in these districts. But even with the FD design, there may still be some OVB in the leftover variation, leading to reverse causality that biases the results upwards.

The third approach utilizes the fact that the stock of pollution in a district is partially due to sources outside the district, notably agricultural fires in this instance. Using the instrument for PM2.5 defined in the previous section with both the panel and FD specifications allows us to tackle the reverse causality challenge. Fires in the winter are much worse for downwind PM due to meteorological conditions that favor longer suspension and entrapment of particulate matter in the lower atmosphere of downwind districts. I also construct the same instrument with other months separately and together for the whole year, and test the hypothesis that fires in the winter are worse for PM2.5 in fire-exposed downwind districts. Before discussing the results on the effect of PM2.5 on district GDP in section 5.0.3, I discuss the results of this first stage in section 5.0.2, based on specifications in equations 4 and 5.

Inference would ideally be conducted using Conley spatial standard errors with arbitrary autocorrelation at district level, given the spatial autocorrelation in PM2.5 and GDP levels. However, I am unable to implement these standard errors for a panel data model with instrumental variables.[16]

[16] The R package **lfe** provides the command **felm** that implements Conley SEs with panel data; however it does not

Instead I cluster standard errors at the region-year level, where regions are groups of contiguous districts that share similarities in economic and geographic fundamentals such as level of development, soil types, weather etc.[17] In this case, the dependence of PM2.5 is fundamentally spatial, and not administrative. In order to test whether clustering at the level of the region is appropriate, I plan to conduct inference using a wild cluster bootstrap later. I also plan to write code to calculate Conley standard errors in a panel IV setting.

5 Results

5.0.1 Effect of groundwater laws on monthly fire patterns

To begin the results section, I refer to figure 2 that shows some growth in the fire count and fire strength for November occurring just before the laws were passed in 2009, with a stronger trend upwards after the passage of the law, before stabilizing by 2015 or so. This suggests some anticipation effects in November, since the count of fires is trending up 2-3 years before the policy came into effect. These anticipation effects can be attributed to the informal implementation of the policy before 2009 that is discussed in section 2.0.1. This may have driven the shifts in fire patterns by slightly delaying the cultivation dates before 2009, with formal implementation inducing a larger shift. The lack of pre-trends on October fires combined with a downward shift after 2009 supports this view. Therefore, the effect for November my be an underestimate.[18]

Table 2 presents estimates of the causal effect of these laws on monthly fires, based on 2. Columns 1 and 3 provide the mean number of fires and measure of biomass burnt in Punjab and Haryana, before the passage of these laws. Those columns show peaks of fires during the months of April, October and November. Fires in the latter two months are used to clear the monsoon season rice residue, as described earlier. Fires in April are used to clear the wheat crop residue after the harvest is done. The time pressure of needing to be rid of the rice residue before wheat plantation that leads to the fires after monsoon rice harvest does not arise after the wheat harvest. Yet we see substantial fire activity in April. This wheat residue burning practice may have come about due to habit formation from setting fires to the rice residue. However, it is less troublesome for downwind

produce these SEs with panel IV estimation.

[17] There are 530 districts and 96 regions in the sample, so that there are 5.5 districts on average in each region. Each district had an area of approximately 100 sq km, on average.

[18] Given the recent literature on the bias of TWFE, I plan to test for conditional parallel trends with anticipation effects as well as the treatment effect of interest using the framework of Callaway and Sant'Anna (2021). Their approach would work well in this setting since they rely on never-treated units to estimate treatment effects. Therefore, I plan to utilize their R *did* package to estimate these effects in the future, better accounting for anticipation effects.

pollution than are fires during early winter, since meteorological conditions in April do not favor suspension of particulate matter over the plains of North India.

Turning to the results in columns 1 and 3 of table 2 now, the main result is that the laws increase the log of expected fire count in November by 0.43, and log of expected fire strength by 0.54. Estimates for October are negative, providing evidence that the laws probably succeeded in pushing rice cultivation by a few weeks to a month, and therefore peak fire activity into November. Estimates for the other months except June and July are negative. For the months from December to May, this would suggest a domino effect of the later rice cultivation on other crop burning, since the entire crop calendar gets pushed back. The spring season fire peak (from the wheat harvest) that used to happen in April and May seems to shift slightly toward June and July, generating the positive estimates for those two months. The negative estimates for August and September probably also come from the enforced delay in rice plantation that would have affected some farmers who would plant rice in early May otherwise. Finally, since there are very few fires to begin with in July, and since July happens to be the rainy season, the shifting of the wheat fire season perhaps does not have the same consequences for downwind pollution that the shifting of the rice season does.

I present robustness results to alternative specifications and sample selection in table A.5. These include the following: OLS estimation rather than Poisson, including fires data from 2000 and 2001,[19] and limiting the analysis to the sample for which GDP data is available.[20] The results are consistent with table 2 in all these robustness checks, with only the fire strength when limiting to the GDP sample becoming insignificant. This lack of power could be due to the effects of policy not having had enough time to accumulate by 2013 or to anticipation effects just before 2009. It should certainly not be taken as an indication that the laws did not increase November fire activity.

5.0.2 Effect of November fires on annual downwind pollution

Before turning to the causal effect of PM2.5 on district GDP, I discuss the effect of fire exposure on district PM2.5 levels. These results are equivalent to the first stage for the 2SLS results on GDP in the next section. As noted in the previous sections, fires in the winter are particularly harmful for PM2.5 levels due to prevailing meteorological conditions over North India that favor slower dispersion of the particulate matter over space. Further, the groundwater laws pushed agricultural fires in Punjab and Haryana toward November (early Winter). Therefore, I focus on the effect of November fire exposure on PM2.5.

[19] The NASA Aqua satellite was launched in 2002 and drastically improved estimates of fire activity in the FIRMS database

[20] 530 district between 2007-2013

The construction of the explanatory variable *log(Nov FRP exposure)* is described in section 3.0.5. Referring to that section, variable F_{ot} is the total fire strength measured by Fire Radiative Power (FRP) of all fires in district o on day t in November. Certain fires can be stronger because more organic material is burnt, thereby producing higher amounts of particulate matter. Therefore I use FRP to maximize signal in the instrument relative to using count of fires.

Next, I implement various distance cut-offs on the exposure measure: origin districts at a larger distance than the cut-off are not used to construct FRP exposure for destination district. This is done for two reasons. Firstly, while wind fraction *times* inverse-distance weighting[21] captures some of the pollution decay over distance, it could miss out on some important features that govern decay, such as (i) rainfall along the path, which can cause the "wet deposition" of particulate matter (Vallero 2014) (ii) meteorological conditions along the path such as wind speed, temperature and relative humidity that could also alter the trajectory or cause further deposition out of the atmosphere and (iii) geographical features such as mountains along the way. For this reason, I hypothesize that larger cut-offs could add more noise to the instrument. Therefore, I test which distance cut-off maximizes the within-R2, in order to quantify the trade-off between signal and noise when increasing the distance cut-offs.

Table 3 shows results for cut-offs between 500 and 1000 km. In panel A, I present results from a fixed effects model that includes a district-specific time trend, equivalent to the first stage for equation 4. Panel B presents results from the first difference model with district fixed effects in equation 5, therefore assuming a district-specific trend in growth of PM2.5. Both these sets of results show strong and robust elasticities of PM2.5 to November FRP exposure, peaking at a cut-off of 900 km (for both the coefficient size and within-R2). The main result here is that a 1% increase in November FRP exposure increases PM2.5 levels by 0.029% (0.032%) with the FE (FD) model. It further illustrates the trade-off between signal and noise when increasing distance to origin in constructing the instrument.[22] Globally, 900 km maximizes within-R2 when explaining PM2.5 using November FRP exposure. I therefore use that as the preferred distance to construct the instrument for PM2.5 in the next section.

Now, I address a concern that distance may be correlated with geographic determinants of PM2.5 (and GDP later). Controlling for district fixed effects in these regressions helps address that concern. But, distance also enters the instrument itself non-linearly; it may be that the district fixed effect does not fully address the issue. Therefore, I construct the instrument for each of these cut-offs by adding up wind fraction-weighted FRP from qualifying origins *without* inverse-distance weighting. Thus distance directly does not enter this instrument. Results for these regressions are

[21] The distance decay could be modeled through a distance elasticity different from -1 too. I plan to do this later.

[22] Results for regressions with a 100 km cut-off to no distance cut-off at all show an increasing within R2 until 900 km when they start dropping of monotonically.

presented in appendix table A.6. They do not suggest any cause for concern that distance entering the instrument non-linearly causes any bias in the first stage.

Lastly, in appendix table A.7, I confirm that higher FRP exposure only from fires during winter months affects annual PM2.5 levels. This can be explained by unfavorable meteorological conditions during winter that cause the particulate matter emissions from agricultural fires to stay suspended for longer. However, fires in the winter months other than November are not affected by the groundwater laws in Punjab and Haryana. Therefore, in order to quantify the effect of increased November fires due to the laws later, I use only November-based FRP exposure instrument in the analysis of the effect of PM2.5 on GDP in the next section.

5.0.3 Effect of PM2.5 on GDP

In this section, I turn to the impact of annual PM2.5 levels on annual GDP in Indian districts in panel A of table 4. I present results with the fixed effects in columns 1-3, and with the first difference specification in columns 4-5. Column 1 presents the OLS estimate controlling for weather and including district and year fixed effects, but without district-specific linear time trends. The coefficient is positive and strongly significant. The causal effect of higher PM2.5 on GDP should be negative, given the harmful effects on human health and productivity, and potential effects on agriculture and machinery. The positive coefficient suggests that this specification is not sufficient to address the concern about omitted variables that jointly determine GDP and PM2.5, such as yearly demand shocks that cause higher GDP growth due to certain districts being more trade-exposed, for example. Higher economic activity in that year would increase PM2.5 levels in that district, and district fixed effects are insufficient to capture the co-movement of these variables. The estimate is biased upwards since the causal chain runs from GDP to PM2.5 in such cases. The sample period witnessed very strong GDP growth in Indian districts, making this a particular concern in this setting.

Column 2 presents results with the addition of a district-specific linear time trend to reduce this concern. The coefficient turns negative now, although it is imprecise, suggesting that this time trend is able to reduce the upward bias from the reverse causality of GDP to pollution. It also suggests the importance of including such time trends for non-stationary GDP data when focusing on the effect of jointly determined variables such as air pollution, as opposed to plausibly exogenous variables such as temperature deviations (Dell et al. 2012).

Before discussing the 2SLS estimates in columns 3 and 5, I focus on column 4 which presents the first difference estimate along with a district fixed effect, in effect assuming a district-specific trend in the growth rate of GDP. The coefficient is -0.03 and significant at the 5% level. The

FD specification works much better with non-stationary data, and therefore this coefficient is less biased and also more precisely measured than the fixed effects regression with time trends in column 2.

Both these approaches solve some of the omitted variable problem plaguing estimation of the effect of PM2.5 on district GDP. However, any joint residual variation from the trend still causes upward bias in the estimates. I turn to the instrumental variable strategy to address this residual concern. In column 3, I present 2SLS results from the fixed effects model with district-specific time trends, instrumenting for PM2.5 using November FRP exposure with a 900 km distance cut-off. The estimate is now much larger, although the IV also increases standard errors as expected.

Panel B reproduces relevant first stage estimates from table 3. To test for weak instruments, I also present two statistics below the first stage estimates. Stock and Yogo (2005) suggest the use of the Cragg-Donald F-stat in a multivariate setting to test for weak instruments, with a rule of thumb that a value less than 10 indicates a potentially weak instrument. The Cragg-Donald F-stat is about 101.4; but this relies on iid assumptions for the errors. Therefore, I also report the Kleibergen-Paap (KP) F-stat which is equivalent to the robust F-stat with one endogenous regressor, as in this setting. The F-stat of 25.3 is comfortably above 10, and therefore concerns about weak instruments do not arise here.[23]

Column 5 presents the 2SLS results from the first difference model. The point estimate is slightly larger than column 3, and is estimated much more precisely. The KP F-stat is 26.4, again comfortably larger than 10. I consider the specification in column 5 as the preferred specification. These estimates suggest that increasing PM2.5 levels by 1% in a given year has a large negative causal effect of 0.18% on district GDP.

5.0.4 Quantifying the impact of groundwater laws on net GDP

The 2SLS estimates from the previous section can be used to estimate the effect of the increase in November fires in Punjab and Haryana on net GDP. Variation in the instrument in the previous section comes from November fires in all districts of India, both before and after passage of the groundwater laws. Therefore, we cannot interpret those estimates directly as the LATE associated with the laws. But we can use the estimates from this paper to calculate the percentage loss in net GDP across Indian districts in the following way.

[23] Andrews et al. (2019) recommend the use of the effective F-statistic (MOP F-stat) of Olea and Pflueger (2013) in the case of a single endogenous regressor. This statistic is not easily calculated in any R or Stata package that implements IV with panel data. However, Andrews et al. (2019) also note that with one single endogenous regressor, the MOP F-stat is equivalent to the KP and robust F-stats. Therefore, the provided F-stat is the correct on to test for weak instruments. In future versions of the paper, also plan to present identification-robust Anderson-Rubin confidence intervals which are efficient regardless of the strength of the instrument.

Results from table 2 demonstrating that the laws increased November FRP in districts of Punjab and Haryana by an average of 0.542 log points. This increase in November FRP would increase FRP exposure within 900 km of each district. Since each Punjab and Haryana district sees the same proportionate increase in fires, and the inverse-distance and wind fraction weights do not change,[24] the proportionate increase in FRP exposure for all districts within 900 km is the same. Using the distance and wind fraction weights, this proportionate increase for each district within 900 km is also 0.542 log points. The increase in PM2.5 from this 0.542 log points higher November FRP exposure is 0.542*0.032 = 0.0173 log points, using the first stage estimate from column 5 of table 4. Finally, the proportionate reduction in GDP for each district is 0.0173*(-0.179) = -0.0031 log points or -0.3%.

The same proportionate reduction in GDP for districts within 900 km can produce different reduction in net GDP based on the initial GDP. This estimate for the average yearly impact of the groundwater laws on net GDP is -0.125%, 54.29 billion Indian Rupees or 1.12 billion USD (2004 values). This estimate is based on the 530 sample districts only, assuming that November fire exposure is limited to 900 km.

6 Conclusion

This paper estimates the unintended consequences of groundwater conservation laws in the two states of Punjab and Haryana on net Indian GDP, due to increased air pollution in downwind districts. In order to arrive at the net impact, I estimate three elasticities. First, I provide evidence that the groundwater conservation laws shifted agricultural fire activity from late monsoon into early winter: biomass burnt in November increased by 72% while it decreased in October by 57%. This increased fire activity during winter more strongly affects PM2.5 levels because lower wind speeds and temperatures along with scant rainfall favor longer suspension of particulate matter in the smoke plumes.

Second, to quantify the impact of higher November fires on annual downwind PM2.5 levels, I construct a novel measure to summarize the exposure of each district to all upwind fires in November. I show that this exposure measure predicts 4% of the year-to-year variation in PM2.5 within each district. Third, I estimate the impact of higher PM2.5 levels on contemporaneous GDP. To solve concerns about omitted variable bias/reverse causality and non-stationarity, I adopt an identification strategy that relies on first differences to control for district-specific trends in PM and GDP, with an instrumental variable that provides exogenous variation in particulate matter. A 10% increase in PM2.5 levels in a given year reduces district GDP in that year by 1.8%. With these three

[24] I construct a 10-year average wind fraction for each day here

elasticities, I calculate the yearly impact of increased PM2.5 levels due to the groundwater laws on net Indian GDP to be 0.125%. This estimate does not include non-monetized impacts of this pollution on health and well-being.

I have conducted some robustness checks that are presented in this chapter, and plan to conduct more. There are also some limitations to this approach. First, the estimate relies on the exposure instrument affecting downwind districts in accordance with its structure. While a chemical transport model could do better in modeling this relationship, it is much more resource-intensive to operate and may not do especially well for seasonal sources such as agricultural fires. Second, it is not a direct LATE of the legislation itself. It relies on the GDP elasticity of pollution that is estimated using November fires both within and outside Punjab and Haryana. In future work, I plan on directly estimating the impact of increased November fires from the groundwater laws on downwind GDP by restricting fires sources to Punjab and Haryana, and leaving out districts outside North India that the exposure instrument does not affect. While this would reduce power and potentially limit external validity to the rest of India, it will also allow me to estimate more directly the impact of groundwater laws on downwind GDP through exposure only to fires in Punjab and Haryana. I also plan to

On a different note, I also plan to explore the mechanisms behind the impact of the groundwater laws on net GDP, driven by the increase in November fires and PM2.5. Does this decrease come from a reduction in industrial production or agricultural output? Is the main channel the health and labor productivity impacts of PM2.5? Can firms adjust to this increased pollution by either moving or reallocating production to other months? One potential issue could be that legislation may affect local GDP in Punjab and Haryana through the costly adaptation to the laws themselves, biasing the estimates. Removing districts in these states from the sample would solve that problem. However, fire exposure is also likely to have the largest impact on PM2.5 on districts within these states and I prefer not to remove those districts from the sample for that reason. Instead, I intend to utilize outcomes such as the Index of Industrial Production that are not likely to be directly affected by the groundwater laws.

These laws were intended to conserve critical groundwater aquifers that have deplted at an alarming rate. In this paper, I do not investigate whether the laws increased groundwater levels, or quantify other benefits of this policy. I plan to do these in the future too. But the unintended consequences of this policy were to exacerbate the effects of agricultural fires on air pollution. Even though the use of fires to clear fields of residue has large costs in India, a combination of factors ranging from weak regulatory capacity or political capture by farm lobbies at the macro level, to credit constraints or lack of trust among smaller farmers may be responsible for this continued practice. By quantifying GDP losses in other states due to the increased pollution from fires in Punjab and

Haryana, this paper suggests a mechanism whereby fiscal transfers from these downwind states affected by increased pollution could be made to farmers in Punjab and Haryana as payment not to burn (Jack et al. 2022).

The design of payments, specifically whether they should be upfront to alleviate credit constraints and combined with more stringent monitoring and enforcement, is another question for further research. While these payments go against the "polluter pays" principle that may be more relevant for the farmers in Punjab and Haryana, who are richer and have larger landholdings than the rest of India, any workable solutions in such second-best environments should consider the existing political and regulatory distortions which make these payments a sensible way to increase welfare. In the long term, incentivizing farmers to plant crops that are more suitable to the available resources, priced appropriately, could be a more sustainable solution.

7 Figures and Tables

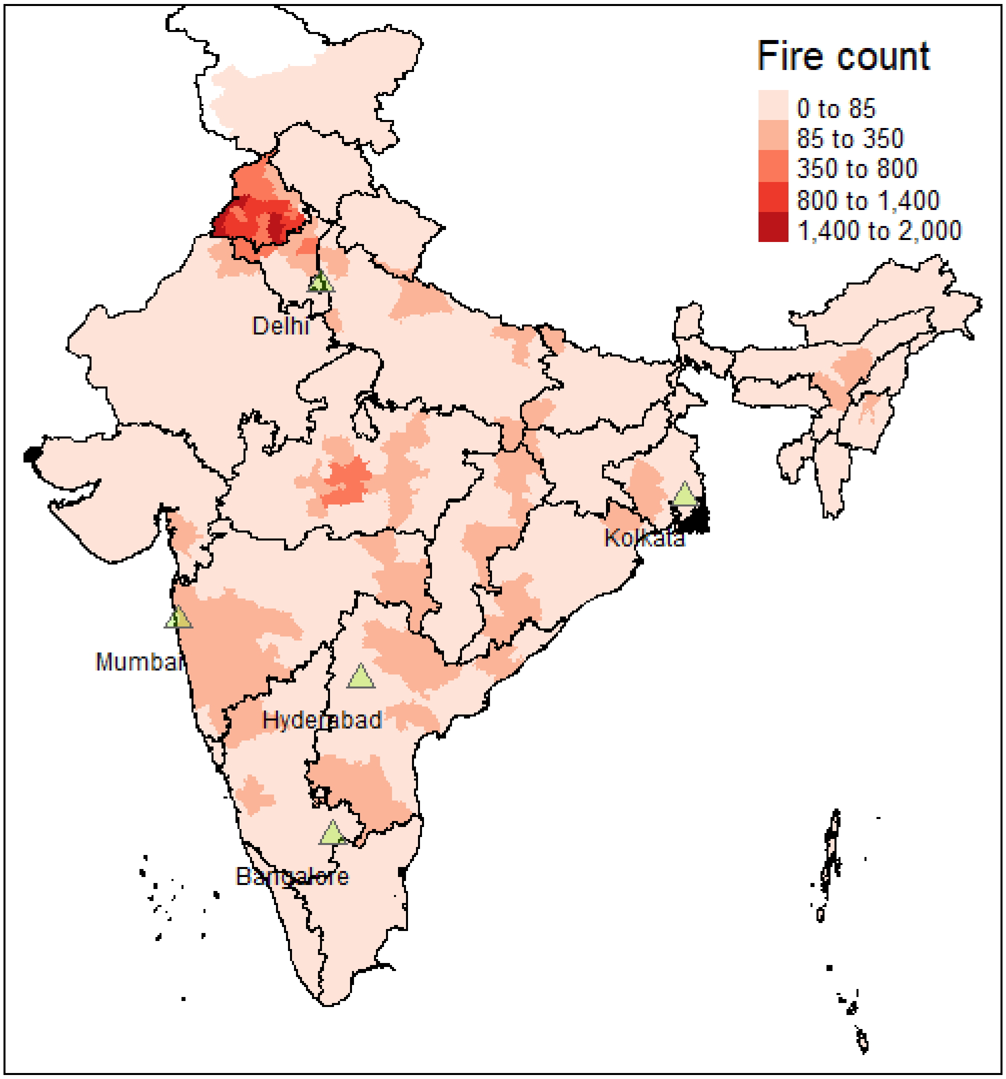
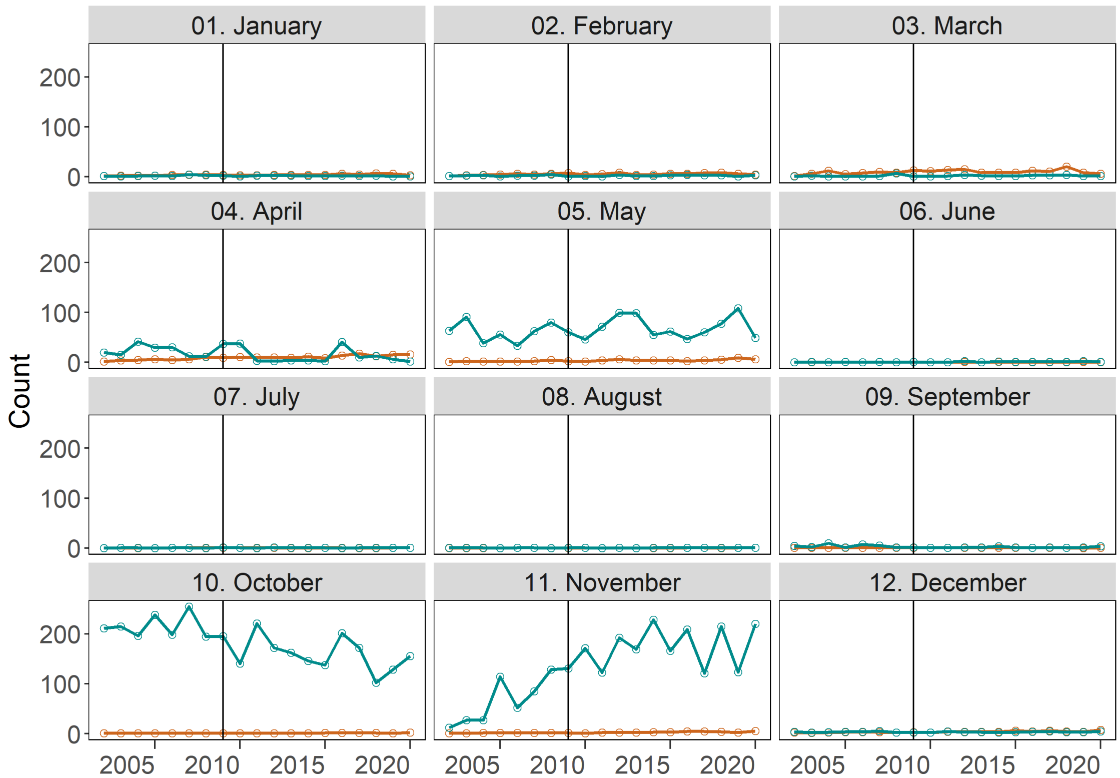


Figure 1: Count of fires in Indian districts (2010) ↩

(a) **Fire count**



(b) **Fire radiative power**

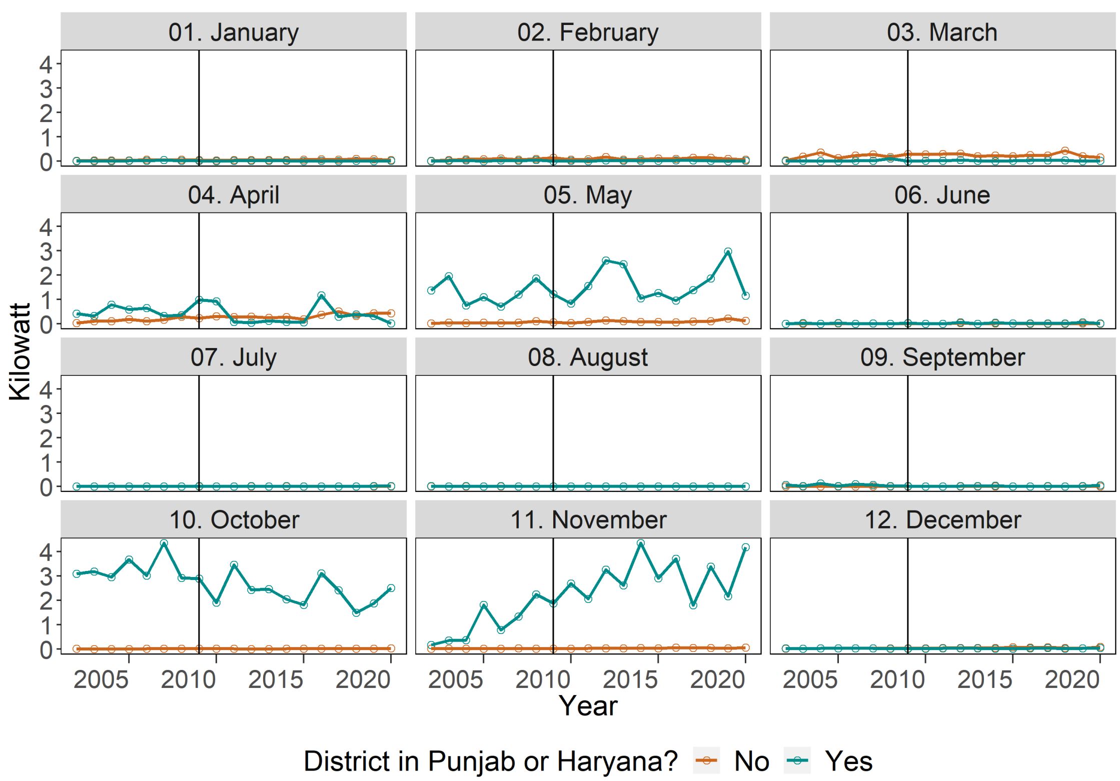


Figure 2: Trends in fire count and fire radiative power (2002-2020) ↩

(a) Average wind directions at origin

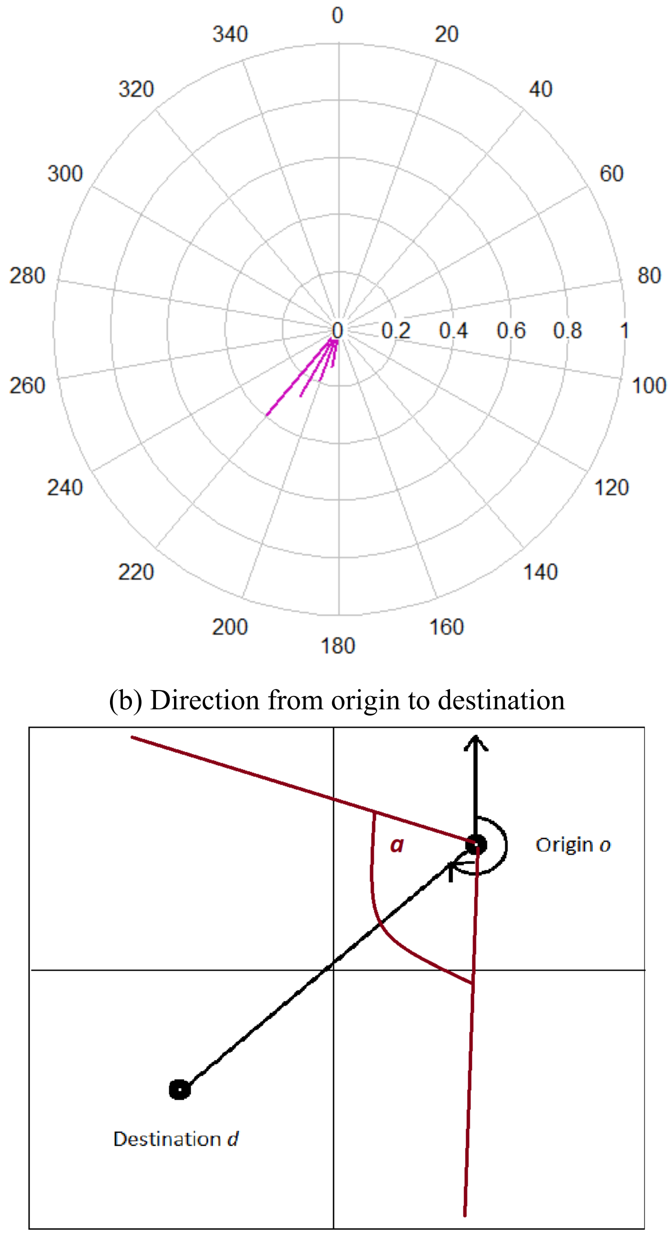


Figure 3: Schematic for Construction of Fire Exposure Instrument. The pink lines on top are fractions of time during the day when the wind was blowing in that bin. ↩

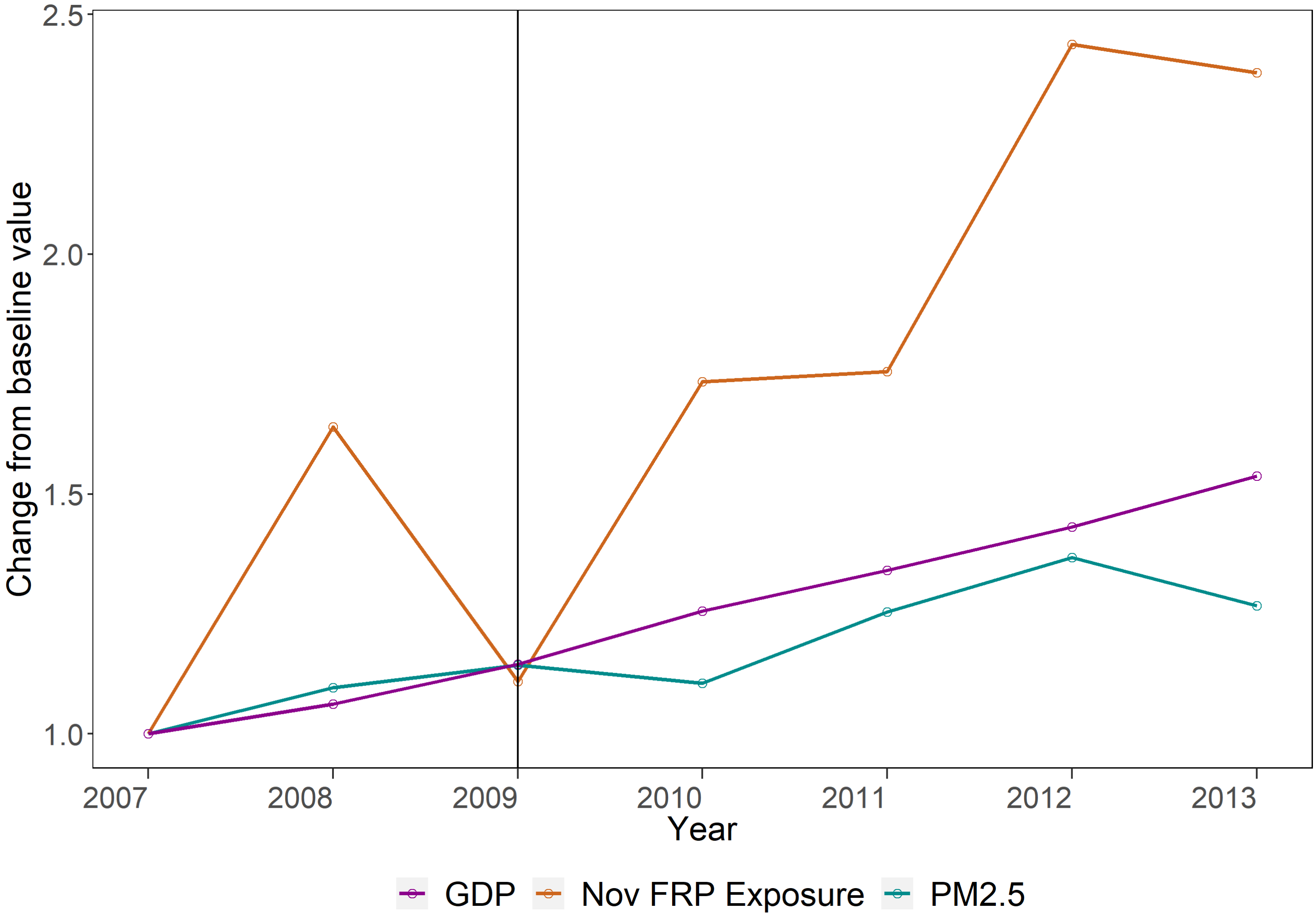


Figure 4: Trend in fire exposure, PM and GDP (2007-2013) ↩

Table 1: Summary Statistics

Variable	N	Mean	SD	Min	Max
Panel A: Monthly Fire Measures and Groundwater Law (2002-2020)					
Count of fires	143640	5.340	32.720	0	1148
Total Fire Radiative Power (mw)	143640	102.75	670.34	0	45044
Groundwater Law Dummy	143640	0.065	0.247	0	1
Panel B: Exposure to Upwind November Fires (FRP-based) with distance cut-off (2007-2013)					
Nov FRP exposure, cut-off = 500	3731	35.167	86.607	0.020	675.725
Nov FRP exposure, cut-off = 600	3731	38.924	87.728	0.053	675.902
Nov FRP exposure, cut-off = 700	3731	42.431	88.123	0.063	675.920
Nov FRP exposure, cut-off = 800	3731	45.904	88.126	0.085	675.985
Nov FRP exposure, cut-off = 900	3731	49.451	87.799	0.087	676.289
Nov FRP exposure, cut-off = 1000	3731	52.759	87.268	0.164	676.300
Panel C: Annual Particulate Matter and GDP (2007-2013)					
Mean PM2.5 (micrograms/m3)	3731	62.517	27.678	17.828	147.946
GDP (Billions of Rupees, Constant 2004)	3731	81.301	164.07	2.414	3728
Panel D: Annual Weather (2007-2013)					
Mean Temperature (°C)	3731	25.011	3.767	-10.369	29.847
Total Rainfall (mm)	3731	2165.9	430.47	0	2809
Mean Relative Humidity (Ratio)	3731	0.640	0.081	0.388	0.852
Mean Surface Pressure (kilo-pascal)	3731	96.85	4.946	56.460	100.83
Mean Windspeed (m/s)	3731	1.437	0.598	0.329	3.831

Notes: All data is at the district level. The sample consists of 530 districts, except for Panel A which consists of 630 districts (out of 640 census 2011 districts). The reduction is due to ICRISAT GDP data only being available between 2007-2013 for a subset of districts. ↩

Table 2: Poisson Estimates of Impact of Groundwater Laws on Monthly Fires

	Fire Count		Fire Radative Power	
	Pre-2009 Mean [SD]	(1)	Pre-2009 Mean [SD]	(2)
January	1.881	-0.749***	18.459	-0.746***
	[3.037]	(0.137)	[38.224]	(0.154)
February	2.384	-0.659**	25.36	-0.809***
	[3.759]	(0.278)	[55.396]	(0.218)
March	2.11	-0.529***	31.073	-0.771***
	[4.482]	(0.145)	[81.6]	(0.154)
April	20.527	-1.09***	440.866	-0.789***
	[27.907]	(0.260)	[601.855]	(0.266)
May	62.546	-0.430***	1330.916	-0.286*
	[72.912]	(0.118)	[1652.531]	(0.157)
June	0.494	0.253	13.74	0.040
	[1.306]	(0.181)	[55.281]	(0.158)
July	0.149	0.542***	2.401	0.726***
	[0.524]	(0.196)	[9.039]	(0.265)
August	0.36	-1.10***	6.031	-1.28***
	[1.077]	(0.289)	[19.759]	(0.249)
September	4.625	-1.83***	58.319	-1.94***
	[10.988]	(0.182)	[143.641]	(0.185)
October	192.287	-0.857***	2946.382	-0.855***
	[268.594]	(0.118)	[4473.063]	(0.158)
November	49.846	0.429***	788.759	0.542***
	[130.006]	(0.116)	[2265.141]	(0.159)
December	3.084	-0.600***	26.838	-0.526***
	[3.997]	(0.162)	[40.39]	(0.175)
Observations	4018	140,372	4018	140,372
Pseudo R2		0.784		0.797
Years	2002-2018	2002-2018	2002-2018	2002-2018
Districts	41	630	41	630

continued

State x Month FE	X	X
Year FE	X	X
District FE	X	X

Notes: Years 2002-2018. Columns 1 and 3 provide mean and SD of fire count and fire strength before 2009 **in Punjab and Haryana**. Columns labeled (1) and (2) provide Poisson estimates. Standard errors are clustered at district and State x Year. *p<0.1; **p<0.05; ***p<0.01. ↩

Table 3: Impact of distance-weighted November fire exposure on PM2.5

	Dependent Variable: log(PM)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Fixed Effects Model						
log(Nov FRP Exposure)	0.011***	0.016***	0.022***	0.027***	0.029***	0.028***
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
Observations	3,731	3,731	3,731	3,731	3,731	3,731
Within R2	0.539	0.542	0.546	0.550	0.551	0.549
Panel B: First Differences Model						
log(Nov FRP Exposure)	0.008***	0.009***	0.010***	0.031***	0.032***	0.031***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)
Observations	3,178	3,178	3,178	3,201	3,201	3,201
Within R2	0.171	0.172	0.173	0.197	0.197	0.193
Distance Cutoff	[500 km]	[600 km]	[700 km]	[800 km]	[900 km]	[1000 km]
Weather Controls	X	X	X	X	X	X
District and Year FE	X	X	X	X	X	X
District x Time Trend	X	X	X	X	X	X

Notes: Years 2007-2013. The sample is limited to districts for which GDP data is available. Each column of Panel A and B provides estimates from the same regression specification but with a different distance cut-off when constructing the FRP exposure instrument. Estimates in each panel are equivalent to the first stage for columns 3 and 5 in table 4. Standard errors are clustered at district and Region x Year. *p<0.1; **p<0.05; ***p<0.01. ↩

Table 4: Impact of Air Pollution (PM2.5) on GDP

	Dependent Variable				
	log(GDP)			Δ log(GDP)	
	(1)	(2)	(3)	(4)	(5)
Panel A: OLS and 2SLS Results					
log(PM2.5)	0.147***	-0.008	-0.159	-0.030**	-0.179***
	(0.035)	(0.016)	(0.097)	(0.014)	(0.069)
Observations	3,731	3,731	3,731	3,201	3,201
R2	0.996	0.999	0.999	0.379	0.326
Weather Controls	X	X	X	X	X
District and Year FE	X	X	X	X	X
District x Time Trend		X	X		
First Differences				X	X
2SLS Estimate			X		X
Panel B: First Stage Results					
log(Nov FRP Exposure)			0.029***		0.032***
			(0.006)		(0.006)
Cragg-Donald F-stat			101.4		116.5
Kleibergen-Paap F-stat			25.3		26.4

Notes: Years 2007-2013. The sample is limited to districts for which GDP data is available. Panel A, columns 1-3, show estimates for both OLS and 2SLS regressions of log GDP level on log PM2.5, starting without a time trend, then controlling for a time trend and finally conducting 2SLS with time trend. Columns 4 of panel A shows an OLS estimate using first differences while column 5 instruments for first difference of log PM with first difference of log Nov Exposure (900 km cut-off). Standard errors are clustered at district and Region x Year. *p<0.1; **p<0.05; ***p<0.01. ↩

8 Bibliography

Adhvaryu, Achyuta, Namrata Kala, and Anant Nyshadham. 2022. "Management and Shocks to Worker Productivity." Journal of Political Economy 130 (1): 1–47. https://doi.org/10.1086/717046.

Andini, Ade, Sébastien Bonnet, Patrick Rousset, and Udin Hasanudin. 2018. "Impact of Open Burning of Crop Residues on Air Pollution and Climate Change in Indonesia." Current Science 115 (12): 2259–66.

Andrews, Isaiah, James H. Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." Annual Review of Economics 11 (1): 727–53. https://doi.org/10.1146/annurev-economics-080218-025643.

Assunção, Juliano, Clarissa Gandour, and Romero Rocha. 2022. "DETERring Deforestation in the Amazon: Environmental Monitoring and Law Enforcement." American Economic Journal: Applied Economics. https://doi.org/10.1257/app.20200196.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." Review of Environmental Economics and Policy 7 (2): 181–98. https://doi.org/10.1093/reep/ret016.

Behrer, A. Patrick. 2019. "The Impact of Anti-Poverty Efforts on Indian Agriculture and Air Pollution." Unpublished manuscript.

Bellemare, Marc F., and Casey J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." Oxford Bulletin of Economics and Statistics 82 (1): 50–61. https://doi.org/10.1111/obes.12325.

Berge, Laurent, Sebastian Krantz, and Grant McDermott. 2022. "Fixest: Fast Fixed-Effects Estimations."

Blackburn, Mckinley L. 2015. "The Relative Performance of Poisson and Negative Binomial Regression Estimators." Oxford Bulletin of Economics and Statistics 77 (4): 605–16. https://doi.org/10.1111/obes.12074.

Blakeslee, David, Ram Fishman, and Veena Srinivasan. 2020. "Way Down in the Hole: Adaptation to Long-Term Water Loss in Rural India." American Economic Review 110 (1): 200–224. https://doi.org/10.1257/aer.20180976.

Bondy, Malvina, Sefi Roth, and Lutz Sager. 2020. "Crime Is in the Air: The Contemporaneous Relationship Between Air Pollution and Crime." Journal of the Association of Environmental and Resource Economists 7 (3): 555–85. https://doi.org/10.1086/707127.

Borgschulte, Mark, David Molitor, and Eric Zou. 2022. "Air Pollution and the Labor Market: Evidence from Wildfire Smoke." {NBER} working paper 29952. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29952.

Burgess, Robin, Francisco J M Costa, and Ben Olken. 2019. "The Brazilian Amazon's Double Reversal of Fortune." Preprint. SocArXiv. https://doi.org/10.31235/osf.io/67xg5.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." Journal of Econometrics 225 (2): 200–230.

Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2019. "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China." American Economic Journal: Applied Economics 11 (1): 151–72. https://doi.org/10.1257/app.20160436.

Dechezleprêtre, Antoine, Nicholas Rivers, and Balazs Stadler. 2019. "The Economic Cost of Air Pollution: Evidence from Europe." {OECD} {Working} paper 1584. Paris. https://doi.org/10.1787/56119490-en.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." American Economic Journal: Macroeconomics 4 (3): 66–95. https://doi.org/10.1257/mac.4.3.66.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." American Economic Review 109 (12): 4178–4219. https://doi.org/10.1257/aer.20180279.

Down To Earth. 1999. "Parched Punjab." Parched Punjab. https://www.downtoearth.org.in/coverage/parched-punjab-20535.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. "The Value of Regulatory Discretion: Estimates From Environmental Inspections in India." Econometrica 86 (6): 2123–60. https://doi.org/10.3982/ECTA12876.

Fu, Shihe, V Brian Viard, and Peng Zhang. 2021. "Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China." The Economic Journal 131 (640): 3241–73. https://doi.org/10.1093/ej/ueab033.

Ghanem, Dalia, and Junjie Zhang. 2014. "'Effortless Perfection:' Do Chinese Cities Manipulate Air Pollution Data?" Journal of Environmental Economics and Management 68 (2): 203–25. https://doi.org/10.1016/j.jeem.2014.05.003.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." American Economic Review 102 (7): 3652–73. https://doi.org/10.1257/aer.102.7.3652.

Greenstone, Michael. 2021. "AQLI India Fact Sheet."

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." American Economic Review 104 (10): 3038–72. https://doi.org/10.1257/aer.104.10.3038.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution." American Economic Review: Insights, April. https://doi.org/10.1257/aeri.20200373.

Grossman, Gene M., and Alan B. Krueger. 1995. "Economic Growth and the Environment*." The Quarterly Journal of Economics 110 (2): 353–77. https://doi.org/10.2307/2118443.

Hagerty, Nick. 2021. "Adaptation to Surface Water Scarcity in Irrigated Agriculture." Unpublished manuscript.

Hammer, Melanie S., Aaron van Donkelaar, Chi Li, Alexei Lyapustin, Andrew M. Sayer, N. Christina Hsu, Robert C. Levy, et al. 2020. "Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018)." Environmental Science & Technology 54 (13): 7879–90. https://doi.org/10.1021/acs.est.0c01764.

Jack, B Kelsey, Seema Jayachandran, Namrata Kala, and Rohini Pande. 2022. "Money (Not) to Burn: Payments for Ecosystem Services to Reduce Crop Residue Burning." Unpublished manuscript.

Jayachandran, Seema. 2022. "How Economic Development Influences the Environment." Annuul Review of Economics 14 (September): 1–30. https://doi.org/doi.org/10.1146.

Kahn, Matthew E., Pei Li, and Daxuan Zhao. 2015. "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives." American Economic Journal: Economic Policy 7 (4): 223–42. https://doi.org/10.1257/pol.20130367.

Kataki, P. K., P. Hobbs, and B. Adhikary. 2001. "The Rice-Wheat Cropping System of South Asia." Journal of Crop Production 3 (2): 1–26. https://doi.org/10.1300/J144v03n02_01.

Lin, Wei, and Jeffrey M. Wooldridge. 2019. "Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models." In Panel Data Econometrics, edited by Mike Tsionas, 21–43. Academic Press. https://doi.org/10.1016/B978-0-12-814367-4.00002-2.

Lipscomb, Molly, and Ahmed Mushfiq Mobarak. 2017. "Decentralization and Pollution Spillovers: Evidence from the Re-Drawing of County Borders in Brazil*." The Review of Economic Studies 84 (1): 464–502. https://doi.org/10.1093/restud/rdw023.

Lohan, Shiv Kumar, H. S. Jat, Arvind Kumar Yadav, H. S. Sidhu, M. L. Jat, Madhu Choudhary, Jyotsna Kiran Peter, and P. C. Sharma. 2018. "Burning Issues of Paddy Residue Management in North-West States of India." Renewable and Sustainable Energy Reviews 81 (January): 693–706. https://doi.org/10.1016/j.rser.2017.08.057.

McDonald, Andrew J., Balwinder-Singh, Alwin Keil, Amit Srivastava, Peter Craufurd, Avinash Kishore, Virender Kumar, et al. 2022. "Time Management Governs Climate Resilience and Productivity in the Coupled Rice–Wheat Cropping Systems of Eastern India." Nature Food 3 (7): 542–51. https://doi.org/10.1038/s43016-022-00549-0.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." Journal of Business & Economic Statistics 31 (3): 358–69. https://doi.org/10.1080/00401706.2013.806694.

Parikh, Kirit S., A. Ganesh-Kumar, and Gangadhar Darbha. 2003. "Growth and Welfare Consequences of Rise in MSP." Economic and Political Weekly 38 (9): 891–95.

Pingali, Prabhu L. 2012. "Green Revolution: Impacts, Limits, and the Path Ahead." Proceedings of the National Academy of Sciences 109 (31): 12302–8. https://doi.org/10.1073/pnas.0912953109.

Ranson, Matthew. 2014. "Crime, Weather, and Climate Change." Journal of Environmental Economics and Management 67 (3): 274–302. https://doi.org/10.1016/j.jeem.2013.11.008.

Rodrik, Dani. 2008. "Second-Best Institutions." American Economic Review 98 (2): 100–104. https://doi.org/10.1257/aer.98.2.100.

Ryan, Nick, and Anant Sudarshan. 2020. "Do India's Farmers Use Too Much Water?" {VoxDev}.

Schlenker, Wolfram, and W. Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." The Review of Economic Studies 83 (2): 768–809. https://doi.org/10.1093/restud/rdv043.

Shyamsundar, P., N. P. Springer, H. Tallis, S. Polasky, M. L. Jat, H. S. Sidhu, P. P. Krishnapriya, et al. 2019. "Fields on Fire: Alternatives to Crop Residue Burning in India." Science (New York, N.Y.), August. https://doi.org/10.1126/science.aaw4085.

Singh, Karam. 2009. "Act to Save Groundwater in Punjab: Its Impact on Water Table, Electricity Subsidy and Environment." Agricultural Economics Research Review 22 (January).

Stern, David. 2017. "The Environmental Kuznets Curve After 25 Years." Journal of Bioeconomics 19 (1): 7–28.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, edited by Donald W. K. Andrews and James H. Stock, 80–108. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511614491.006.

Subramanian, Kapil. 2017. "How Delhi's Air Pollution Crisis May Have Origins in the Green Revolution." The Hindu, November.

Vallero, Daniel. 2014. Fundamentals of Air Pollution - 5th Edition. 5th ed. Elsevier.

Wooldridge, Jeffrey M. 1999a. "Quasi-Likelihood Methods for Count Data." In Handbook of Applied Econometrics Volume 2: Microeconomics, 321–68. John Wiley & Sons, Ltd. https://doi.org/10.1111/b.9780631216339.1999.00009.x.

———. 1999b. "Distribution-Free Estimation of Some Nonlinear Panel Data Models." Journal of Econometrics 90 (1): 77–97. https://doi.org/10.1016/S0304-4076(98)00033-5.

———. 2010. Econometric Analysis of Cross Section and Panel Data. 2nd ed. Cambridge, MA, USA: MIT Press.

World Bank. 2018. "Power Sector Distortions Cost India Billions." World Bank.

———. 2021. "Addressing Groundwater Depletion: Lessons from India, the World's Largest User of Groundwater."

9 Appendix

Table A.5: Impact of Groundwater Laws on Monthly Fires in Punjab and Haryana - Robustness

	Fire Count			Fire Radative Power		
	(1)	(2)	(3)	(4)	(5)	(6)
January	-0.143**	-0.747***	-0.746***	-0.164***	-0.765***	-0.918***
	(0.059)	(0.133)	(0.078)	(0.054)	(0.156)	(0.080)
February	-0.368***	-0.737***	-0.786***	-0.407***	-0.867***	-0.949***
	(0.080)	(0.252)	(0.205)	(0.077)	(0.202)	(0.229)
March	-0.398***	-0.668***	-0.914***	-0.500***	-0.957***	-1.27***
	(0.049)	(0.142)	(0.080)	(0.057)	(0.156)	(0.137)
April	-0.790***	-1.06***	0.198	-0.735**	-0.766***	0.028
	(0.289)	(0.254)	(0.186)	(0.335)	(0.263)	(0.183)
May	-0.118*	-0.534***	-0.097	-0.072	-0.410***	-0.086
	(0.059)	(0.118)	(0.076)	(0.072)	(0.158)	(0.080)
June	-0.078	0.243	0.591***	-0.233***	0.054	0.577***
	(0.052)	(0.165)	(0.170)	(0.077)	(0.157)	(0.118)
July	-0.152*	0.517***	1.20***	-0.159***	0.698***	1.55***
	(0.076)	(0.150)	(0.401)	(0.055)	(0.220)	(0.224)
August	-0.495***	-1.05***	-0.035	-0.699***	-1.25***	0.333*
	(0.082)	(0.295)	(0.070)	(0.065)	(0.254)	(0.182)
September	-1.09***	-1.95***	-1.42***	-1.28***	-2.07***	-1.52***
	(0.258)	(0.193)	(0.144)	(0.250)	(0.195)	(0.141)
October	-0.223***	-0.817***	-0.380***	-0.190*	-0.821***	-0.535***
	(0.074)	(0.119)	(0.070)	(0.107)	(0.159)	(0.079)
November	1.05***	0.509***	0.238***	1.15**	0.613***	0.124
	(0.382)	(0.120)	(0.081)	(0.430)	(0.161)	(0.091)
December	-0.370**	-0.677***	-0.323**	-0.359*	-0.640***	-0.355***
	(0.154)	(0.157)	(0.132)	(0.178)	(0.166)	(0.103)
Observations	56,082	149,257	43,904	56,082	149,257	43,904
Specification	OLS	Poisson	Poisson	OLS	Poisson	Poisson
Years	2002-2018	2000-2018	2007-2013	2002-2018	2000-2018	2007-2013

continued

Districts	630	630	630	630	630	630
State x Month FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
District FE	X	X	X	X	X	X

Notes: Provides robustness checks for table 2. Columns 1 and 4 conduct OLS estimation with log(fire count) and log(FRP) as the dependent variables. Columns 2 and 5 conduct the Poisson estimation with fires data from 2000 and 2001, when the fires are less reliable. Columns 3 and 6 conduct Poisson estimation by restricting sample to data from the 530 districts over 2007-2013 which have GDP data available. Standard errors are clustered at district and State x Year. *p<0.1; **p<0.05; ***p<0.01. ↩

Table A.6: Impact of November fire exposure without distance weighting on PM2.5

	Dependent Variable: log(PM)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Fixed Effects Model						
log(Nov FRP Exposure)	0.014***	0.021***	0.028***	0.032***	0.032***	0.028***
	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	3,731	3,731	3,731	3,731	3,731	3,731
Within R2	0.542	0.545	0.551	0.554	0.553	0.549
Panel B: First Differences Model						
log(Nov FRP Exposure)	0.010***	0.011***	0.011***	0.034***	0.033***	0.030***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.006)
Observations	3,178	3,178	3,178	3,201	3,201	3,201
Within R2	0.173	0.174	0.174	0.201	0.198	0.191
Distance Cutoff	[500 km]	[600 km]	[700 km]	[800 km]	[900 km]	[1000 km]
Weather Controls	X	X	X	X	X	X
District and Year FE	X	X	X	X	X	X
District x Time Trend	X	X	X	X	X	X

Notes: Years 2007-2013. Robustness to dropping distance from construction of exposure instrument in table 3. The sample is limited to districts for which GDP data is available. Each column of Panel A and B provides estimates from the same regression specification but with a different distance cut-off when constructing the FRP exposure instrument. Standard errors are clustered at district and Region x Year. *p<0.1; **p<0.05; ***p<0.01. ↩

Table A.7: Impact of distance-weighted Monthly fire exposure on annual PM2.5

	Dependent Variable: log(PM)					
	Jan	Feb	Mar	Apr	May	Jun
	Panel A: Estimates for January to June					
log(Monthly FRP Exposure)	0.007	0.017***	0.011**	-0.002	0.007	-0.006***
	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)	(0.002)
Observations	3,731	3,731	3,731	3,731	3,731	3,718
Within R2	0.535	0.540	0.537	0.534	0.535	0.537
	Jul	Aug	Sep	Oct	Nov	Dec
	Panel B: Estimates for July to December					
log(Monthly FRP Exposure)	-0.004	-0.002	0.003	-0.010*	0.029***	0.012*
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.007)
Observations	3,697	3,726	3,730	3,731	3,731	3,731
Within R2	0.534	0.534	0.535	0.536	0.551	0.536
Distance Cutoff	[900 km]	[900 km]	[900]	[900 km]	[900 km]	[900 km]
Weather Controls	X	X	X	X	X	X
District and Year FE	X	X	X	X	X	X
District x Time Trend	X	X	X	X	X	X

Notes: Years 2007-2013. The sample is limited to districts for which GDP data is available. Each single column in Panel A and B displays estimates for the regression of annual PM2.5 on exposure to fires during that month of the year only, using the same specification as in table 3. Standard errors are clustered at district and Region x Year. *p<0.1; **p<0.05; ***p<0.01. ↩