

Pollution, agricultural productivity, and development: Evidence from coal plants in India*

Joshua D. Merfeld[†]

2023-02-20

Abstract

I show that emissions from a common energy source lead to decreased productivity and growth. The rapid increase in population in India from the 1990s through the 2010s led to the rapid construction of new coal power plants throughout the country. Using this roll-out combined with random changes in wind direction, I document negative effects of pollution from coal power plants on agricultural productivity throughout the country. Results suggest this decrease may be driven by a decline in labor allocation of the oldest half of my sample, who greatly decrease labor allocated to both farm and non-farm employment when exposed to pollution. These impacts spill over into aggregate economic growth, as well; nightlight growth is slower following years of higher exposure to pollution.

Keywords: pollution, productivity, agriculture, labor, India

JEL Codes: H40, I15, J22, O13, Q52, Q53

*I thank Marc Bellemare and David Sungho Park for comments and suggestions that greatly improved the quality of the paper. The usual disclaimer applies. All scripts used in this paper are publicly available on my GitHub page: github.com/JoshMerfeld/pollution_development

[†]KDI School of Public Policy and Management and IZA; merfeld@kdis.ac.kr

1 Introduction

Coal is one of the most common energy sources throughout the world. In 2021, energy generation from coal reached an all-time high,¹ driven in part by rapid population growth and rising energy prices. India is no exception to these trends. From 1990 to 2010, India's population grew by more than 40 percent.² This population growth led to a large increase in demand for power. While natural gas has become a popular alternative in much of the world, the Indian government met much of the increased demand through the construction of coal power plants. During those same two decades, power generation from all coal units of at least 30 megawatts in the country increased by almost 140 percent, from 42.4 gigawatts to more than 100 gigawatts.³ There are no doubt positive effects from this increase in power generation; for example, electricity may lead to economic and productivity growth (Dinkelman 2011; Kline and Moretti 2014; Rud 2012; Van de Walle et al. 2017), though electrification alone may not be a sufficient condition (K. Lee, Miguel, and Wolfram 2020).

These improvements do not come without downsides, however, especially when it comes to power derived from coal. Pollutants from coal contribute to climate change and the phase out of coal energy is therefore a major climate goal (IEA 2022). Emissions from the production and burning of coal are also directly harmful to health and the environment, including pollutants like sulfur dioxide, nitrogen oxides, and mercury, to name but a few.⁴ While the use of coal had been declining, recent events – including Russia's invasion of Ukraine – have led to an increase in the demand for coal.⁵

I study the effects of emissions from coal power plants on agricultural productivity, labor allocation, and economic growth in India. Previous work has suggested that there may be harmful direct effects of air pollution (Heck et al. 1982; Marshall et al. 1997) and water pollution (Reddy and Behera 2006) on agricultural productivity, including work showing that gold mines – which release large amounts of pollutants into the surrounding environment – lead to lower agricultural productivity in Ghana (Aragón and Rud 2016). These effects can work through the well known effects on human health and productivity, but also directly through land productivity. Black carbon in soot, for example, can directly affect solar radiation reaching plants on the surface (Ramanathan and Carmichael 2008; Burney and Ramanathan 2014). Relative to previous work in economics, this paper tackles the broader case of coal plants, which are a leading source of power throughout the world. I show that air pollution from these plants can lead to lower agricultural productivity and even

¹www.iea.org/reports/coal-fired-electricity

²data.worldbank.org

³globalenergymonitor.org/

⁴www.eia.gov/energyexplained/coal/coal-and-the-environment.php

⁵www.npr.org/2022/08/15/1117560560/a-rising-demand-for-coal-amidst-war-in-ukraine

decreased economic growth.

Identifying the causal effects of pollution is difficult. For starters, the construction of coal plants is likely to be endogenous, though the relationship between construction and economic conditions is not *a priori* clear. On the one hand, governments may locate new plants in fast-growing areas, while on the other hand, they may decide the exact location of plants based on the political power of local citizens (Kopas et al. 2020). In the case of the data used here, coal plants are constructed in areas with higher population and higher literacy (Table 1). In terms of agriculture, higher productivity predicts more coal plants in a given year as well as the construction of a coal plant over the next decade, indicating possible bias when estimating OLS regressions of agricultural productivity on pollution levels.

To identify the effects of air pollution on agricultural productivity, I instead combine the roll out of coal plants in India with plausibly exogenous changes in wind direction – in a similar spirit to Deryugina et al. (2019) – to measure the effects of pollution from coal plants on key development outcomes. I first plot the location of coal plants in India by year. Using a village-level shapefile, I identify all villages that are located within 30km of a coal plant at any time during the sample period, which comprises the 1990s to the 2010s, with differences depending on data availability. To measure exposure to emissions from coal plants, I calculate the direction from each coal plant to all villages within 30km. Using daily wind data and these spatial variables, I create an identifier separately for each of the one-hundred-thousand-plus villages for whether a given village is located downwind from a coal plant on any given day. I then aggregate this exposure variable based on the desired analysis; for agricultural productivity, for example, I create an exposure variable measuring the total number of exposure days in the five-month agricultural season, separately for the main monsoon season and the smaller winter season. Consistent with expectations, wind direction is a strong predictor of particulate matter at the monthly level. More days being located downwind from a coal power plant leads to higher average concentrations of particulate matter 2.5.

Using newly available estimates of India-wide agricultural (land) productivity, results show that the total number of days of being downwind from a coal power plant has a negative effect on overall agricultural productivity. Just a single additional day of being downwind from a coal power plant leads to a decrease in predicted yields of 0.3 percent. The within-village absolute deviation of wind exposure is around eight days, meaning that median changes in wind direction are responsible for fluctuations in agricultural productivity of around 2.4 percent, with changes of more than five percent from season to season being common. Using wind direction as an instrument for particulate matter, an increase of PM 2.5 by one microgram per cubic meter leads to an even larger decrease in agricultural productivity, at around 20 percent. This IV strategy is similar to that in Deryugina et al. (2019), though they use general “high-pollution” wind directions while

I concentrate only on wind directions from the location of a coal power plant. The IV estimates indicate that the effect of pollution is more negative than naïve OLS estimates indicate, which is consistent with the evidence on the construction of coal plants; higher levels of pollution are correlated with higher levels of agricultural productivity and economic growth, a relationship that the IV strategy adjusts downwards.

These effects are apparently concentrated in villages that do not experience the most extreme exposure to pollution; villages with higher maximum exposure experience effects only around 10 percent as large as villages with lower maximum exposure. Additional heterogeneity analysis shows that, despite the results with respect to the maximum, the effects are increasing in *mean* exposure. I also show that the effects are largest in villages with the highest initial agricultural productivity and that pollution and rainfall shocks are compounding; higher exposure along with less rainfall increases the negative effect of either shock individually. As weather shocks become more common across the globe, the rise of pollution in India may be particularly problematic for farmers.

Robustness checks show that this effect is concentrated in the years after the construction of the power plant, which helps decrease concerns around endogeneity since, conditional on the village fixed effect and weather (rainfall and temperature), wind direction across years and seasons is exogenous. Additionally, leads for wind direction do not predict current agricultural productivity, again supporting the identifying assumptions.

While I am not able to explicitly test effects on land productivity, I look at how adults allocate labor over the last seven days using the National Sample Survey and creating a district-level exposure measure based on the timing of the survey for each household. An additional day of wind over the previous week decreases overall labor allocation by 0.055 days. However, the overall results show stark heterogeneity based on age; the older half of the sample greatly decreases labor allocation in response to pollution exposure while the younger half increases overall labor allocation. Older adults decrease labor allocation to both farm and non-farm labor, while younger adults increase labor allocated to farm labor. This is suggestive evidence of productivity effects being driven by the health and productivity effects on individuals, since older adults are more susceptible to pollution and they also greatly decrease non-farm employment, which presumably will no suffer any possible direct effects of pollution on productivity. The preponderance of results seems to indicate that effects on labor are the most likely mechanism.

The final set of results explores whether there are impacts of pollution on aggregate economic development. Using nightlights, I show that lagged wind – i.e. pollution exposure in the prior year – has a negative impact on economic development as proxied by nightlights. Going from the 25th percentile to the 75th percentile of wind exposure is synonymous with a decrease in the growth of nightlights within a village by around 0.7

percent. While this is a relatively small magnitude, it is worth noting that there are more than 100,000 villages in this sample. While the effect for a single village might be small, the overall, aggregate effect on growth is quite substantial. To put this number in perspective, Castelló-Climent, Chaudhary, and Mukhopadhyay (2018) argue that a standard-deviation increase in the share of adults with a university education leads to 12 percent higher nightlights in India, while Storeygard (2016) shows that a four-fold increase in oil prices over six years leads to seven percent lower nightlights at a distance of around 500km from the nearest large port in sub-Saharan Africa. In other words, the 0.7 percent change in nightlights is quite large given this is annual variation.

This paper contributes to several strands of literature. First, we already have evidence of effects of pollution on different forms of labor activity. For example, general levels of pollution decrease productivity of call center works in China (Chang et al. 2019), lead people to perform worse on cognitive functions (Ebenstein, Lavy, and Roth 2016; Wen and Burke 2022), and also lead to lower levels of farm labor productivity in California, driven by changes in ozone concentration (Graff Zivin and Neidell 2012). Also in China, He, Liu, and Salvo (2019) find small but significant negative effects of pollution on manufacturing productivity, while Chen, Oliva, and Zhang (2022) show pollution leads to labor out migration. Perhaps relevant to my labor allocation results, Hanna and Oliva (2015) show that decreases in pollution lead to increases in labor allocation, while I show the same relationship among the oldest cohort. The papers most related to this one are Aragón and Rud (2016), who show that gold mines in Ghana have substantial negative impacts on agricultural productivity in the surrounding areas, and Burney and Ramanathan (2014), who model the effects of climate change and pollutants on agricultural productivity in India. This paper differs in two important ways. First, I explore effects of pollution from coal plants, which are much more ubiquitous than gold mines across the world. Second, I use a different identification strategy from either paper, relying on wind blowing pollution from coal plants.

I also contribute to the more general literature on the effects of pollution on health. Pollution has substantial negative effects on health in both developing and developed countries (Arceo, Hanna, and Oliva 2016) and we have known of these negative effects for many years (Brunekreef and Holgate 2002; Kampa and Castanas 2008; Pope III and Dockery 2006). Currie et al. (2014) review the literature on the health effects of pollution during adolescence; we have robust evidence of pollution exposure leading to increases in infant mortality, including in developing countries (Heft-Neal et al. 2018, 2019). Since these effects are often driven by heart and lung disease and asthma (Kampa and Castanas 2008), this may explain the labor allocation effects I identify in this paper.

The rest of the paper is organized as follows. I first discuss data and methodology. I break this section

up into subsection 2.1, which covers the multiple sources of data used in this paper; subsection 2.2, which discusses the empirical strategy and validates the wind direction variable using data on particulate matter. I present the main results in section 3 before concluding in section 4.

2 Data and methods

In this section, I describe the different data sources and the identification strategy. I also spend time validating the use of wind direction as a proxy for particulate matter as well as looking at the correlates of construction of coal power plants.

2.1 Data

The main goal of this paper is to examine whether exposure to emissions from coal plants affects agricultural productivity, labor allocation, and economic growth. I use several sources of data, which I describe here and in Table A1 in the appendix.

The first set of data lists the location of coal plant across the globe. This data comes from Global Energy Monitor⁶ and lists all units generating at least 30 megawatts of electricity. The data on coal plants includes the year of opening (and, if applicable, the year of retirement), the GPS (latitude/longitude) location of each plant, and the amount of power produced by the units at each plant. For this paper, I do not use the information on the capacity of the plant.

Figure 1 displays the location of coal plants in India for two specific years: 1990 and 2010. There is a clear increase in the prevalence of coal plants across the country over the two decades. Additionally, the overall capacity from coal plants increased from 42.4 gigawatts to 100.4 gigawatts, an increase of 136.6 percent in just 20 years. This increase is driven both by the roll out of new plants as well as the construction of new units at existing plants, the latter of which is not shown in Figure 1.

The second dataset includes agricultural productivity for both the monsoon and winter seasons, from 2002 to 2018. To match other data, I only use the data up to 2013. This data comes from Gangopadhyay et al. (2022) and estimates land productivity (i.e. yield, in tons per hectare) for the major crops in India, using satellite-derived data. While there is a wealth of literature on measurement error in land (Desiere and Jolliffe 2018; Abay et al. 2022), the authors use a static cropland mask for the productivity estimates. This means that any measurement error will be constant across time and subsumed by the village fixed effects. While the

⁶globalenergymonitor.org/projects/global-coal-plant-tracker/

Figure 1: Coal plants in India from 1990 to 2010

Figure A: 1990

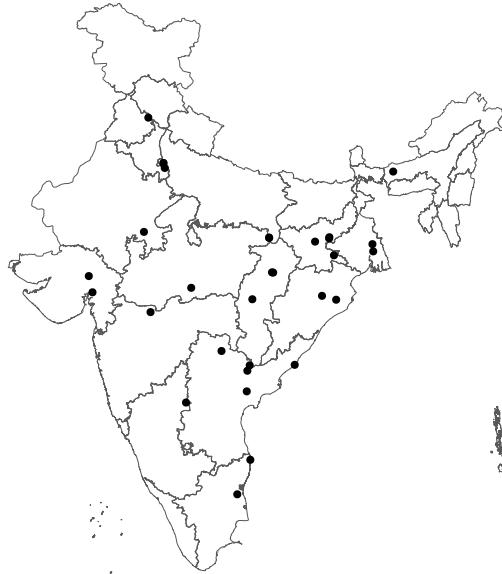
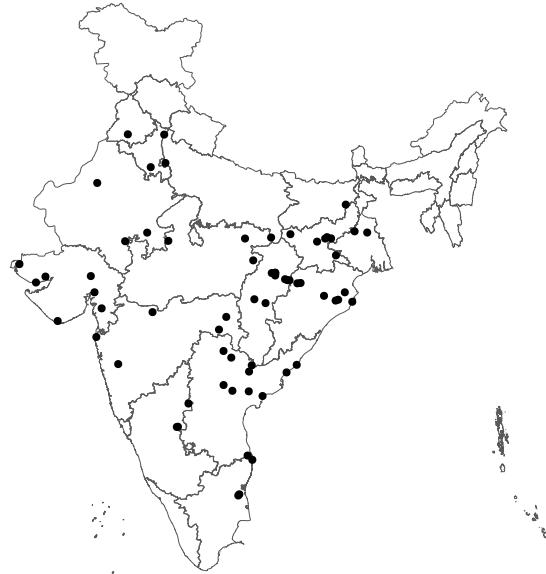


Figure B: 2010



Note: The top figure shows the location of coal plants in 1990. The bottom figure shows the location of coal plants in 2010.

measure of output is also noisy, the use of satellite-derived data and a static cropland mask reduces concerns about measurement error being responsible for the results below. The authors define the monsoon season as June to October and the winter season as November to March. I keep these definitions when matching across data below. This data is also publicly available from Nature’s data-sharing website.⁷ Since the main regressions below use village fixed effects, any differences in cropping patterns across areas should not bias estimates on changes, as long as these patterns do not change markedly in response to pollution.

The agricultural productivity data is available as raster files with a resolution of 500m. To aggregate this data up to a useful administrative unit, I use village-level shapefiles provided by Asher et al. (2021) and publicly available on the SHRUG platform.⁸ I aggregate the agricultural productivity data to the village level by extracting mean productivity for each feature in the shapefile. I do this separately for each season – monsoon and winter – and each year, from 2002 to 2013.

To measure exposure to pollution from coal plants, I first locate all village centroids located within 30km from a coal plant in a given year. I choose 30km due to previous research on the effects of (air) pollution (Aragón and Rud 2016; Bencko and Symon 1977; Li and Gibson 2014). I then calculate the direction from coal plants to all village centroids within that 30km radius. To define exposure, I then pull daily wind direction data

⁷springernature.figshare.com

⁸www.devdatalab.org/shrug_download/

from the National Center for Atmospheric Research.⁹ For each day, I document whether the wind is blowing towards each village centroid.¹⁰ I then temporally aggregate this daily data depending on the temporal definition of the corresponding outcome. For example, agricultural productivity is defined across five months (e.g. the monsoon season is from June to October) so I count the total days a given village is exposed to wind blowing from *any* coal plant within 30km during those five months.

The top panel in Figure 2 is an example wind direction raster. The raster shows wind direction for the entirety of India on January 1st, 2010, as well as the location of coal power plants on that date. The prevailing winds on the date differ across the country and, though not shown in the figure, across days. This means that the overall exposure for a given area to coal plant emissions changes across time. I also pull data on particulate matter to help validate this proxy for exposure. This data comes from Hammer et al. (2020) and I also aggregate this to the village using the same method detailed above.

To help understand some of the mechanisms driving the relationship between exposure to emissions and agricultural productivity, I also include household survey data. Unfortunately, I am not aware of any publicly available survey data that covers multiple years and has identifiers below the district level. For example, the ARIS-REDS data¹¹ includes GPS identifiers for one year, but that only allows GPS matching for a total of two years: 1999 and 2007/2008. The DHS, on the other hand, only has state-level identifiers (except for the newest two waves) and is also not a labor survey. Instead, I use the National Sample Survey (NSS) from the Ministry of Statistics and Programme Implementation.

The NSS is a nationally representative survey that was conducted every couple of years until relatively recently. Many rounds of the NSS also included a module on employment that asked about labor allocation over the previous seven days. I use the 61st, 62nd, 64th, 66th, and 68th rounds of the NSS, all of which include the employment module. The publicly available data also includes the date of survey, which allows me to match pollution exposure to the exact day of the survey.

The biggest issue with using the NSS is that it includes only district-level identifiers. The main exposure variables, on the other hand, are defined at a lower level of aggregation, at the village. I choose to aggregate this village-level exposure to the district level as shown in the second panel of Figure 2 – the district is a randomly selected district in the data, Guna district in Madhya Pradesh. I use the area overlap of the villages with the district to define a weighted average of exposure for the district at a given time. I calculate the area of overlap for each village and the district as well as the overall area of the district and use these variables to

⁹climatedataguide.ucar.edu/

¹⁰I define “towards” as within five degrees to help take into account that I am using village centroids, using the x and y components of wind speed.

¹¹www.ncaer.org/project/additional-rural-incomes-survey-aris-rural-economic-demographic-survey-reds

Figure 2: Wind direction and aggregation examples (2010-01-01)

Figure A: Wind direction

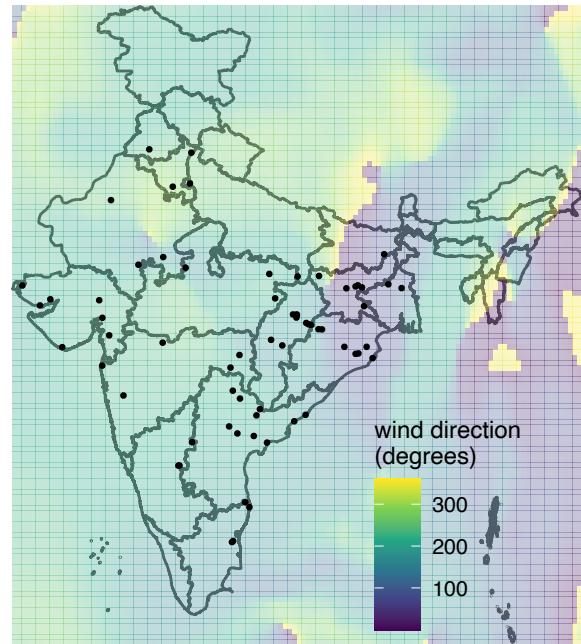
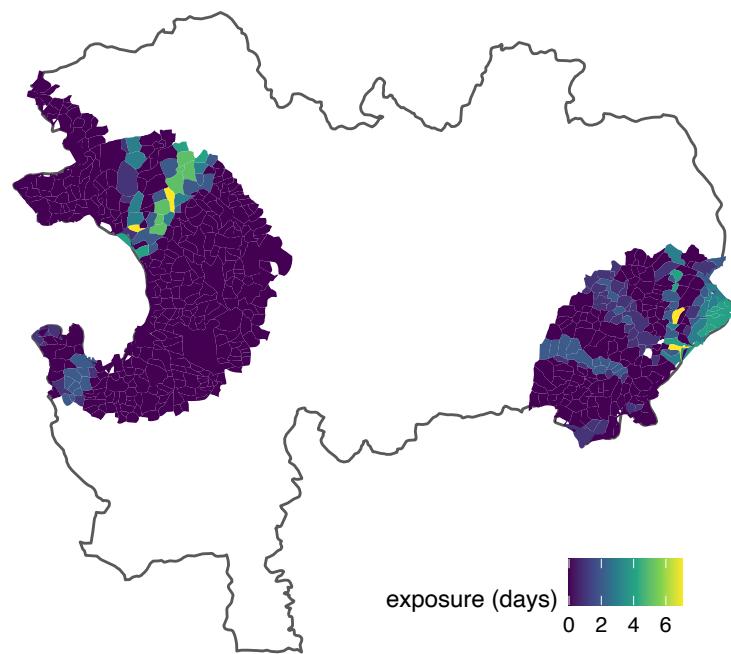


Figure B: Aggregation to district



Note: The top figure shows the average wind direction on January 1st, 2010, with zero degrees indicating north. The points are the location of coal plants on that date. The bottom figure shows the distribution of pollution exposure in a specific district – Guna district in Madhya Pradesh – on the same date.

define weights for individual villages. Any part of the district that does not overlap with one of the villages – the white area in Figure 2 – does not have any exposure and, as such, receives a value of zero for exposure.

Finally, I use nightlights as a measure of economic growth (Henderson, Storeygard, and Weil 2012). I also take this data directly from SHRUG and Asher et al. (2021). This covers 1992 through 2013. Some of the older data may not be ideal (Gibson et al. 2021), but this is the only way to document changes in economic growth at a more disaggregated level and at a regular basis. Although satellite imagery is an attractive alternative (Burke et al. 2021), high-resolution satellite imagery does not go back to the early 1990s.

2.2 Identification

To calculate the effects of exposure to pollution on different outcomes, I rely on fixed effects and the differential exposure driven by plausibly exogenous wind direction, conditional on total rainfall and average temperature. The regression is of the form

$$y_{it} = \alpha_i + \gamma_t + \beta wind_{it} + \delta rain_{it} + \psi temp_{it} + \varepsilon_{it}, \quad (1)$$

where y_{it} is outcome y for unit i at time t , α is a geographic fixed effects, γ_t is time fixed effects, $wind_{it}$ is exposure to pollution – defined using wind direction – $rain_{it}$ is total rainfall during the season, and $temp_{it}$ is average temperature during the season. Although I use i and t as subscripts for the fixed effects, note that these change depending on the unit of analysis and the outcome. I explicitly state the level of fixed effects in the tables and results below.

Identification relies on changes in wind direction across time for the same geographic units. In other words, conditional on the fixed effects and weather, I assume changes in wind direction are as good as random and uncorrelated with the outcome – e.g. agricultural productivity – except through exposure to pollution. One major threat to identification is that the building of a coal plant could lead to a spurious correlation between the exposure variable and the outcome. Specifically, since exposure is defined as zero prior to the building of a coal plant, the exposure variable itself might be endogenous even though wind direction is not. To help allay these concerns, I estimate additional regressions restricting the sample to only time periods after which a given location has a coal plant. I also present results using leads to show that future wind direction does not seem to predict current outcomes.

The inclusion of rainfall and temperature is meant to reduce concerns that wind direction may not affect agricultural productivity but, instead, be correlated with it through changes in weather. For example,

Table 1: Local characteristics and the construction of coal plants

	1991 census		2001 census	
	1991	2001	2001	2011
pop (log)	0.068*** (0.000)	0.000 (0.000)	0.009*** (0.000)	0.002*** (0.000)
literacy (prop)	0.127*** (0.002)	-0.005*** (0.001)	0.044*** (0.002)	0.030*** (0.002)
has paved road	-0.037*** (0.001)	0.003*** (0.000)	-0.008*** (0.001)	-0.000 (0.001)
ag productivity			0.033*** (0.001)	0.005*** (0.000)
sub-sample	all	no plant	all	no plant
observations	490,882	447,288	516,798	476,517

Note: Robust standard errors are in parentheses. The outcome in the first column is whether the village is within 30km of a coal plant in 1991. The second column is whether a village in 1991 will have a coal plant in 2001, conditional on not having one in 1991. The last two columns are similarly defined, except using 2001 and 2011 as the years. * p<0.10 ** p<0.05 *** p<0.01

prevailing winds from one direction may be associated with dry weather, while winds from another direction may be more likely to carry moisture and rain. It is not a priori clear how this would bias the estimates with respect to pollution and coal plants, however. The differences in weather patterns is very context specific and can differ even for villages located near one another. Nonetheless, adding rainfall and temperature should alleviate some of these concerns.

A large question looms with respect to identification: what are the characteristics of areas near the construction of new coal plants? While I examine the identifying assumptions in a panel context below, I begin with this static question here. Table 1 presents four separate regressions. The first two columns use 1991 census data – downloaded from SHRUG – and the second two columns use 2001 census data, all at the village level. The first column presents results from regressing a dummy for whether there is a coal plant within 30km in 1991 on several census variables. Villages with higher populations and higher literacy in 1991 are more likely to be near a coal plant at the time. Interestingly, however, they are less likely to have a paved road.

The second column keeps the 1991 census data but changes the outcome variable to a dummy for whether there *will be* a coal plant constructed within 30km before 2001. This column restricts the sample to only villages that do not already have a coal plant within 30km in 1991. Population is no longer a significant predictor while the coefficients on literacy and the presence of a paved road have flipped.

The last two columns are particularly relevant to the current results. They present similar results to those in the first two columns, but use the 2001 census data and a dummy for the presence of a coal plant in 2001 and

Table 2: Wind direction and particulate matter

	1998-2015		2002-2013	
	(1)	(2)	(3)	(4)
wind (10s)	0.002*** (0.000)	0.002*** (0.000)	0.006*** (0.001)	0.001*** (0.000)
fixed effects:				
village	Yes	Yes	Yes	Yes
month	Yes	No	Yes	No
district-month	No	Yes	No	Yes
observations	24,376,458	24,376,458	16,250,973	16,250,973

Note: Standard errors are in parentheses and are clustered at the village level. * p<0.10
** p<0.05 *** p<0.01

2011 in columns three and four, respectively. Importantly, I include agricultural productivity in the monsoon season of 2002 – the earliest year available – as an additional predictor. In 2001, agricultural productivity is a strong predictor for the presence of a coal plant. Moreover, it is also a significant predictor of the construction of a coal plant between 2001 and 2011 (column four). Given that the coefficient is positive, this indicates that coal plants are more likely to be built in areas with higher agricultural productivity. In other words, any selection in the overall location of coal plants would bias against finding negative effects of pollution on agricultural productivity. However, note that this is not the same as looking at wind direction, though comparing the OLS and IV results below leads to the same conclusion regarding selection.

It is worth taking the time to validate the use of wind direction as a measure of exposure to pollution. Consider the data used for the regressions presented in Table 2. The outcome variable is particulate matter – specifically, PM_{2.5}, which is particulate matter no larger than 2.5 micrometers in diameter – which comes from Hammer et al. (2020). PM_{2.5} is one of the harmful byproducts from coal plants, along with sulfur dioxide, different types of nitrogen oxides, and mercury.¹² This data is available at the monthly level, so in Table 2, I aggregate wind exposure to the monthly level, as well.

Both the particulate matter data and the wind data is at the village level, which means the regressions in Table 2 are at the village level. As such, the geographic fixed effects are village fixed effects and the temporal fixed effects are month fixed effects. Since the exposure variable is at the village level and we follow villages over time, the standard errors are also clustered at the village level. In columns two and four, I switch the month fixed effects for district-month fixed effects, which considers only villages in the same district in the same month.

In addition to serving as an example, Table 2 also helps validate the use of wind direction as a proxy for

¹²www.epa.gov/airmarkets/power-plants-and-neighboring-communities

exposure to pollution. The first two columns include all available years of data while the last two columns restrict estimation only to the years for which the agricultural productivity regressions are estimated. Across all four columns, the story is the same: when the wind is blowing in the direction of a village from a coal plant within 30 kilometers, estimated PM_{2.5} is substantially higher. The regressions indicate that one additional day of wind in the direction of the village increases mean PM_{2.5} for the month by between 0.00 and 0.00 $\mu\text{g}/\text{m}^3$.

To put these numbers into perspective, the World Health Organization's updated guidelines are that average annual exposure should not exceed 15 $\mu\text{g}/\text{m}^3$. For the month, an increase at the midpoint of the estimated range is equal to an increase of 0.2 percent of the maximum recommended mean concentration from the WHO. This is just a single day of wind in the direction of a village. Since this includes all villages located within 30km of a coal plant, this is evidence that wind direction can have serious repercussions on the health of hundreds of millions of people in India.

3 Results

This section presents the main results of this paper. I present three sets of results: one set looking at agricultural productivity, another looking at labor allocation, and a final set looking at nightlights, which are a proxy for economic development. I go through each of these in turn.

3.1 Agricultural productivity

I first present results for agricultural productivity. The outcome of all the regressions in this section is the log of agricultural land productivity, which is defined as tons per hectare. The first set of results are in Table 3. The unit of analysis is the vilalge-year-season – from 2002 to 2013 – the geographic fixed effect is the village-season, and the temporal fixed effect is the year. The exposure variable is defined for the entire season. Since Gangopadhyay et al. (2022) define both the monsoon (kharif) season and the winter (rabi) season as five months long, I do as well, using the same five months as the authors. This implies that the exposure variable ranges from zero to slightly more than 150.

The first column presents the most simple results, with only the wind exposure variable and the fixed effects. The second column adds a rainfall – defined as the total rainfall in the season – and average temperature variable, the addition of which is motivated by the fact that changes in wind direction could be correlated with changes in weather. However, the coefficient on the wind exposure variable is completely unchanged by

Table 3: Wind direction and agricultural productivity

	all			monsoon	winter
	(1)	(2)	(3)	(4)	(5)
wind	-0.003*** (0.000)	-0.003*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.004*** (0.000)
rain (z)		0.029*** (0.000)	0.008*** (0.001)	0.075*** (0.002)	0.006*** (0.000)
mean temp (10s)		-0.000*** (0.000)	-0.007*** (0.000)	-0.002*** (0.000)	-0.008*** (0.000)
fixed effects:					
village-season	Yes	Yes	Yes	No	No
year	Yes	Yes	No	Yes	Yes
district-year-season	No	No	Yes	No	No
village	No	No	No	Yes	Yes
observations	2,387,417	2,371,221	2,371,221	1,256,951	1,114,270

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is estimated agricultural yield, defined as tons per hectare. * p<0.10 ** p<0.05 *** p<0.01

the addition of the two weather variables, indicating that weather may not be as worrisome of a confounder as thought. The direction of the weather coefficients is also consistent with expectations; higher rainfall is good for agricultural productivity while higher temperature is not.

Column three adds district-by-year-by-season fixed effects as an additional robustness checks. This addition decreases the magnitude of the exposure variable but it remains negative and significant. However, part of this may be due to the fact that the new fixed effects soak up much of the variation in the exposure variable, as shown in Figure 2. The last two columns separate the effects into the monsoon season and the winter season, but the overall effect is roughly similar, though rainfall has a larger effect in the monsoon season than the winter season, which is unsurprising given the winter season is driven by irrigation in many areas.

To put the size of the coefficient in context, it is worth digging a little deeper into the exposure variable. The mean within-village absolute deviation in wind exposure is approximately 8.06 days, meaning that the average change in agricultural productivity from year-to-year due to nothing but changes in wind patterns carrying particulate matter is around 2.4 percent relative to the mean. This implies that swings in agricultural productivity – due to wind direction – of more than five percent are probably quite common, since the absolute deviation includes deviations below and above the mean.

Of course, it is not wind, *per se*, that we are interested in. Instead, we are actually interested in better understanding the effects of exposure to pollution on agricultural productivity. Table 4 presents the same regressions as Table ??, with a key difference: instead of wind exposure, it uses average concentration of

Table 4: Particulate matter (PM 2.5) and agricultural productivity

	all			monsoon	winter
	(1)	(2)	(3)	(4)	(5)
particulate matter (log PM 2.5)	0.0000 (0.0000)	0.0329*** (0.0028)	-0.0787*** (0.0040)	-0.0509*** (0.0033)	0.0216*** (0.0052)
rain (z)		0.0290*** (0.0004)	0.0079*** (0.0013)	0.0760*** (0.0017)	0.0064*** (0.0005)
mean temp (10s)		-0.0003*** (0.0001)	-0.0071*** (0.0004)	-0.0020*** (0.0001)	-0.0080*** (0.0001)
fixed effects:					
village-season	Yes	Yes	Yes	No	No
year	Yes	Yes	No	Yes	Yes
district-year-season	No	No	Yes	No	No
village	No	No	No	Yes	Yes
observations	2,387,417	2,371,221	2,371,221	1,256,951	1,114,270

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is estimated agricultural yield, defined as tons per hectare. * p<0.10 ** p<0.05 *** p<0.01

PM_{2.5} – from Hammer et al. (2020) – during the season as the key explanatory variable.

Several interesting patterns emerge. We only observe a negative relationship between particulate matter and agricultural productivity in column three; the coefficient in column two is actually positive. Breaking out the effects into the monsoon and winter seasons, we also see a positive relationship between pollution and agricultural productivity.

Recall, however, the results from above: coal plants were placed in places with higher average agricultural productivity. This means that the simple regressions in 4 might actually be biased upwards: although pollution might negatively affect productivity, pollution is higher in more productive areas due to the endogenous placement of coal plants in the study areas. To recover the actual effect of pollution, we need an instrument. If, conditional on the fixed effects and weather variables, wind direction is exogenous and only affects agricultural productivity through pollution – which seems reasonable since it is hard to imagine what other variables it might affect – then it is a valid instrument for pollution.

I present the IV results in Table 5, with the second-stage results in the top panel and the first-stage results in the bottom panel. The first-stage results make clear how strong of an instrument wind is; the t-statistic in column one, for example, is almost 48 while it is still well above 10 in column three, which translates to a first-stage F-statistic of 195, meeting even the higher bar for the first-stage F-test argued for by D. S. Lee et al. (2022). Also note that the coefficients are quite different from those in Table 2 because the unit of analysis in Table 5 is the season, while in Table 2 the unit of analysis is the month.

Table 5: Pollution and agricultural productivity, IV estimates

	all			monsoon	winter
	(1)	(2)	(3)	(4)	(5)
particulate matter (log PM 2.5)	-1.190*** (0.070)	-1.027*** (0.071)	-1.671*** (0.235)	-0.525*** (0.043)	-1.741*** (0.139)
rain (z)		0.028*** (0.000)	0.012*** (0.001)	0.080*** (0.002)	0.001** (0.001)
mean temp (10s)		-0.001*** (0.000)	-0.009*** (0.000)	-0.002*** (0.000)	-0.007*** (0.000)
fixed effects:					
village-season	Yes	Yes	Yes	No	No
year	Yes	Yes	No	Yes	Yes
district-year-season	No	No	Yes	No	No
village	No	No	No	Yes	Yes
observations	2,387,417	2,371,221	2,371,221	1,256,951	1,114,270
first stage:					
wind	0.003*** (0.000)	0.003*** (0.000)	0.000*** (0.000)	0.004*** (0.000)	0.002*** (0.000)
rain (z)		-0.001*** (0.000)	0.002*** (0.000)	0.009*** (0.000)	-0.003*** (0.000)
mean temp (10s)		-0.000*** (0.000)	-0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is estimated agricultural yield, defined as tons per hectare. * p<0.10 ** p<0.05 *** p<0.01

The second-stage results first indicate that PM_{2.5} is a very strong predictor of agricultural productivity. The smallest coefficient – that for the monsoon season in column (4) – and the mean absolute deviation for PM_{2.5} point to a change in agricultural productivity of around 23 percent, which is much higher than the wind results documented above. This makes sense, of course, since the IV is essentially scaling up the effects of wind.

The comparison of the results to those in the previous table also confirm the concerns around endogenous placement of coal plants. The coefficients in Table 5 are always more negative than those in Table 4, and sometimes substantially so. In other words, the simple regression of agricultural productivity on pollution results in coefficients on pollution that are upwards biased, since higher levels of pollution occur in more productive areas, on average. The IV strategy corrects this bias and leads to a robust, negative relationship between pollution and agricultural productivity.

3.2 Heterogeneity

Having established that exposure to pollution has a negative effect on agricultural productivity, I now move to heterogeneity in these effects. Table 6 presents three different analyses: heterogeneity by median exposure, heterogeneity by initial yield, and possible multiplicative effects of shocks. All regressions use the IV strategy detailed above and I include the F statistic from the first stage regressions (of which there are two in the last column) at the bottom of the table. Table A2 in the appendix presents the first-stage results for all columns.

The first two columns split the sample based on median exposure. In other words, for each village, I calculate the median exposure (based on wind) and then use this single value for each village to split the sample into villages above and below median exposure. The coefficient in the first column is 2.5 times larger than the coefficient in the second column, indicating much larger effects of pollution in places where it is more common. One possible explanation is that exposure is cumulative, leading to larger and larger effects. I return to this in column five.

Columns two and three split the sample based on median yield at the very beginning of the sample (the first year available, 2002-2003). Column three presents results for villages above the median while column four presents results for villages below the median. The largest effects are apparently concentrated in the lower half of the distribution. Perhaps more productive areas are more likely to have ways to combat the effects of shocks – e.g. by adjusting the use of other inputs – or maybe less productive areas are inherently more susceptible to the effects of pollution, perhaps due to direct effects on land productivity.

Given ongoing changes in the climate, a key question is how welfare will change as shocks become more and more common. In the last column looks at possible multiplicative effects of shocks by interacting pollution with rainfall and seeing whether the effects compound one another. Since there are now two endogenous regressors, I instrument for both using wind exposure and an interaction between wind exposure and rainfall.

The coefficient on the level term for pollution is negative. Since rainfall is defined as a z score, this represents the effect of pollution only when rainfall is at its mean value for a given season. The interaction term, on the other hand, is positive. Since higher values of rainfall are generally good, this positive coefficient represents a compounding effect of pollution and rainfall; when rainfall is low (bad), the effects of particulate matter on agricultural productivity is larger (more negative). Another way to think about this, however, is that higher values of rainfall dampen the effects of pollution shocks. This could have very profound effects on welfare, since this essentially increases the variance of agricultural productivity across years (which is already increasing due to climate change).

Table 6: Heterogeneity in the effects of pollution on productivity

	exposure		yield		
	>p(50)	<=p(50)	>p(50)	<=p(50)	
PM 2.5	-1.082*** (0.073)	-2.523*** (0.624)	-0.439*** (0.046)	-1.711*** (0.146)	-0.809*** (0.060)
PM time rain					0.180*** (0.008)
weather	Yes	Yes	Yes	Yes	Yes
fixed effects:					
village-season	Yes	Yes	Yes	Yes	Yes
year	Yes	Yes	Yes	Yes	Yes
F (1st stage, PM)	6,488	60	2,867	2,150	2,958
F (1st stage, PM times rain)	—	—	—	—	1,697
observations	1,051,010	1,320,211	1,193,764	1,177,457	2,371,221

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is estimated agricultural yield, defined as tons per hectare. The first two columns split the sample based on the median exposure (by village). The second two columns split the sample based on initial – the first year available – agricultural productivity. * p<0.10 ** p<0.05 *** p<0.01

3.3 Robustness checks

I present two separate robustness checks in the appendix. The first set of analyses attempts to separate the possible endogenous location of coal plants from outcomes. Specifically, prior to the opening of a new coal plant in an area, all villages receive a zero for wind. This means that, even if wind direction is exogenous, the opening of the coal plant could lead to correlations between the error term and the wind variable.

Although the results with district-year-season help mitigate this concern, Table A3 in the appendix checks this possibility in a different way. Specifically, I restrict the sample to only village-years *after* the introduction of a coal plant. This drops any pre-plant-opening observations – all of which are mechanically zeros for wind – to prevent endogeneity due to the opening. The coefficients in Table A3 show that the overall effects are essentially unchanged across any specifications. Comparing changes in coefficients can also help reduce any effects driven by other polluting industries located near coal plants; since the pre-opening exposure variable is zero, the coefficients in Table A3 would likely increase in magnitude if the effects were driven by non-coal pollution.

The second robustness check includes leads and lags for wind to see if they predict current year productivity. I present these results in A4. The first column is the reduced form effect and shows that only current wind predicts productivity; leads and lags have no impact. We see similar results in column two – the IV results – with a negative coefficient for current pollution, only. However, the standard errors are much larger than the main results – around five times as large – making it hard to come to any conclusions.

Table 7: Wind direction and labor allocation by rural status

	all	all	self	wage	farm	non-farm
rural						
wind	-0.112*** (0.036)	-0.118*** (0.029)	-0.050* (0.030)	-0.068** (0.031)	-0.002 (0.028)	-0.116*** (0.034)
DV mean	3.499	3.499	2.172	1.327	1.260	2.240
observations	529,562	529,466	529,466	529,466	529,466	529,466
urban						
wind	-0.020 (0.033)	-0.030 (0.041)	-0.011 (0.039)	-0.019 (0.031)	0.023* (0.014)	-0.053 (0.048)
DV mean	3.040	3.040	1.393	1.647	0.101	2.939
observations	369,483	369,390	369,390	369,390	369,390	369,390
controls	No	Yes	Yes	Yes	Yes	Yes
fixed effects:						
district	Yes	Yes	Yes	Yes	Yes	Yes
state-month	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. The outcome in each regression is days worked in the previous seven days. * p<0.10 ** p<0.05 *** p<0.01

3.4 Labor allocation

The previous set of results show a consistent negative effect of wind-driven pollution on agricultural productivity. A key question is the mechanism through which this happens; is this driven by changes in labor allocation/productivity or by changes in land productivity? While I cannot definitely prove the mechanism, I can shed some light on possibilities by looking at labor allocation decisions. Since these shocks generally happen after initial planting decisions, it seems unlikely that changes in land *allocation* could drive the results.

Table ?? presents results for labor allocation over the last seven days using the National Sample Survey and the methodology described above. Since the labor recall is for seven days, labor allocation ranges from zero to seven, as does the wind variable. Recall that the data is at the district level, not the village level. However, it is also more temporally disaggregated, with rolling interviews throughout the year-long survey. Figure A1 in the appendix shows this temporal distribution across the five waves of the survey. Since I can no longer include district-by-month fixed effects, I instead opt for district-level trends to help take into account aggregate temporal changes in outcomes.

There are some interesting results in Table ???. The first two columns show a marginally significant decrease in overall labor allocation (where “labor” is defined as non-domestic self and wage employment). The third and fourth columns show that this overall decrease is predominantly from wage employment, not self employment. An additional day of wind over the last seven decreases wage employment by 0.034 days in the district. Mean

wage employment is 1.458 days, meaning one day of wind decreases wage employment by 2.3 percent of the mean. Since much of the district is not covered by “treated” villages, this is actually a rather substantive effect on the villages involved.

The last two columns split the effects by sector, into farm and non-farm employment. Although agricultural (land) productivity decreases when wind blows in the direction of villages, overall labor allocation to farm labor actually goes *up*, by around 4.5 percent of the mean of 0.784 days. Non-farm labor, on the other hand, goes down, by around 2.4 percent of the mean of 2.527 days.

There are two possible explanations for this pattern of results. First, it could be that land productivity goes down, leading farmers to increase their farm labor allocation in an attempt to compensate for the decreased productivity. A second possibility is that pollution does not affect land productivity at all, at least not directly. In this scenario, the worsening air quality could lead to decreased labor productivity, meaning that “effective labor” essentially decreases when pollution is higher.

Table A5 and Table A6 in the appendix disaggregate these effects into the monsoon and winter seasons. Apparently, the entirety of the effect is driven by changes in labor allocation during the monsoon season. The dynamics across seasons apparently lead to large differences in the effects of pollution on labor allocation.

One way to provide suggestive evidence of mechanisms is to explore heterogeneity in the effects. Previous research has shown larger effects of pollution on the health of the oldest (Deryugina et al. 2019). The labor sample includes all adults between 15 and 60 years of age. To look at heterogeneity by age, I present an additional set of results interacting the wind variable with an indicator for whether an individual is over 37 years of age.¹³ These results are in Table ??, which otherwise presents the same specifications in the last five columns as ??.

Table ?? presents a very clear set of results: the effects are always larger (more negative) for older adults. Particularly interestingly, we now see that the mean-zero effect on self employment in the previous table is driven by two stark contrasts: a large increase for the younger group and a large decrease for the older group. For the youngest group, this increase appears in farm labor, indicating likely increases in farm self-employment. The oldest, on the other hand, have large decreases in both farm and non-farm labor, with larger decreases in the latter.

To put some of these changes in context, an additional day of exposure to pollution from wind increases self-employment for the younger half of the sample by around six percent of the mean and increases farm employment for the younger half by almost 10 percent of the mean. For the older half, on the other hand, an

¹³This splits the sample into two equal bins of age: those who are 15-37 and those who are 38-60.

Table 8: Wind direction and labor allocation by age (rural only)

	all	all	self	wage	farm	non-farm
younger than 35						
wind	-0.077*	-0.085**	-0.061*	-0.024	-0.001	-0.085**
	(0.046)	(0.034)	(0.032)	(0.032)	(0.034)	(0.038)
DV mean	3.106	3.106	1.794	1.312	1.061	2.044
observations	263,965	263,923	263,923	263,923	263,923	263,923
45 or older						
wind	-0.131**	-0.165***	-0.054	-0.111***	-0.070*	-0.095**
	(0.054)	(0.043)	(0.054)	(0.041)	(0.036)	(0.042)
DV mean	3.584	3.584	2.481	1.103	1.444	2.140
observations	166,184	166,155	166,155	166,155	166,155	166,155
controls	No	Yes	Yes	Yes	Yes	Yes
fixed effects:						
district	Yes	Yes	Yes	Yes	Yes	Yes
state-month	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. The outcome in each regression is days worked in the previous seven days. * p<0.10 ** p<0.05 *** p<0.01

additional day of exposure decreases self-employment and non-farm employment by around five percent of their means.

While speculative, one possible explanation for this pattern is a decrease in labor productivity. While the youngest may be able to readjust and increase their overall self-employment labor allocation, perhaps the oldest cannot due to, for example, health effects of pollution, which are well documented (Bentayeb et al. 2012; Pope III 2000).¹⁴ We see decreases in overall agricultural productivity, despite an increase in overall farm labor. We also see decreases in non-farm employment, indicating that individuals are reallocating labor from the non-farm sector to the farm sector.

It is worth noting that the larger negative effects for older adults is more consistent with a labor story than a land story; it is hard to imagine these patterns emerging if the only effects were through land productivity changes. The overall decrease in non-farm employment is more compatible with a labor health/productivity story, though with general equilibrium effects in agriculture spillover over into the non-farm sector, this need not necessarily be the case.

In addition, Table A7 in the appendix presents one additional set of results to support a labor story. Changes in land productivity should be relatively fixed in a short period of time. For example, if we focus on just a single month in a single district, the overall effects driven by changes in land productivity should be

¹⁴Much of the literature also suggests the youngest are at significant risk of mortality and morbidities. However, given my focus on labor allocation, I only discuss older adults here.

constant even though pollution may have differential effects throughout the month on labor productivity, which responds more immediately to changes in air quality. Table A7 changes the specification to include district-by-year-by-month fixed effects. In other words, the results compare people in the same district in the same month, but who may have differential exposure to pollution simply because they were interviewed at a slightly different time.

The overall pattern of results stays the same when we shift the focus. Overall labor allocation increases for the youngest, but decreases for the oldest, and this is driven completely by self employment. There are decreases in both farm and non-farm employment for the oldest, while the youngest show increases in non-farm employment, but not farm employment – which is a change from the previous results. The overall story remains the same, however, which is further evidence of effects being driven primarily through changes in labor allocation and productivity.

Finally, Table A8 in the appendix shows that there are no significant changes in daily wages, though the overall direction of wage changes would be ambiguous, anyway. While reduced productivity might lead to lower wages, decreased labor supply could also put upward pressure on wages.

3.5 Month-specific exposure

Pollution leads to lower agricultural productivity as well as changes in labor allocation. This leads to an obvious question: might higher levels of pollution also directly affect economic growth? To explore this question, I look at changes in nightlights relative to changes in pollution. It seems reasonable to think that pollution might affect nightlights with a lag, so I include the lag of pollution rather than contemporaneous pollution.

I present these results in Table ???. To help simplify the coefficients, I recode wind to be the proportion of days in the year that a given village experiences wind in their direction, from a coal plant. Mean wind in this new specification is 0.066, indicating that a village in the sample experiences wind in its direction on around 6.6 percent of days, or 24 days of the year.

We see a strong, negative relationship between pollution (lagged wind) and nightlights. Going from the 25th percentile to the 75th percentile of wind exposure is synonymous with a decrease in the growth of nightlights within a village by around 0.7 percent. While this is a relatively small magnitude, it is worth noting that there are more than 100,000 villages in this sample. While the effect for a single village might be small, the overall, aggregate effect on growth is quite substantial. To put this number in perspective, Castelló-Climent, Chaudhary, and Mukhopadhyay (2018) argue that a standard-deviation increase in the

Table 9: Month-specific exposure

	monsoon	winter
Month 1 (June/Nov.)	-0.808*** (0.066)	0.788 (0.588)
Month 1 (July/Dec.)	0.049 (0.060)	-0.351 (0.967)
Month 3 (Aug./Jan.)	-0.229*** (0.034)	4.758 (3.756)
Month 4 (Sept./Feb.)	-0.215*** (0.035)	-2.086*** (0.503)
Month 5 (Oct./March)	0.149* (0.082)	-3.788 (3.210)
fixed effects:		
village	Yes	Yes
year	Yes	Yes
F (month 1)	1,430	566
F (month 2)	1,029	897
F (month 3)	2,713	358
F (month 4)	2,131	389
F (month 5)	592	565
observations	1,266,922	1,120,495

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is logged nightlights. * p<0.10 ** p<0.05 *** p<0.01

share of adults with a university education leads to 12 percent higher nightlights in India, while Storeygard (2016) shows that a four-fold increase in oil prices over six years leads to seven percent lower nightlights at a distance of around 500km from the nearest large port in sub-Saharan Africa. In other words, the 0.7 percent change in nightlights is quite large given this is annual variation.

4 Conclusion

The rapid growth of coal power plants has also led to increases in emissions from the burning of coal. In this paper, I show some of the negative effects associated with this burning: decreased agricultural productivity. I find clear indications that pollution from coal plants leads to lower agricultural productivity. Importantly, I use exogenous changes in wind direction to measure exposure to pollution and show that these effects are robust to a range of alternative specifications.

Results indicate that these effects may be driven by changes in labor allocation patterns. In addition, the negative labor effects are concentrated in older adults, which is consistent with previous research on the effects of pollution on health in adults. Despite this, I am unable to rule out direct effects on land productivity with

the data available.

Finally, these negative effects are also seen in aggregate economic development, with pollution leading to decreases in the growth of nightlights. These effects, though small for an individual village, are quite large in overall terms, indicating that pollution from coal plants can have effects on overall economic development.

These results, of course, do not address possible positive effects of the building of additional power plants, which can lead to increased electricity access. As such, the main conclusion is that the location of new power generation facilities may be particularly important for multiple aspects of economic activity, not just health and mortality.

References

- Abay, Kibrom A, Christopher B Barrett, Talip Kilic, Heather Moylan, John Ilukor, and Wilbert Drazi Vundru. 2022. "Nonclassical Measurement Error and Farmers? Response to Information Reveal Behavioral Anomalies."
- Aragón, Fernando M, and Juan Pablo Rud. 2016. "Polluting industries and agricultural productivity: Evidence from mining in Ghana." *The Economic Journal* 126 (597): 1980–2011.
- Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City." *The Economic Journal* 126 (591): 257–80.
- Asher, Sam, Tobias Lunt, Ryu Matsuura, and Paul Novosad. 2021. "Development Research at High Geographic Resolution: An Analysis of Night Lights, Firms, and Poverty in India using the SHRUG Open Data Platform." *The World Bank Economic Review*.
- Bencko, Vladimir, and Karel Symon. 1977. "Health Aspects of Burning Coal with a High Arsenic Content." *Environmental Research* 13 (3): 378–85.
- Bentayeb, Malek, Marzia Simoni, Nour Baiz, Dan Norback, Sandra Baldacci, Sara Maio, Giovanni Viegi, Isabella Annesi-Maesano, Geriatric Study in Europe on Health Effects of Air Quality in Nursing Homes (GERIE) Group, et al. 2012. "Adverse Respiratory Effects of Outdoor Air Pollution in the Elderly." *The International Journal of Tuberculosis and Lung Disease* 16 (9): 1149–61.
- Brunekreef, Bert, and Stephen T Holgate. 2002. "Air pollution and health." *The Lancet* 360 (9341): 1233–42.
- Burke, Marshall, Anne Driscoll, David B Lobell, and Stefano Ermon. 2021. "Using satellite imagery to understand and promote sustainable development." *Science* 371 (6535): eabe8628.
- Burney, Jennifer, and V Ramanathan. 2014. "Recent climate and air pollution impacts on Indian agriculture." *Proceedings of the National Academy of Sciences* 111 (46): 16319–24.
- Castelló-Climent, Amparo, Latika Chaudhary, and Abhiroop Mukhopadhyay. 2018. "Higher education and prosperity: From Catholic missionaries to luminosity in India." *The Economic Journal* 128 (616): 3039–75.
- 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.
- Chen, Shuai, Paulina Oliva, and Peng Zhang. 2022. "The effect of air pollution on migration: evidence from China." *Journal of Development Economics* 156: 102833.
- Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. 2014. "What do we know about short-and long-term effects of early-life exposure to pollution?" *Annual Review Resource Economics* 6 (1): 217–47.
- 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.
- Desiere, Sam, and Dean Jolliffe. 2018. "Land productivity and plot size: Is measurement error driving the inverse relationship?" *Journal of Development Economics* 130: 84–98.
- Dinkelmann, Taryn. 2011. "The effects of rural electrification on employment: New evidence from South Africa." *American Economic Review* 101 (7): 3078–108.
- Ebenstein, Avraham, Victor Lavy, and Sefi Roth. 2016. "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution." *American Economic Journal: Applied Economics* 8 (4): 36–65.
- Gangopadhyay, Prasun K, Paresh B Shirsat, Vinay K Dadhwal, and Pramod K Aggarwal. 2022. "A new two-decade (2001–2019) high-resolution agricultural primary productivity dataset for India." *Scientific Data* 9 (1): 1–12.
- Gibson, John, Susan Olivia, Geua Boe-Gibson, and Chao Li. 2021. "Which night lights data should we use in economics, and where?" *Journal of Development Economics* 149: 102602.
- Graff Zivin, Joshua, and Matthew Neidell. 2012. "The impact of pollution on worker productivity." *American Economic Review* 102 (7): 3652–73.
- 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.

- Hanna, Rema, and Paulina Oliva. 2015. "The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City." *Journal of Public Economics* 122: 68–79.
- He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. "Severe air pollution and labor productivity: Evidence from industrial towns in China." *American Economic Journal: Applied Economics* 11 (1): 173–201.
- Heck, Walter W, OC Taylor, Richard Adams, Gail Bingham, Joseph Miller, Eric Preston, and Leonard Weinstein. 1982. "Assessment of crop loss from ozone." *Journal of the Air Pollution Control Association* 32 (4): 353–61.
- Heft-Neal, Sam, Jennifer Burney, Eran Bendavid, and Marshall Burke. 2018. "Robust relationship between air quality and infant mortality in Africa." *Nature* 559 (7713): 254–58.
- Heft-Neal, Sam, Jennifer Burney, Eran Bendavid, Kara Voss, and Marshall Burke. 2019. "Air pollution and infant mortality: Evidence from Saharan dust." National Bureau of Economic Research.
- Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028.
- IEA. 2022. "Coal-fired electricity." International Energy Association.
- Kampa, Marilena, and Elias Castanas. 2008. "Human health effects of air pollution." *Environmental Pollution* 151 (2): 362–67.
- Kline, Patrick, and Enrico Moretti. 2014. "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics* 129 (1): 275–331.
- Kopas, Jacob, Erin York, Xiaomeng Jin, SP Harish, Ryan Kennedy, Shiran Victoria Shen, and Johannes Urpelainen. 2020. "Environmental justice in India: incidence of air pollution from coal-fired power plants." *Ecological Economics* 176: 106711.
- Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2022. "Valid t-ratio Inference for IV." *American Economic Review* 112 (10): 3260–90.
- Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2020. "Does Household Electrification Supercharge Economic Development?" *Journal of Economic Perspectives* 34 (1): 122–44.
- Li, Ya-Ru, and Jacqueline MacDonald Gibson. 2014. "Health and air quality benefits of policies to reduce coal-fired power plant emissions: a case study in North Carolina." *Environmental Science & Technology* 48 (17): 10019–27.
- Marshall, Fiona, Mike Ashmore, Fiona Hinchcliffe, et al. 1997. *A hidden threat to food production: Air pollution and agriculture in the developing world*. International Institute for Environment and Development.
- Pope III, C Arden. 2000. "Epidemiology of Fine Particulate Air Pollution and Human Health: Biologic Mechanisms and Who's at Risk?" *Environmental Health Perspectives* 108 (suppl 4): 713–23.
- Pope III, C Arden, and Douglas W Dockery. 2006. "Health effects of fine particulate air pollution: lines that connect." *Journal of the Air & Waste Management Association* 56 (6): 709–42.
- Ramanathan, Veerabhadran, and Gregory Carmichael. 2008. "Global and regional climate changes due to black carbon." *Nature Geoscience* 1 (4): 221–27.
- Reddy, V Ratna, and Bhagirath Behera. 2006. "Impact of water pollution on rural communities: An economic analysis." *Ecological Economics* 58 (3): 520–37.
- Rud, Juan Pablo. 2012. "Electricity provision and industrial development: Evidence from India." *Journal of Development Economics* 97 (2): 352–67.
- Storeygard, Adam. 2016. "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa." *The Review of Economic Studies* 83 (3): 1263–95.
- Van de Walle, Dominique, Martin Ravallion, Vibhuti Mendiratta, and Gayatri Koolwal. 2017. "Long-term gains from electrification in rural India." *The World Bank Economic Review* 31 (2): 385–411.
- Wen, Jeff, and Marshall Burke. 2022. "Lower test scores from wildfire smoke exposure." *Nature Sustainability* 5 (11): 947–55.

Appendix A

Table A1: Data sources

	source	geographic coverage	temporal coverage
shapefile	Asher et al. (2021)	India	
coal plants	Global Energy Monitor	global	yearly
wind	NCAR	global	daily
pollution	Hammer et al. (2020)	global	monthly
agriculture	Angopadhyay et al. (2022)	India	two seasons/year
nightlights	Asher et al. (2021)	India	yearly
weather	TerraClimate	global	monthly
labor	National Sample Survey (NSS)	India	daily interviews

Global Energy Monitor: globalenergymonitor.org/projects/global-coal-plant-tracker. NCAR: climatedataguide.ucar.edu/. TerraClimate: www.climatologylab.org/terraclimate.html

Table A2: Heterogeneity in the effects of pollution on productivity, first-stage results

	(1)	(2)	(3)	(4)	(5)	
	PM	PM	PM	PM	PM	PM*rain
wind	0.003*** (0.000)	0.000*** (0.000)	0.003*** (0.000)	0.002*** (0.000)	0.001*** (0.000)	-0.032*** (0.001)
wind times rain					0.001*** (0.000)	0.009*** (0.000)
weather	Yes	Yes	Yes	Yes	Yes	Yes
fixed effects:						
village-season	Yes	Yes	Yes	Yes	Yes	Yes
year	Yes	Yes	Yes	Yes	Yes	Yes
observations	1,051,010	1,320,211	1,193,764	1,177,457	2,371,221	2,371,221

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions is estimated agricultural yield, defined as tons per hectare. The first two columns split the sample based on the median exposure (by village). The second two columns split the sample based on initial – the first year available – agricultural productivity. * p<0.10 ** p<0.05 *** p<0.01

Table A3: Wind direction and agricultural productivity, post-coal plant only

	all			monsoon	winter
	(1)	(2)	(3)	(4)	(5)
particulate matter	-1.153*** (0.079)	-1.028*** (0.083)	-2.270*** (0.386)	-0.538*** (0.044)	-1.688*** (0.166)
(log PM 2.5)					
rain (z)		0.017*** (0.001)	0.018*** (0.002)	0.073*** (0.002)	-0.002*** (0.001)
mean temp (10s)		-0.000*** (0.000)	-0.008*** (0.001)	0.005*** (0.000)	-0.007*** (0.000)
fixed effects:					
village-season	Yes	Yes	Yes	No	No
year	Yes	Yes	No	Yes	Yes
district-year-season	No	No	Yes	No	No
village	No	No	No	Yes	Yes
observations	1,886,956	1,872,177	1,872,177	996,509	875,668
first stage:					
wind	0.002*** (0.000)	0.002*** (0.000)	0.000*** (0.000)	0.004*** (0.000)	0.002*** (0.000)
rain (z)		-0.002*** (0.000)	0.004*** (0.000)	0.023*** (0.000)	-0.003*** (0.000)
mean temp (10s)		-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	0.000*** (0.000)

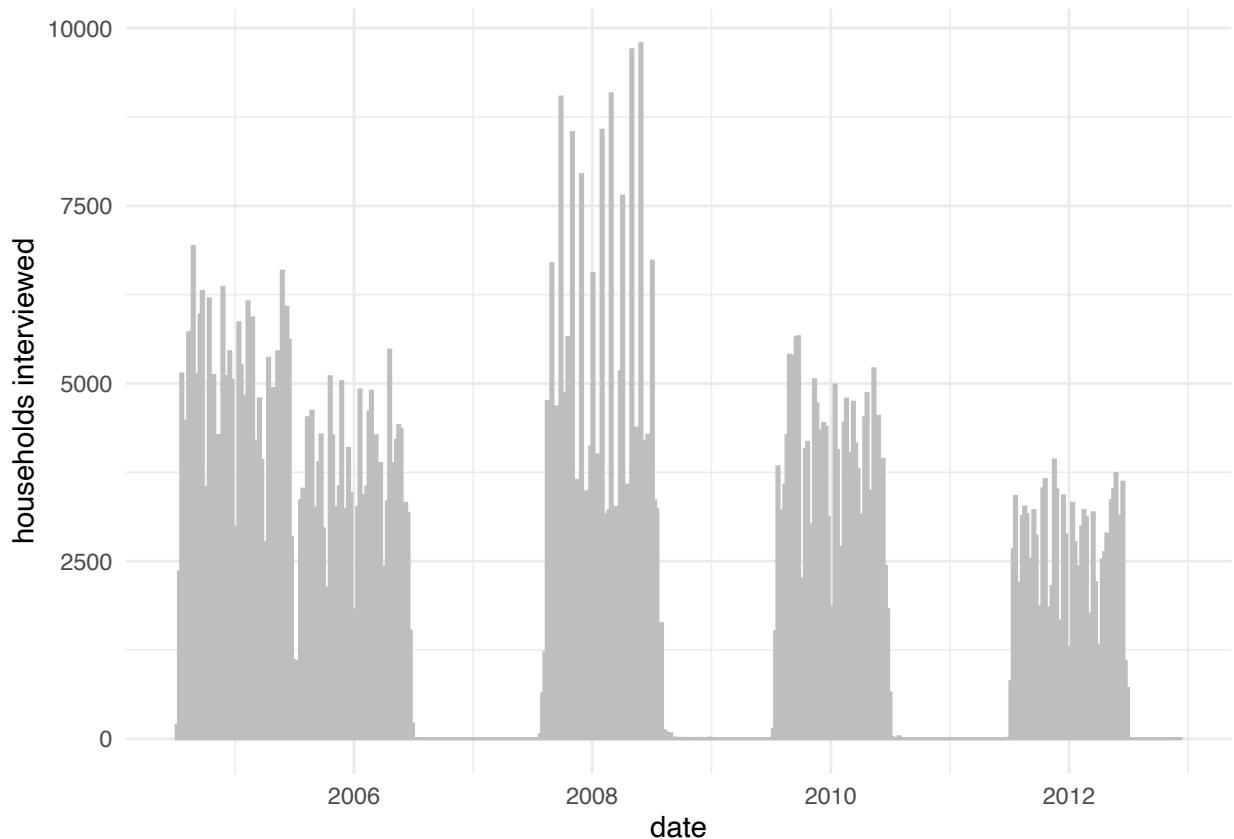
Note: Standard errors are in parentheses and are clustered at the village level. The sample is restricted to village-years after a coal plant was built near the village. The outcome in all regressions estimated agricultural yield (tons per hectare). * p<0.10 ** p<0.05 *** p<0.01

Table A4: Agricultural productivity and pollution leads

	(1)	(2)
wind (lag)	0.0005 (0.0003)	
wind	-0.0006*** (0.0002)	
wind (lead)	-0.0001 (0.0001)	
PM (lag)		0.4340 (0.8666)
PM		-0.6942*** (0.1497)
PM (lead)		-0.0958 (0.5137)
weather	Yes	Yes
fixed effects:		
village-season	Yes	Yes
district-year-season	Yes	Yes
F (lag)	—	162
F	—	410
F (lead)	—	289
observations	1,976,205	1,976,205

Note: Standard errors are in parentheses and are clustered at the village level. The outcome in all regressions estimated agricultural yield (tons per hectare). * p<0.10 ** p<0.05 *** p<0.01

Figure A1: Timing of household surveys for the National Sample Survey



Note: The histogram shows the distribution of survey dates for the four waves of the National Sample Survey (NSS) for the five waves used in this paper.

Table A5: Wind direction and labor allocation, monsoon season only

	all	all	self	wage	farm	non-farm
wind	-0.120*** (0.046)	-0.119** (0.046)	-0.067 (0.042)	-0.052 (0.045)	0.025 (0.036)	-0.144** (0.062)
controls	No	Yes	Yes	Yes	Yes	Yes
fixed effects:						
district	Yes	Yes	Yes	Yes	Yes	Yes
state-month	Yes	Yes	Yes	Yes	Yes	Yes
DV mean	3.324	3.324	1.883	1.441	0.810	2.513
observations	359,652	359,576	359,576	359,576	359,576	359,576

Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. * p<0.10 ** p<0.05 *** p<0.01

Table A6: Wind direction and labor allocation, winter season only

	all	all	self	wage	farm	non-farm
wind	-0.012 (0.040)	-0.032 (0.037)	-0.026 (0.040)	-0.006 (0.028)	-0.020 (0.030)	-0.012 (0.036)
controls	No	Yes	Yes	Yes	Yes	Yes
fixed effects:						
district	Yes	Yes	Yes	Yes	Yes	Yes
state-month	Yes	Yes	Yes	Yes	Yes	Yes
DV mean	3.339	3.339	1.865	1.474	0.790	2.550
observations	375,956	375,883	375,883	375,883	375,883	375,883

Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. The outcome in each regression is days worked in the previous seven days. * p<0.10 ** p<0.05 *** p<0.01

Table A7: Wind direction and labor allocation, heterogeneity by age

	all	self	wage	farm	non-farm
wind	0.100*** (0.027)	0.117*** (0.029)	-0.018 (0.026)	0.026 (0.024)	0.074*** (0.028)
wind times age ≥ 38	-0.203*** (0.037)	-0.201*** (0.032)	-0.002 (0.039)	-0.070*** (0.018)	-0.133*** (0.031)
controls	Yes	Yes	Yes	Yes	Yes
fixed effects:					
district-year-month	Yes	Yes	Yes	Yes	Yes
observations	898,856	898,856	898,856	898,856	898,856

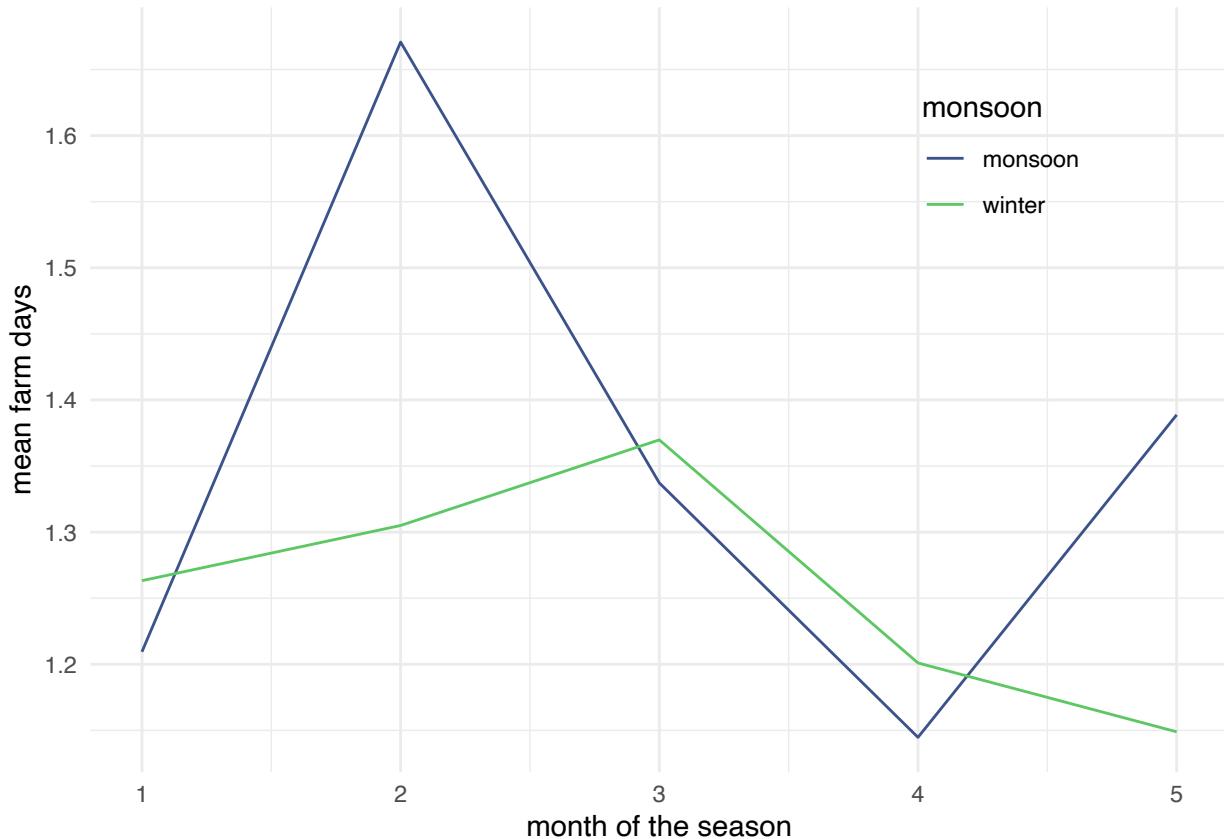
Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. The outcome in each regression is days worked in the previous seven days. * p<0.10 ** p<0.05 *** p<0.01

Table A8: Wind direction and wages

	farm	non-farm	all
wind	0.037 (0.031)	0.010 (0.010)	0.009 (0.011)
controls	Yes	Yes	Yes
fixed effects:			
district	Yes	Yes	Yes
year	Yes	Yes	Yes
varying slopes:			
year (by district)	Yes	Yes	Yes
observations	16,534	149,973	133,526

Note: Standard errors are in parentheses and are clustered at the district level. Control variables include female, age, age squared, and (years of) education. The outcome in each regression is the daily wage (log Rs per day). * p<0.10 ** p<0.05 *** p<0.01

Figure A2: Average farm days by season, rural only



Note: The line graph shows the mean number of farm days worked per week based on the month of the relevant agricultural season. The monsoon season is defined as June (month 1) to October (month 5) while the rainy season is defined as November (month 1) to March (month 5) of the following year. The data is from the NSS.