Three Essays in Environmental and Development Economics

Dissertation Advisors:
Professor Rema Hanna
Professor Robert N. Stavins

Author:
Arnold Patrick Behrer

Three Essays in Environmental and Development Economics

Abstract

In these three essays I examine the relationship between environmental quality and economic development in a variety of settings. In Chapter 1, I examine the impact of the world's largest anti-poverty program (NREGA) on agricultural burning and its subsequent contribution to air pollution in India. I find that agricultural burning increases substantially after the implementation of NREGA. I find evidence that this is due to farmers mechanizing part of the production process in response to higher wages induced by NREGA. The increase in agricultural burning leads to a substantial increase in the emissions of particulate pollutants from biomass burning. In Chapter 2, I and a co-author examine how chronic exposure to particulate air pollution in the United States may worsen the mortality impacts of a pandemic. Using an instrumental variables approach based on shifts in electric power generating capacity due to the shock of hydraulic fracturing, we show that mortality from COVID19 increases in counties that have experienced higher levels of $PM_{2.5}$ pollution in the ten years prior to the pandemic. In Chapter 3, co-authors and I examine how exposure to high temperatures during schooling may reduce student learning. Using data from both the United States and 58 other countries we find that hotter days in the year(s) leading up to an exam substantially reduce performance on that exam. The impact of hot days is concentrated on school days as compared to weekends or summer days. The impact of heat is larger in low income districts of the United States and low income countries. We discuss the implications that this may have for climate change and long-run economic growth.

Contents

List of Tables

List of Figures

Acknowledgments

Completing a Ph.D. dissertation is a long process that no one does alone. I am deeply indebted to many people for the help and support they have provided throughout this journey.

First, my advisors, Rema Hanna, Robert Stavins, Edward Glaeser, and Gabriel Kriendler.

Rob's example and guidance throughout the Ph.D. shaped not only this dissertation but also set an example that will guide the rest of my career. Rema and Ed joined my committee later but made an immediate and lasting impact. They taught me to be a sharper, more careful, and much better researcher. Finally, Gabes generosity with his time, and willingness to talk through any question, helped me get through some of the most difficult parts of this process. Joe Aldy and Andrei Shleifer also gave me generous feedback and help despite not being on my committee. Megan Mauter encouraged me to pursue the Ph.D. and offered advice freely at every stage.

My classmates were essential sources of advice, guidance, and support throughout this process. In particular Jisung Park, Valentin Bolotnyy, and Sam Stolper all offered wisdom gained from going through the market before me and made my path easier. Sharan Mamidipudi and Asad Liaqat went through it with me and reminded me we'd make it. MD Magini shared both his office and his advice whenever I was stuck on a technical problem. I cannot name everyone in the Harvard Environmental Economics Lunch and Ph.D. community individually but all were valuable sources of feedback and support throughout my Ph.D.

Nicole Tateosian kept me on track and was always there to help make the day-to-day of being a Ph.D. student easier. We are lucky to have her helping us in this process.

Simon Schröder gave excellent research assistance and both Rakesh Kumar and Ryan Lee provided invaluable technical assistance collecting data. That effort was aided by support from the Harvard Environmental Economics Program. I also gratefully acknowledge funding support from the EPA STAR Fellowship, the Harvard Climate Change Solutions Fund, and Harvards Vicki Norberg-Bohm Fellowship.

Finally, I cannot thank my family enough for their support. Devon, mom, and dad were consistent and vocal in their belief that I would succeed in finishing the Ph.D. Curiosity and persistence are the most important traits in research; I get both from my parents. They, and Isbell Behrer, also inspired my love for nature. Christopher has always been there to talk and argue about ideas and, despite being younger, has always pushed me to be better. This was no exception.

And Katie, thank you for everything.

To mom and dad who made me believe this, and anything, was possible.

Introduction

Environmental protection or economic development? This is the trade-off that policy-makers frequently face. Economists have worked to add nuance to this simple trade-off for many years. In particular, they highlight the ways that environmental quality can impact economic development and how economic development can occur in ways that do not require as severe of a trade-off.

Contributing to this effort is the focus of my doctoral dissertation. In the three chapters that follow, I examine three different settings in which development policy has impacted environmental quality in unexpected ways or in which environmental quality has had important impacts on human development and well-being.

In the first chapter I examine the unintended consequences of a development policy in India on environmental quality. India's Mahatma Gandhi National Rural Employment Guarantee Act (NREGA) is the worlds largest anti-poverty program. It was intended to offer an avenue for economic development and poverty reduction in rural Indian communities by providing a Federal jobs guarantee to residents of rural districts. NREGA offered workers jobs on local public works projects financed by the Federal government. The program succeeded in providing employment to rural workers and in raising rural wages. However, it had the unintended consequence of increasing the number of fires to clear post-harvest residue on rural farms.

Using a difference-in-differences framework that takes advantage of the sequential implementation of NREGA across Indian districts over three years, I show that fires on agricultural land increased by 21% after the implementation of NREGA in a district. This increase in fires occurs only among fires that occur on agricultural land; there is no increase in fires on plantations, scrubland, or in forests.

It appears that the use of fires increased because farmers responded to the higher wages that resulted from NREGA implementation by harvesting their crops with mechanical combines instead of manual labor. Mechanical combines leave more residue on the fields after the harvest which must be removed before planting the next seasons crops. Burning is the least expensive way for farmers to remove this residue. I find that burning increased more, after the implementation of NREGA, in areas that had more ability to mechanize and more experience with burning prior to NREGA.

The increase in agricultural fires after NREGA leads to an increase in the emission of particulate pollution. Using the same difference-in-differences framework, I find that the number of months in which the Indian Ambient $PM_{2.5}$ standard was exceeded increases by 11% after the implementation of NREGA. Crop burning is a substantial source of particulate pollution in urban India during the burning season and the increase in pollution from NREGA added to pollution levels that were already among the highest in the world.

In Chapter 2, I continue my focus on air pollution and examine how chronic exposure to high levels of particulate air pollution in the United States increased mortality from the 2020 COVID19 pandemic. COVID19 is a respiratory disease whose transmission and lethality may both be increased by particulate air pollution. Specifically, chronic exposure to high levels of particulate air pollution may increase the ability of the virus to enter the cells of the body and infect an individual. Chronic exposure to particulate air pollution may also increase the chance that patients who are infected suffer a hyperinflammatory response to the disease, a key cause of mortality from COVID19.

In order to identify the impact of air pollution on COVID19 mortality specifically as opposed to other determinants of mortality that may be correlated with air pollution I use an instrumental variables approach that relies on the shift away from coal fired power generation that occurred in the United States in response to hydraulic fracturing. This approach relies on the shift from relatively dirty coal-fired electricity generation to relatively clean natural gas generation due to changes in the natural gas price that occurred after the hydraulic fracturing revolution. The shift away from coal generation provides an exogenous change in pollution levels in downwind counties.

Using this change, I find that a 1% increase in pollution levels increases COVID19 mortality by approximately 0.75%. This is impact is consistent with existing work that finds substantial mortality impacts of increased air pollution in both direction and magnitude. It suggests an additional consequence of elevated air pollution increased mortality during pandemics that regulators and policy-makers may wish to consider in choosing the appropriate level of particulate pollution to tolerate.

Finally, in Chapter 3, I shift my attention to the impacts of climate change and specifically to the impact of exposure to high temperatures during the learning process. Heat negatively impacts cognitive performance across a range of environments. We know from previous research that students who take exams on hot days perform worse on those exams relative to their peers who take exams on cooler days. However, we know little about how heat impacts what students learn on a normal school day and how much of that they retain. We know even less about the impacts of heat outside of the United States.

To shed light on that question I use data from two data sets combined with natural variation in temperature across countries during the school year to show that students appear to learn substantially less during hotter years. I combine data on nearly all U.S. students from 2009 to 2015 with data from the PISA international assessment on student performance on a standardized exam across 58 countries. I show that in the United States student performance declines by 0.04% of a standard deviation for every school day above 80°F in the year leading up to an exam. In the international data an additional 80°F day reduces performance by 0.22% of a standard deviation.

These declines are much larger in lower income school districts in the United States and in countries in the bottom half of the income distribution in the international sample. Further, minority students in the United States appear to suffer larger negative consequences from heat exposure relative to white students. Internationally, I examine whether the decline in performance can be explained by correlated shocks to agriculture and I find no evidence that declines in nutrition can explain the decline in student performance.

These results suggest that heat can have a long-term impact on economic growth by reducing the rate at which human capital is created. Much more research is needed to clearly establish that link and measure its magnitude. These results indicate a clear additional cost of climate change however and one that may need to be included in calculations of the social cost of carbon.

Chapter 1

Earth, Wind, and Fire: The Impact of Anti-Poverty Efforts on Indian Agriculture and Air Pollution

1.1 Introduction

What drives the relationship between income growth and environmental quality? Economists have observed a correlation between income growth and reductions in environmental quality since at least the early 1990s (Grossman and Krueger, 1995) but the mechanisms driving that relationship remain unresolved (Harbaugh *et al.*, 2002). One possibility is that rising incomes and wages lead to a transition from labor intensive production to capital and pollution intensive production (Arrow *et al.*, 1995; Kuznets, 1973). The transition from human powered cotton production to steam powered production in England is a classic example: as firms replaced human labor with coal-fired powered cotton, production increased and shifted to Manchester which suffered notable declines in air quality (Longhurst and Conlan, 1970; Rodgers, 1960).[1] In London the replacement of human labor with power from coal

[1] Longhurst and Conlan (1970) quotes the Manchester police commission in the 1800s: "the increase of steam engines as well as smoak issuing from chimnies used over stoves, foundries, dressers, dyehouses and bakehouses are become a great nuisance to the town." There is no support for the claim that the downward sloping portion of the "environmental kuznets curve" reflects a causal relationship (Harbaugh *et al.*, 2002), but the correlation between increasing environmental damage and income growth over some range of starting income has robust empirical support (Stern, 2018; Dinda, 2004; Dasgupta

throughout the 19^{th} century increased pollution and led to the appearance of the infamous "London Fog" (Clay and Troesken, 2010).

A central challenge in demonstrating that increasing incomes has a causal impact on pollution by increasing mechanization has been the simultaneity of income growth and mechanization (Ebenstein *et al.*, 2015). As I've suggested, and show in a theoretical exercise in Section 2, rising incomes could drive increases in mechanization by raising labor wages and leading firms to invest in labor-saving mechanization.[2] Allen (2011) argues that firms choice to invest in labor-saving production techniques in response to higher relative wages in England is the central reason the Industrial Revolution began there. However, exogenous innovation in production techniques also leads to labor-saving mechanization that raises labor productivity with wage increases coming as a consequence of higher productivity (Solow, 1957).[3]

I show empirically that a policy raising incomes led to an increase in pollution. I address the problem of simultaneity by measuring how air pollution changes after an exogenous shock to wages generated by India's Mahatma Gandhi National Rural Employment Guarantee Act (NREGA), the world's largest anti-poverty program. To explain the mechanism, I start by outlining a model showing that rising incomes could lead to firms to invest in labor-saving, but polluting, mechanization.[4] My model suggests that farmers may have responded to NREGA by mechanizing harvest. This is consistent with both Hornbeck and Naidu (2014) and Clemens *et al.* (2018), who show that farmers mechanize after shocks to the low-skill labor market in the U.S. Mechanization could lead to an increase in cropland fires in India because it leaves between 80% and 120% more biomass on a field relative to manual harvesting (Yang *et al.*, 2008; Jitendra *et al.*, 2017). Biomass must be removed prior to

---

et al., 2002; Cole *et al.*, 1997; Cuaresma and Heger, 2019; Wilebore *et al.*, 2019).

[2] In section 2 I also discuss how Pigouvian policy could induce producers to adopt a cleaner labor-saving technology. Rising incomes might also lead to worse environmental quality as consumption increases with higher incomes.

[3] Humphries (2013) and Kelly *et al.* (2014) argue that wages were not in fact higher in England at the time of the Industrial Revolution and adoption of labor-saving technology was driven by concerns other than wages. Higher subsequent wages were a side-effect of increased labor productivity. Zheng and Kahn (2017) point out that the causality could also go in the opposite direction: firms in polluted cities in 19^{th} England had to pay higher wages to attract workers *because* of the pollution.

[4] The mechanism increasing pollution in my model is different from, but not mutually exclusive to, arguments that income's causal impact operates by increasing consumption (e.g. Alix-Garcia *et al.* (2013)). I show evidence that the consumption mechanism does not appear to be operating in my context.

planting and fires are the least expensive way to remove biomass (Ministry of Finance, 2018).

To explore these ideas, I use a difference-in-differences framework, taking advantage of the staggered roll-out of NREGA coupled with data on nearly one million fires, to show that of NREGA increased cropland fires by between 9% and 21%.[5] NREGA is representative of the common workfare approach to raising incomes in which governments guarantee work in return for some combination of cash and/or aid.[6] NREGA statutorily covers all rural districts in India and guarantees residents employment on public projects that use low-skill labor. Wages on NREGA projects are set by the statute and low-skill wages are paid by the Federal government. NREGA was rolled out sequentially in roughly one third of the districts in India each year in 2006, 2007, and 2008. After the implementation of NREGA, low-skill wages increased by between 4% and 8% (Imbert and Papp, 2015).

The contribution of cropland fires to air pollution is an acute policy challenge in India. As much as 40% of the pollution in Delhi during the winter may be due to crop residue burning (Bikkina *et al.*, 2019). Increases in cropland fires are believed to be a major reason that winter air quality in Delhi is among the worst in the world (Shyamsundar *et al.*, 2019) with levels of $PM_{2.5}$ that exceed WHO standards by as much as 1,000% (Liu *et al.*, 2018). To my knowledge this is the first paper that uses quasi-experimental methods and data from all of India to identify potential causes of the rise in cropland fires. Satellite data allows me to construct measures that focus precisely on the outcome of interest - cropland fires - as opposed to all fires.[7]

I show that air emissions also increase after the implementation of NREGA. Focusing specifically on air pollution from burning biomass I show that NREGA increased the emission rates of particulate

[5] I solve a second problem, in addition to simultaneity, that has made studying the relationship between income growth and environmental quality difficult. In many contexts, especially in developing countries, data on environmental quality is sparse and mis-measured (Donaldson and Storeygard, 2016). This limits the ability of researchers to measure outcomes over large spatial and temporal horizons. I use data on fire location from NASA's MODIS satellite platform that provides more than a decade of data on my primary outcome covering all of India to overcome this obstacle. I combine this with remotely sensed data from the European Space Agency on land use across India to isolate fires that occur on cropland. By focusing on cropland fires I am explicitly not examining the impact of NREGA on swidden or "slash-and-burn" agriculture. In all my analysis I focus on mainland India and omit disputed territories.

[6] The World Bank records 94 labor-based programs active in 2018 among the 142 (66%) countries for which they have programmatic data (Bank, 2018).

[7] The use of cropland fires to clear residue after harvest is a long-standing practice globally and contributes substantially to local air pollution. Agricultural burning is widely used in Pakistan and China and is used more intensively in Africa than anywhere else (Korontzi *et al.*, 2006). The drivers of cropland fires appear to be similar across countries (Cassou, 2018; Andini *et al.*, 2018) thus understanding these drivers is important beyond India.

pollution. I join a growing body of literature in both development and environmental economics that uses remotely sensed data to measure pollution over large areas with otherwise poor monitoring coverage (Auffhammer, 2018; Barrows *et al.*, 2018). I show that the rate of emissions from biomass burning of black carbon, organic carbon, and SO_2 – pre-cursor pollutants to $PM_{2.5}$ and PM_{10} – increased by between 30% and 50% after the implementation of NREGA.[8] I also show that the number of months in which the Indian ambient standard for $PM_{2.5}$ is exceeded increases by 11% after implementation of NREGA.

In showing that NREGA increases pollution emissions I contribute to the literature on the health impacts of NREGA (Thomas, 2015; Dasgupta, 2017; Nair *et al.*, 2013). Emissions from cropland fires have been shown to have negative health consequences (Rangel and Vogl, 2016; Pullabhotla, 2018) and so my results suggest that there may have been important health impacts from NREGA, particularly in downwind districts, that are not captured by existing work.[9]

To use the variation in implementation timing to identify the impact of NREGA I must assume that, in the absence of NREGA, the frequency of cropland fire would have evolved similarly in districts that received NREGA in a given year and those that did not. This is a fundamentally untestable assumption. However, as a proxy I find no evidence of pre-trends in cropland fires prior to the implementation of NREGA. My results add to the growing literature on the impacts of NREGA by expanding that literature to environmental and agricultural impacts.[10] I also provide further evidence that large-scale interventions in agricultural labor markets can have perverse effects (Lee *et al.*, 2017).[11]

Finally, I return to the mechanisms driving the increase in cropland fires. My model predicts that if mechanization is the mechanism through which NREGA increased fires, districts that had more

[8]Biomass emissions are calculated from satellite emissions data (van der Werf *et al.*, 2006).

[9]It is unclear whether my results suggest these previously estimated health impacts are an over or under-estimate of the true effect. Previous estimates will be net of any negative effect *within* a district which suggests they underestimate the true effect. However, downwind pollution created by NREGA means the SUTVA assumption may not hold in previous studies employing difference-in-differences, which suggests that their estimates may be biased upwards.

[10]The large literature on the impacts of NREGA on labor markets and other outcomes motivated and informed this project. For a comprehensive review see Sukhtankar (2016). Bhargava (2014) and Gehrke (2013) are particularly notable for examining the impact of NREGA on agricultural production practices and output.

[11]These effects are not always negative. If development depends on increasing the productivity of the agricultural sector shifting to mechanization may be a positive outcome (Herrendorf *et al.*, 2013). The negative effects arise here because of a failure to internalize social costs of pollution.

fires prior to NREGA should have seen larger increases after the implementation of NREGA. Further, districts where more farms had mechanized prior to NREGA should also see a greater increase in the number of fires after NREGA. To test these predictions I conduct two heterogeneity analyses. In the first I examine the differential impact of NREGA by the number of pre-NREGA fires in a district. In the second I assign each district a score in a mechanization index and examine how the impact of NREGA varies across this index.[12]

I show that NREGA leads to a large increase in cropland fires in districts with high levels of pre-NREGA fires using both the national set of districts in which NREGA was implemented and a subset of districts in which an RCT randomly improves the implementation of NREGA.[13] Next, I show in the full sample that fires increase by 27% in districts with the highest mechanization score compared to no statistically significant change in districts with low mechanization scores. I measure the same pattern of results in the RCT data. This pattern of impacts is consistent with my model predictions and suggests that mechanization in response to higher wages may be the channel through which NREGA increased cropland fires. This complicates policy recommendations encouraging distribution of land among many small farms (Sanchez *et al.*, 2019). There is a clear positive relationship between farm size and fire use in India, consistent with larger farms being more able to mechanize.[14] On the other hand, precisely because larger farms are more likely to enjoy economies of scale and be able to afford capital investments, they may be more able to invest in capital equipment that enables shifts away from fire use (e.g. seed drills) (Shyamsundar *et al.*, 2019; Cassou, 2018).

To further isolate the mechanization channel I test for evidence of two alternative mechanisms: (1) an increase in production that may have led to more fires and (2) a shift in cropping patterns due to NREGA's role as implicit insurance. Increases in income driven by NREGA may have led to changes in local demand that raised production. I show reduced form estimates of NREGA's impact on total hectares planted and total tonnage produced that suggest there were not large changes in the

[12] I construct the mechanization index based on predictors of mechanization cost, including the level of mechanical harvesters, collected from the Indian Agricultural Input Survey.

[13] I analyze data from the RCT conducted by Muralidharan *et al.* (2016) in Andhra Pradesh.

[14] Evidence from both China (Wang *et al.*, 2018) and Indonesia (Yamauchi, 2016) support the claim that larger farms are more able to substitute mechanization for labor.

area or tonnage of crops specifically associated with the use of fires.[15] Alternatively, NREGA provided implicit insurance that may have allowed farmers to shift into higher value but higher variance crops. While others have found that NREGA induced shifts into higher value crops (Raghunathan and Hari, 2014) I do not find evidence of an increase in the overall volume of agricultural output due to NREGA. Further, the crops that are shifted into are not associated with the use of fires in my data.

My results speak to the growing body of work on "envirodevonomics" outlined by Greenstone and Jack (2013) and the general question of how economic development – and concurrent changes in incomes, consumption, and production – impacts environmental quality.[16] I contribute most directly to the work examining the environmental impacts of raising incomes (Alix-Garcia *et al.*, 2013; Gertler *et al.*, 2006).[17] I move this work forward by showing, using an exogenous shock, that raising incomes led to an increase in pollution and introducing evidence that the up-sloping portion of the EKC may capture a causal relationship in some settings.[18]

The trade-off between income growth and environmental quality does not imply that governments should not strive to raise incomes. Rather, the trade-off necessitates understanding the mechanisms linking income growth and environmental quality to promote policies that maximize both. Producers may respond to income raising policy by choosing production processes with higher environmental externalities.[19] Government's should employ appropriate Pigouvian policies (Weitzman, 1974; Stavins, 1996; Blackman, 2010; Kremer and Willis, 2016) at all levels of development to ensure producers

[15]Data on agricultural outcomes comes from the ICRISAT Meso dataset (Rao *et al.*, 2012).

[16]Pollution, in particular air pollution, tends to be substantially above recommended limits in developing countries (Alpert *et al.*, 2012; Liu *et al.*, 2018) and these elevated levels of pollution lead to meaningful negative impacts on health and other economic outcomes (Cropper *et al.*, 2012; Ebenstein *et al.*, 2017; Barrows *et al.*, 2018). Work examining the causes of these elevated levels of pollution has focused on institutional failures (Greenstone and Hanna, 2014) and potentially lower willingness to pay for environmental quality (Kremer *et al.*, 2011).

[17]This work has often focused explicitly on the impact of anti-poverty programs. Alix-Garcia *et al.* (2013) show evidence of an anti-poverty and environment trade-off in places with little market access. But the same low levels of market access that drive the negative effects they observe may generate positive environmental change in other settings (Barbier, 2010). Determining the extent of this trade-off is especially important because the poor disproportionately live in more environmentally degraded areas (Dogo *et al.*, 2017).

[18]The need for government policy to encourage firms to choose less polluting technology in my model supports arguments that the downward sloping part of the EKC is driven by policy changes and does not result causally from income growth (see e.g. Frankel (2003))

[19]The opening example of England in the Industrial Revolution provides a non-agrarian example where increased labor costs led to increases in mechanization, and power demand, in manufacturing that results in higher pollution.

consider the full costs associated with their choices.[20]

1.2 A model of mechanization driven by wage changes

The setting in which I examine the impact of income growth on pollution is the Indian agricultural sector, but the model I present here is more general. It could apply broadly to any setting in which firms face a choice among multiple production technologies, some more labor intensive than others, and where the labor saving production technologies produce at least weakly more pollution externalities than the labor intensive technology. I start by outlining the general model and then discuss the predictions it makes in my specific context of Indian agriculture.[21]

Farmers produce crops that they sell at fixed prices, normalized to one. Each farmer has a fixed quantity of land $A \geq 1$ with fixed productivity. They face a choice between production technologies they can use to produce crop Y. One is a labor intensive production technology $G(L)$ and one is a capital intensive technology $F(L)$. F relies on mechanization to produce output so also includes the use of fire to clear residual biomass. There is no fire used in G. Farmers do not consider the social costs of F and G but I assume that F imposes a cost of emissions $e > 0$ on society. e is monotonically increasing in the number of farmers who employ F. Workers earn wages w and capital is purchased at a fixed cost. Farmers are price takers in both labor and capital markets, reflecting the relatively small size of most agricultural operations in India.

I assume that $F(L) > G(L)$ for all L, $0 < F'(L) < G'(L)$, and $F''(L) < 0, G''(L) < 0$. Further, $F'^{-1}(x) < G'^{-1}(x)$. Farmers face the associated profit conditions $\pi^G = AG\left(L\right) - wL$ and $\pi^F = AF\left(L\right) - K - wL$ and jointly choose a production technology and optimal level of labor L to maximize profits. Farmers will adopt $F(L^{**})$ when the increase in profits exceeds the cost of

[20] Kremer and Willis (2016) discuss the details of how governments should design Pigouvian subsidy policies to encourage the uptake of durables that reduce social ills. Their context is latrines to limit open defecation but the model applies to tractors that reduce air pollution as well. They argue that they dynamically optimal Pigouvian subsidy may be declining over time.

[21] I draw heavily on both Hornbeck and Naidu (2014) and Clemens *et al.* (2018). Hornbeck and Naidu (2014) show an increase in out migration of relatively unskilled labor leads to an increase in the use of labor-saving capital in response. Clemens *et al.* (2018) show how farmers rapidly mechanize agricultural production in crops with a readily available mechanization technology in response to a change in U.S. immigration policy that substantially reduces the supply of migrant labor from Mexico.

switching to capital intensive production (K). The gap in profits between $F(L^{**})$ and $G(L^{*})$ is larger when farmers have more land, which leads to the following proposition:

Proposition 1. *There exists an $\hat{A} > 0$ such that iff $A = \hat{A}$ then $\pi^F = \pi^G$, iff $A > \hat{A}$ then $\pi^F > \pi^G$ and farmers choose F, and iff $A < \hat{A}$ then $\pi^F < \pi^G$ and farmers choose G. $\hat{A} = \frac{K - wL^* + wL^{**}}{F(L^{**}) - G(L^*)}$ such that $\hat{A}$ is increasing in K and decreasing in w. Further, e is decreasing in $\hat{A}$. For proof see appendix A.1.*

NREGA acts as a shock to the agricultural labor market that makes it more difficult for farmers to hire labor. As a result, NREGA raises the equilibrium wage in the agricultural labor market (Imbert and Papp, 2015). Farmers face the same choice between G and F but the increase in w implies that $\hat{A}_{Post-NREGA} < \hat{A}_{Pre-NREGA}$ and so the number of farmers using F increases.

Finally, as an extension, consider a third production technology J^C where C denotes that this is a clean technology. Specifically, $e^G \leq e^C < e^F$ so that emissions associated with the sustainable technology are lower than with the new technology that farmers adopt. This could be because the new technology does not require the removal of residue (i.e. no till) or mechanizes the removal so that burning is not necessary (Shyamsundar *et al.*, 2019). However, consistent with the experience of Indian farmers, $K^{LE} > K$ such that the low emissions technology is more expensive to adopt than the existing labor saving technology. This highlights that without Pigouvian policy setting a price τ on the emissions associated with the existing labor saving technology (so that the farmer faces a choice of K^C versus $K + \tau$) the adoption of J^C will be below the social optimum.

1.2.1 Heterogeneity in fire response

Assume A is distributed such that $A = A_D + \mu$ where μ has a distribution $H(\mu)$ that is constant across space and the median A, A_D, varies by district. Then

Proposition 2. *The level of pollution e in a district is increasing with A_D if $\hat{A}$ is everywhere greater than A_D and μ has a constant mean zero distribution that is single-peaked at 0. For any two districts D and D' the increase in pollution for an ϵ shift downwards in $\hat{A}$ will be greater in D' if $A_D < A_{D'}$. For proof see appendix A.1.*

This implies that for a given wage shock some districts may see larger changes in fires. In particular, those districts where the A_D is higher prior to NREGA will see larger increases.

1.3 Agricultural fires in India and NREGA

In this section, I provide brief background on the use of fires in agriculture to clear crop residue in India and on the NREGA workfare program. I begin with a discussion of the use of fires.

1.3.1 Fires in Indian agriculture

While slash-and-burn agriculture is still widely used in some parts of the world (notably Africa and Indonesia) (Korontzi *et al.*, 2006; Andini *et al.*, 2018) the predominant use of fire in agriculture in India today is to clear harvest residue off of fields in order to prepare the field for the subsequent season's planting (Jain *et al.*, 2014; Bhuvaneshwari *et al.*, 2019).[22] Fires are widely used despite being nominally illegal since the mid-1990s (Lohan *et al.*, 2018). While governments in some states have begun to enforce these laws, the expected cost of violation is small. In 2012 the state government of Haryana handed out a total of roughly $12,000 in fines (Anand, 2016). That works out to an expected fine of 0.75USD per fire in the state. Farmers face an average marginal cost of roughly 50USD to clear their fields of residue without using fire (Ministry of Finance, 2018).

While fires offer an easy and inexpensive means of clearing crop residue, they have substantial negative effects. The primary negative effect is the increase of both local and global air pollutants. The largest source of carbonaceous particulate matter globally is crop residue burning (Cassou, 2018). NASA data on the practice suggests that agricultural fires contribute an average of 12.5 million tons of carbon emissions annually, roughly 1% of the global emissions from agriculture (NASA, 2017a; Smith *et al.*, 2014). More significant than their contribution to climate change is the impact that cropland fires have on local air pollution and health. Source-apportionment studies have suggested that pollution from cropland fires can raise local concentrations of $PM_{2.5}$ to more than 1,000% above the WHO 24-hour guideline of $25\mu g/m^3$ (Bikkina *et al.*, 2019; Balakrishnan *et al.*, 2019; Liu *et al.*, 2018).

[22] For a brief discussion of the history of fire use and details of its global use see appendix section A.2.

Exposure to these elevated levels of pollution leads to reductions in child height for age and weight for age scores (Singh *et al.*, 2019; Rangel and Vogl, 2016) and increased infant mortality (Pullabhotla, 2018).

There is a large body of literature that shows exposure to high levels of general particulate pollution – not solely pollution from crop burning – has negative economic consequences aside from health effects. Hanna and Oliva (2015) show reductions in pollution levels in Mexico City due to a refinery closure increased weekly hours worked. Chang *et al.* (2016) show substantial declines in worker productivity as exposure to $PM_{2.5}$ increases in California and Deschenes *et al.* (2017) show that individuals make substantial defensive investments to avoid the consequences of air pollution. Air pollution exposure may also have significant long-run effects, particularly on the formation of human capital. Ebenstein *et al.* (2016) shows that exposure to particulate pollution reduces student test scores while Voorheis *et al.* (2017) argues that early life exposure to elevated air pollution levels reduces college attendance by 19-22 year-olds in the U.S. There is also evidence that elevated levels of particulate pollution have direct negative consequences for agricultural yields (Burney and Ramanathan, 2014).

Frequent and long-term use of crop residue burning may also decrease the productivity of agricultural land (Smil, 1999; Vasilica *et al.*, 2014; Sawyer, 2019; Prasad *et al.*, 1999; Mandal *et al.*, 2004). It does so by destroying micro-nutrients in the soil and removing valuable fertilizer including nitrogen and phosphorus. Others have argued, however, that the extent to which burning negatively impacts soil quality is highly dependent on the type of soil. Further, farmers who shift from a production process that includes burning to one that removes residue from the field without burning may suffer short-term yield losses if they fail to adjust their use of fertilizer as well (Jain *et al.*, 2014; Bhargava, 2014).

The use of fire is particularly prevalent in the parts of India that grow crops in a coupled rice-wheat cropping system (Jain *et al.*, 2014; Prasad *et al.*, 1999) because of the short turn-around time between harvest of rice and planting of wheat. In this system farmers plant rice during the monsoon season (“kharif"), roughly from August to December, and wheat immediately following rice harvest during the pre-monsoon (“rabi") season from January to March or April. This system of agriculture is particularly widespread in Punjab, Haryana, Uttar Pradesh and Uttarakhand (NAAS, 2017). It is worth noting that Andhra Pradesh, the location of the RCT I analyze, is not a major producer of wheat or rice and the rice-wheat cropping process is not widespread in the state.[23]

The other major Indian crop in which fire plays a role in the production process is sugarcane. Here the fields are burned prior to harvest, and residue is burned after harvest, to make it easier to conduct the harvest (Rangel and Vogl, 2016; Jain *et al.*, 2014). In India the primary states producing sugarcane and burning residue are Uttar Pradesh, Karnataka, and Gujarat (FLA, 2012).

Figure 1.1 shows the general pattern of cropland fires across India in the years from 2003-2005, prior to the implementation of NREGA. Consistent with the expected pattern of fire use by crop type, fires are concentrated in the northwest and Indo-Gangetic plain. The sugarcane producing areas of the country also show some local hot spots. Areas that predominantly produce oilseeds, namely Rajasthan and Maharashtra, have low levels of fire.

My data confirms that districts that plant more rice, wheat, and sugarcane have more fires in the pre-NREGA period. Areas with more land in larger farms also appear to have more fires. In table 1.1 I show the correlation between several district characteristics in the pre-NREGA period and the level of fires over the same time period. Column one shows the univariate relationship while column two shows the results including all predictors in the same regression. Areas that have a higher per capita GDP in agriculture see more fires, although this relationship becomes much weaker and less significant when accounting for the total area planted in fire-associated crops in the district. The same pattern can be seen in the relationship between combines and fires. The presence of combines is important because in the rice-wheat production process existing work has suggested that farmers who harvest with combines are more likely to use fires (Yang *et al.*, 2008; Jain *et al.*, 2014; NAAS, 2017). Like with agricultural GDP this relationship becomes much weaker when incorporating cropped area, likely because of colinearity between combine presence and area in rice/wheat and the amount of land in larger farms.

[23] In appendix figures A.7a-A.7c I show the distribution of wheat, rice and sugarcane production across India in the pre-NREGA period. I also show maps of the average crop coverage on October 31^{st} each year as an approximation of which areas most heavily engage in rice-wheat production in appendix figures A.8 and A.9. It is clear from these that Andhra Pradesh, the location of the RCT improving NREGA implementation, does not heavily use rice-wheat production.

1.3.2 NREGA

Roll-out of NREGA

The objective of NREGA is explicitly to provide employment to rural households on projects that create public assets that "address causes of chronic poverty...so that the process of employment generation is on a sustainable basis" (GOI, 2007). The essential feature of the program is a guarantee of 100 days of employment for rural households in a given financial year. The first districts received NREGA in February of 2006 ("phase 1"). The targeting formula used to select these first districts is unknown (Sukhtankar, 2016).[24] However, the government had an explicit goal of including the poorest districts in the country in the first wave and every state had to have at least one district in the first wave (Shah and Steinberg, 2015).[25] After the initial roll-out another 130 districts received the program in April 2007 ("phase 2") with the remaining roughly 270 districts receiving the program in April 2008 ("phase 3").

Table 1.2 summarizes the pre-NREGA (pre-2006) level of a number of measures of economic development and the primary outcome variables in this study by NREGA phase. The table highlights that earlier districts were on average poorer, more rural, slightly more agricultural and had less land in cash crops than later districts. Differences between each phase are not equal however, with the largest differences between districts in the first two phases and those in phase three.

Impact of NREGA in the literature

Despite, or perhaps because of, the size of NREGA, program implementation has been inconsistent. This inconsistency has resulted in heterogeneous impacts across states and difficulty in precisely assessing the true impact of NREGA.[26] In the most comprehensive review of the research on NREGA, Sukhtankar (2016) suggests that there are four aspects of NREGA that the substantial literature

[24] For details on the history of workfare in India as well as background on the structure of NREGA see appendix section A.3.

[25] States may have explicitly considered the poverty level of districts when assigning them to phases (Zimmermann, 2013; Shah and Steinberg, 2015; Commission *et al.*, 2003).

[26] One reason for the heterogeneity in implementation has been substantial corruption and rent-seeking by implementing officials (Niehaus and Sukhtankar, 2013; Jha *et al.*, 2009).

agrees on: (1) the above mentioned heterogeneity in impact, (2) despite the statutory requirement that employment be provided on demand there has been meaningful rationing of employment provision, (3) NREGA has increased rural, private sector wages, and (4) the overall impact on rural productivity is ambiguous.

There is wide consensus in the published estimates that NREGA increased unskilled wages by between 5% and 8%, increased labor participation in public works, and may have led to declines in the supply of labor to the private sector (Azam, 2011; Sukhtankar, 2016; Berg *et al.*, 2012; Imbert and Papp, 2015). The most commonly cited study (Imbert and Papp, 2015) shows that NREGA increased wages of low-skill workers by 5% on average. This increase is concentrated in the dry season, when they show the bulk of NREGA work occurs, and is accompanied by a decline in private sector labor supply of 1.3%. When they focus on the states in which the fraction of time spent on public works projects by rural, prime age adults was above one percent ("star" states), they find wages increased by 9% and private sector employment fell by 3%. Deininger and Liu (2013) shows that NREGA led to an increase in the accumulation of non-financial assets in the medium run. They suggest that the overall increase in wage income exceeded the program costs. Raghunathan and Hari (2014) shows that farmers plant riskier crops after the implementation of NREGA which increases their incomes above the direct wage support of the program. Berg *et al.* (2012) suggests that the wage increase takes between 6 and 11 months after program implementation to materialize and is biased towards low-skill labor.

In the only large-scale RCT on NREGA to-date (Muralidharan *et al.*, 2016), the data of which I re-examine here, (MNS) shows that improving the implementation of NREGA increases wages by 7%. They focus only on Andhra Pradesh but note that the similarity in the size of their estimates on the impact of improving NREGA to the impacts of initial implementation highlight the importance of implementation heterogeneity. They also do not find substantial evidence of impacts on migration but note that their migration data differs from the data used in previous studies.[27]

[27] Other attempts to estimate the impact of NREGA on migration have been limited but suggest it may have had important impacts. A government review of the impacts of NREGA claims that the program led to a 27% decline in cross-district migration caused by economic distress. One of the few academic studies of the impact of NREGA on migration finds a reduction in rural to urban migration of around 8% (Imbert and Papp, 2014). While not direct evidence of reductions in migration, several studies have estimated a positive spillover of NREGA on wages in neighboring districts that is hypothesized

The impact of NREGA on education also suggests NREGA tightened local labor markets. Shah and Steinberg (2015) show that NREGA decreased the educational attainment of 13-16 year-olds by decreasing school enrollment. The effects are similar for boys and girls. They suggest that NREGA induces these changes by tightening local labor markets which raises the opportunity cost of schooling. As a result, boys leave school to provide labor in the market and girls substitute into domestic work.

Thomas (2015) finds positive impacts of NREGA on a number of health outcomes. Infant mortality declines by 6% and maternal mortality declines by 10%. These declines appear to be driven by an increase in the rate of breast feeding, institutional delivery, and immunization. Similarly Dasgupta (2017) shows that NREGA reduces the negative impact of droughts on height for age by improving child nutrition.[28] Overall the evidence on health outcomes is broadly consistent with the hypothesis that, by increasing incomes, NREGA also led to improved health outcomes.

Overall the existing literature on NREGA suggests that it had positive impacts on wages, incomes, and potentially health outcomes. It is clear that NREGA had meaningful impacts on local labor markets. In particular it appears to have tightened labor markets, especially for unskilled labor, by providing an outside option for unskilled labor in the form of public works at a wage that may have been above the prevailing agricultural wage and therefore made unskilled labor more expensive (OKeefe, 2005).

1.4 Research design and implementation

I follow much of the existing literature analyzing NREGA and utilize a difference-in-differences framework that takes advantage of the phased roll-out of NREGA across the country to examine its overall impact on the use of fires in agriculture. As a robustness check, I also utilize the treatment pattern from the MNS experiment to determine if improving the implementation of NREGA leads to

to operate via reductions in inter-district migration (Prasann, 2016; Muralidharan *et al.*, 2017). These estimates suggest that wages in neighboring districts may have increased by as much as 9% and that the estimated impact of NREGA on wages in districts in which it was implemented may be substantially underestimated due to these spillovers.

[28]Beyond Thomas (2015), comprehensive research on the health impacts of NREGA is difficult to find. Nair *et al.* (2013) use a small sample of households in Rajasthan to show that the incidence of child stunting and malnutrition is lower among families that participate in NREGA projects. Uppal (2009) finds weak evidence that participation in NREGA improves a variety of child health measures in Andhra Pradesh. Sharma (2015) shows that participation in NREGA improves child health outcomes as proxied by height and weight measures. Banerjee and Maharaj (2019) shows that NREGA reduces the negative impact of extreme heat on infant mortality.

an increase in cropland fires.

1.4.1 Data

In this section, I describe the primary outcome measure – cropland fires – as well as the construction of several key control variables.

Cropland Fires: The source of the raw data on fire presence comes from from NASA's Fire Information for Resource Management System (FIRMS).[29] The FIRMS data provides the latitude and longitude and detection time of fires around the world using imagery from the MODIS and VIIRS imaging platforms. For all the primary analysis I use only data from MODIS because VIIRS does not provide sufficient temporal coverage.[30] MODIS Aqua imagery is available from mid-2002 to the present. Imagery from the satellites are collected every 6-12 hours for every point on earth and are processed using NASA's image processing algorithm to identify fires based on the emissions of mid-range infrared radiation. The algorithm is designed to filter out spurious signals (e.g. solar glare and gas flaring). NASA suggests that the imagery can detect fires as small as 50m^2 if conditions are ideal and at sizes around 100m^2 under average conditions. They also report that fires are located at the correct location with a spatial margin of error of less than 100m on average. Unfortunately, the resolution of MODIS is such that the data available in FIRMS only measures whether at least one fire exists in given square kilometer. As a result, MODIS does not provide any information about the size of the detected fire or the total burned area. Further, it does not distinguish between pixels with only one fire and those with multiple fires.

I combine the FIRMS data on fire presence with remotely sensed landcover data from the European Space Agency's Copernicus system to determine the land uses on which fires occur.[31] Copernicus land cover data assigns each pixel to a land class based on imaging that measures its reflectivity. Classes include water/ice, urban, wetland, irrigated cropland, non-irrigated cropland, various classes of forest,

[29] I primarily use data from MODIS, available here: https://earthdata.nasa.gov/firms

[30] See the appendix for additional details on the difference between MODIS and VIIRS imagery.

[31] Copernicus data is available here:https://land.copernicus.eu/imagery-in-situ

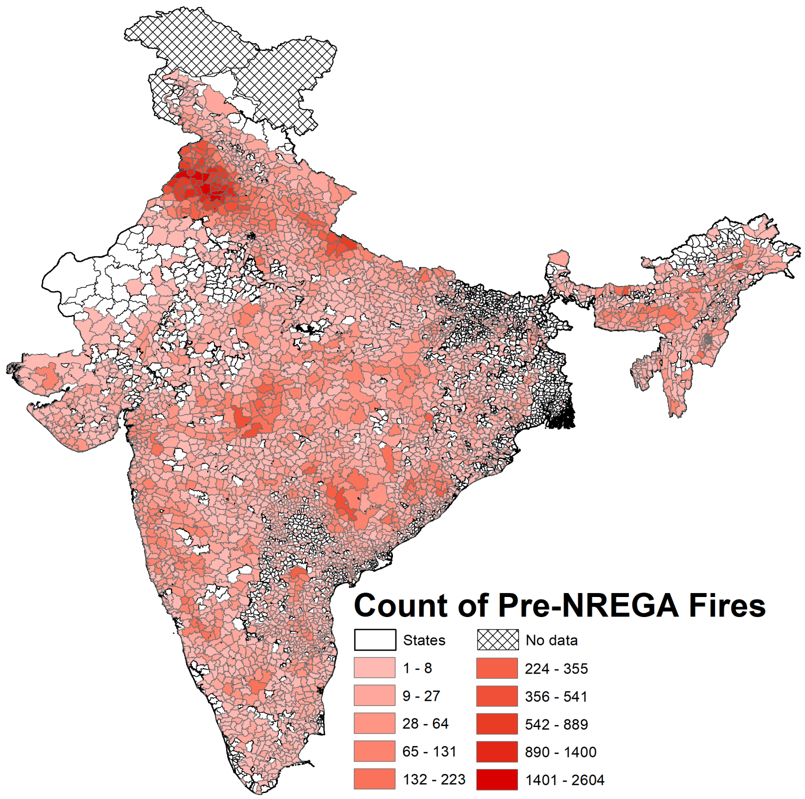
and various classes of dry shrubland. I use the cropland classes to determine which fires occur on cropland and focus my analysis on these cropland fires. I assign cropland fires based on land use in 2006. I supplement the land use data from Copernicus with additional polygon layers from the Harvard Center for Geographic Analysis (CGA) that identify roads and urban areas in India. I verify that my cropland fires do not include any fires that occur in locations that the CGA defines as urban. Finally, I aggregate the assigned cropland fires to the subdistrict or district by month level.

Figure 1.1 shows the annual average number of monthly fires per subdistrict prior to NREGA. Fires are concentrated in the northwest of the country, particularly in Punjab, and along the Indo-Gangetic Plain. However, it is clear that burning occurs on cropland throughout India. Both of these facts are confirmed by figure 1.2 showing the distribution of states by mean monthly fires in subdistricts. While fires are used in most states in India, their use in Punjab and Haryana is far more widespread than in other states (see notes on figure 1.2). Note that Andhra Pradesh, the location of the RCT conducted by MNS, is roughly in the middle of the distribution. As a result it may be broadly representative of the median state in terms of the frequency of cropland fires but it does not appear to be representative of the areas in which fire use is most widespread. However, splitting the subdistricts that are included in the MNS RCT into quartiles based on the frequency of cropland fires prior to NREGA suggests that the RCT subdistricts in the fourth quartile more closely resemble the states in which fires are used most frequently. For example, the subdistricts in Haryana have a monthly average of 3.37 fires prior to NREGA compared to a monthly average of 2.6 in subdistricts in the fourth quartile of RCT subdistricts in Andhra Pradesh.

Weather Controls: I collect weather re-analysis data from ERA5. ERA5 is a weather re-analysis product produced by by the European Commission's Copernicus Climate Change Service.[32] When working with weather data there is a trade-off between using re-analysis data, which combines observed data with a physics model to provide data at a fine resolution over long time periods, and data collected from weather stations. Station data has the advantage of being based only on observation and not including a modeled component. However, station networks often lack complete geographic coverage of a given study area and station records may be incomplete, introducing

[32]Data available here: https://cds.climate.copernicus.eu/cdsapp!/dataset/reanalysis-era5-land?tab=overview

FIGURE 1.1: PRE-NREGA FIRES BY SUBDISTRICT

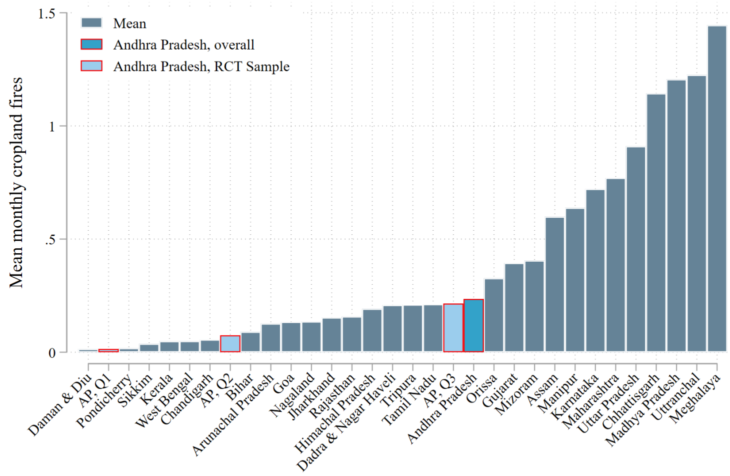


NOTES*: The count of total fires by subdistrict in the years 2003-2005, prior to implementation of NREGA. Darker areas had more fires. White areas had no fires. Data comes from the NASA FIRMS database.*

temporal gaps in coverage as well. Re-analysis solves these problems but relies on models to do so. Despite the reliance on models, re-analysis data is broadly believed to provide a reasonable best estimate of weather variables (Auffhammer *et al.*, 2013). As a result, it is widely used in both environmental and development economics (Schlenker and Lobell, 2010; Hsiang, 2016; Emerick, 2018). I chose re-analysis data for this project because comprehensive station data is not available.

Specifically I use data from the ERA5 Land hourly product. This provides data at an hourly level on a grid of $0.1^{\circ} \times 0.1^{\circ}$, which translates to a 9km resolution. ERA5 Land is an improvement over the existing, widely used ERA-Interim product (Garg *et al.*, 2018; Barrows *et al.*, 2018). Hersbach (2016) discusses the technical improvements of ERA5 Land over ERA Interim. I collect data on

FIGURE 1.2: MEAN MONTHLY CROPLAND FIRES BY STATE FROM 2003-2005



NOTES*: The mean of the average number of fires within each subdistrict in a state. The mean for Andhra Pradesh, the location of the RCT, are shown with a lighter color bar and highlighted in red. The mean number of fires is calculated by finding the average number of fires by subdistrict across the pre-NREGA years and taking the mean of these by state. The overall average for Andhra Pradesh, the location of the MNS RCT, is shown as well as the mean for the subdisricts in each quartile of the distribution of the number of pre-treatment fires. These are highlighted in red. The following are omitted from the figure because their levels are so high including them would make it difficult to see variation in the remaining states: Punjab (Mean: 18.88), Haryana (Mean:3.37), and Andhra Pradesh, Q4 (Mean: 2.61).*

cloud cover, temperature, and precipitation over the full sample from 2003 to 2014.[33] I aggregate these weather variables to the district level in the primary analysis and the subdistrict level when I work with the RCT sample.

Agricultural Data: I collect data on Indian agriculture from a number of sources. The first is the ICRISAT Meso data (Rao *et al.*, 2012). This data is collected by the International Crops Research Institute for the Semi-Arid Tropics and measures the performance, structure, and behavior of the agricultural economy at a district level in India since 1966. I use data from 2003 to 2014. Because

[33]In on-going work I am collected data on wind speed and direction for the same grid points.

Indian district boundaries have changed over time they provide both an apportioned, where they adjust data for boundary changes, and unapportioned data in which data is not re-apportioned based on changes. I use the unapportioned data and manually re-apportion data to the district boundaries as they were recorded in the 2001 census to align with my other data sources. I use data on district level cropping patterns and land holdings from ICRISAT.

I supplement agricultural data from ICRISAT with data from the Indian Ministry of Agriculture. I scrape agricultural census data from 2001, 2005, and 2011 at the district and subdistrict level. This provides additional data on characteristics of agricultural holdings and planting patterns at the district level to supplement ICRISAT data and provides resolution at the subdistrict level not available in the ICRISAT data. I also scrape the agricultural input survey data from 2001, 2006, and 2011 to collect data on machine inputs to production.[34] I also scrape data from the cost of cultivation survey to measure trends in input costs at the state level.

SHRUG Data: To measure baseline conditions in the districts in the primary analysis I use data from the 2001 Census compiled in the new Socioeconomic High-resolution Rural-Urban Geographic panel for India (SHRUG) dataset (Asher *et al.*, 2019). I use the night lights data provided in SHRUG in robustness checks as well.

Pollution Data: Previous work on the relationship between agricultural fires and pollution has relied on data from air quality monitoring stations (Pullabhotla, 2018). Like with weather data this has the disadvantage of limiting the analysis of pollution to areas with monitors that have been active over the full time period. To get around this issue, I use satellite data from the Modern-Era Retrospective analysis for Research and Applications (MERRA) database provided by NASA (Rienecker *et al.*, 2011). This is a satellite based product used by economists to study air pollution from coal fired power in India (Barrows *et al.*, 2018) that provides data on the monthly average emissions rates for black carbon, organic carbon, and sulfur dioxide (SO_2) on a $0.5^\circ \times 0.625^\circ$ grid. Importantly for my study, MERRA separately identifies the emissions of the above pollutants by source, including biomass

[34] The agricultural census can be found here: http://agcensus.dacnet.nic.in/districtsummarytype.aspx. The agricultural input survey can be found here: http://inputsurvey.dacnet.nic.in/districttables.aspx. Cost of cultivation data is here: https://eands.dacnet.nic.in/Cost_of_Cultivation.htm.

burning. I also collect concentrations of black carbon, organic carbon, sulfur dioxide, and sulfate which allows me to calculate the concentration of $PM_{2.5}$ (He *et al.*, 2019).

1.4.2 Empirical framework

Following the difference-in-differences method of Shah and Steinberg (2015) I estimate the effect of NREGA on fire use from the change in fires before and after implementation of NREGA within a district controlling for month by year and district fixed effects.[35] In the primary framework, districts are treated when NREGA becomes statutorily effective in the district. In the RCT robustness checks, treatment occurs after the baseline surveys occurred as in MNS.[36] As a result, in both cases I am estimating the Intent to Treat (ITT) effect.

I assume that the number of fires F_{imy} in district i in month m of year y follows a Poisson distribution. This is appropriate given both the count nature of the data and the skewness of the distribution of monthly fires. The standard probability density function is:

$$P = \frac{\lambda^k e^{-\lambda}}{k!} \tag{1.1}$$

which becomes:

$$f\big(F_{imy}|\mathbf{X}_{imy}\big) = exp\big(-\mu(\mathbf{X}_{imy})\big)\mu(\mathbf{X}_{imy})^{F_{imy}}/F_{imy}! \tag{1.2}$$

where $\mathbf{X}_{imy}$ is the set of observed covariates and $\mu(\mathbf{X}_{imy}) \equiv \mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] \equiv \mathbb{E}\left[F_{imy}|N_{imy}, \delta_{my}, \psi_i, \mathbf{W}_{imy}\right]$ is the link function that defines the conditional mean of F_{imy} given $\mathbf{X}_{imy}$ in parametric form. I assume the standard exponential form for $\mu\big(\mathbf{X}_{imy}\big)$ (Ranson, 2014) and taking logs of both sides get:

$$log\Big(\mu(\mathbf{X}_{imy})\Big) = \beta N_{imy} + \omega_i \mathbf{W}_{imy} + \delta_{my} + \psi_i \tag{1.3}$$

[35] In my setting all districts are eventually treated. There has been an explosion of econometric research on the use of difference-in-differences in settings like mine including De Chaisemartin and DHaultfŒuille (2017); Roth (2018); Goodman-Bacon (2018); Athey (2018); Abraham and Sun (2018). Although several propose adjustments to the standard difference-in-differences model they are not straightforward to implement in my Poisson fixed effects model. I am in the process of implementing the fuzzy difference-in-differences approach suggested by De Chaisemartin and DHaultfŒuille (2017). Goodman-Bacon (2018) and Abraham and Sun (2018) suggest that even if the difference-in-differences estimator is biased by varying treatment effects over time, the event-study, which I show below, accurately illustrates treatment effects even if every unit is treated eventually.

[36] Appendix figure A.2 shows the pattern of NREGA roll-out. Appendix figure A.3 shows the treatment pattern from Muralidharan *et al.* (2016).

where N_{imy} is an indicator that NREGA had been implemented in district i in month m and year y. $\mathbf{W}_{imy}$ is a vector of weather controls including minimum and maximum temperature, cloud cover, and total precipitation. δ_{my} is a month $\times$ year fixed effect and ψ_i is a district fixed effect. I estimate versions of equation 1.3 with and without $\mathbf{W}_{imy}$. Because of the impact that cloud cover has on the ability of satellites to detect fires and the impact that temperatures and precipitation can have on the presence of fire, my preferred specification includes $\mathbf{W}_{imy}$.[37] β is the estimate of interest and measures the approximate percentage change in monthly fires after the implementation of NREGA.[38] I cluster standard errors at the district level (Abadie *et al.*, 2017).

I estimate this fixed effects Poisson model using maximum likelihood (Hausman *et al.*, 1984; Wooldridge, 1999; Correia *et al.*, 2019). I choose a Poisson model both because of the count and skewed nature of the outcome fire data and because of the properties of the fixed effects Poisson estimator.[39] Since fires are not used in every district or every month my data has many zeros. The Poisson model accounts for these zeros more easily than a linear fixed effects model with $\log(F_{imy})$ and avoids the bias caused, when the share of zeros is non-trivial, by some common methods of transforming data to account for zeros (Nichols *et al.*, 2010). Further, the fixed effects Poisson, estimated using maximum likelihood, produces unbiased estimates of the coefficients even if my fire data does not exactly match the Poisson distributional assumptions (Wooldridge, 1997, 1999; Lin and Wooldridge, 2019). The same cannot be said for other common approaches to dealing with data with many zeros like the negative binomial or zero-inflation model (Blackburn, 2015).[40] Another advantage of the fixed effects Poisson is that it avoids the incidental parameters problem (Charbonneau, 2012; Cameron and Trivedi, 2001) which allows me to estimate a model with a large number of geographic

[37]Cloud cover introduces non-classical measurement error into my estimates and failing to control for it may lead to attenuation bias (see appendix section A.4). There is also evidence that large-scale burning induces cloud creation (Fromm *et al.*, 2010; Gatebe *et al.*, 2012; Jain *et al.*, 2014). This will exacerbate the attenuation effect.

[38]The precise interpretation of β is the difference in the logs of the expected count of fires. For small changes this is approximately equal to the percent change in the count of fires.

[39]As a robustness check I run a standard OLS fixed effects specification as well. Results are qualitatively the same.

[40]Ranson (2014) points out that this robustness does not hold for the covariance matrix. I run robustness checks where I bootstrap my standard errors as suggested in Ranson (2014). The results of the bootstrapping confirm the results from the primary estimation technique.

and temporal fixed effects.

In any difference-in-differences study the crucial identifying assumption is that the trends in the outcome variable, in this case cropland fires, would have been similar across all groups without the treatment. This is a fundamentally untestable assumption. Instead, common practice is to test whether the trends were parallel prior to the implementation of the policy being evaluated. Here that requires that the trend in cropland fires in the years leading up to treatment in phase 1 districts was similar to that in phase 2 and phase 3 districts. Like Imbert and Papp (2015) and Shah and Steinberg (2015) I am relying on the assumption that the assignment of districts to NREGA phases was based on features of the district that do not include the trend in fires *and* are not correlated with trends in fires.

I show evidence in figure 1.3 that the trends in pre-NREGA fires did not differ across the phases. Figure 1.3 shows the results of an event study on the year of NREGA implementation where the outcome is monthly cropland fires. There is a clear and significant increase in fires after the implementation of NREGA but little evidence of trends in the number of fires prior to NREGA's implementation. I show event studies for a number of other outcomes (e.g. area planted in various crops) in the appendix.

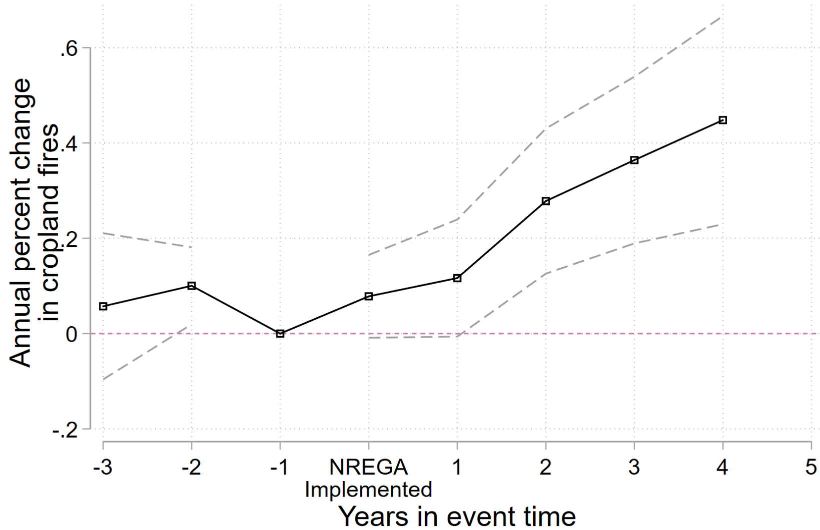
Framework for analyzing the RCT data

The RCT conducted by MNS focused on improving the implementation of NREGA by improving the provision of biometric smartcards connected to bank accounts that enabled electronic payment of NREGA wages.[41]. Electronic wage payments reduced the opportunity for corruption in the payment process, reduced the time between work and payment, and increased the likelihood that workers received payment for their participation in NREGA. This in turn increased participation in NREGA projects and wages received from NREGA projects.

The government of Andhra Pradesh (GoAP) began the initial smartcard program in 2006 with the beginning of NREGA. However, early implementation was heterogeneous because different banks were used to implement the program in different districts. In 2010 the GoAP restarted the program in

[41] I provide a brief description of the RCT here. For more details see the appendix or Muralidharan *et al.* (2016) and Muralidharan *et al.* (2017)

FIGURE 1.3: IMPACT OF NREGA ON FIRES



NOTES*: Each point is the estimated ω_τ coefficient from the regression* $\log\left(\mathbb{E}\left[F_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau\in T}\omega_\tau Y_{\tau i} + \mathbf{W}_{imy} + \psi_i + \delta_{my}$*, where Y_τ is an indicator for event-time year τ in the set $T = \{-3,-2,-1,0,1,2,3,4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\mathbf{W}_{imy}$ are weather controls. F_{imy} is the number of cropland fires in month m in year y in district i. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year prior to NREGA implementation.*

eight districts in which initial implementation had been particularly poor. MNS were able to randomize the timing with which subdistricts in these eight districts received the new program. Specifically, 112 of the subdistricts were assigned to a treatment group, 45 to a control group, and the remaining 139 to a buffer group. The treated subdistricts received the program beginning in June 2010 and there was a two year lag between implementation in the treated and control subdistricts. Baseline surveys were conducted in the treated and control subdistricts prior to implementation and endline surveys were conducted in treated subdistricts prior to implementation in control subdistricts. MNS show balance across the treated and control subdistricts on their outcomes variables as well as a range of baseline socioeconomic characteristics. I replicate their balance table in the appendix (appendix table A.1).

Unfortunately, the balance does not extend to the frequency of cropland fires. While the difference in pre-treatment fires between treated and control subdistricts is not significant the relative difference is sizable. Control subdistricts have approximately 30% fewer fires than treated subdistricts.

Because the assignment of subdistricts to treated and control in the MNS RCT did not result in perfect balance in the frequency of pre-treatment fires across the treated and control subdistricts (see appendix table A.2) I use the same difference-in-differences approach when analyzing the RCT sample. I modify the estimating equation from 1.3 to be:

$$log\Big(\mu(\mathbf{X}_{imy})\Big) = \beta T_{imy} + \omega_i \mathbf{W}_{imy} + \delta_{my} + \psi_i \tag{1.4}$$

where in the RCT sample I replace N_{imy} with T_{imy} – an indicator for treatment having occurred in subdistricts i in month m in year y where treatment occurs in the treated subdistricts after the baseline survey in 2010 as in Muralidharan *et al.* (2016). Further, in the RCT sample ψ_i becomes a subdistrict, rather than district, fixed effect.

1.5 Main results

I present the main results in three subsections. First, I discuss the impact of implementing NREGA on fires across the entire country using my primary difference-in-differences specification and the full country sample. Second, I present the results from the RCT that improved the implementation of NREGA in Andhra Pradesh. Finally, I present results that show how the effect of NREGA varies in both samples when (sub)districts are divided based on the level of fires in the unit prior to the implementation of NREGA.

1.5.1 Nationwide mean impacts

Table 1.3 shows that NREGA increased fires by approximately 21% after implementation (column 2 of table 1.3). Although the confidence interval is wide – I cannot rule out an increase of between 11% and 30% at 95% confidence – I can easily reject a zero effect at 99% confidence. This suggests that NREGA had a sizable impact on the frequency of cropland fires. To put these estimates in perspective, the average number of monthly fires increased by approximately 40% from the beginning of my sample in

2003 to the end in 2014. The estimates here suggest that between 25% and 75% of that increase can be explained by NREGA.

The estimates of NREGA's impact when I do not to control for the weather conditions in the district at the time of fire detection are substantially smaller than the estimates in the preferred specification (column 1 of table 1.3). This is consistent with the expectation that failure to control for cloud cover biases the estimates towards zero because it introduces non-classical measurement error.[42] Even with this potential bias, when I do not control for weather I estimate that NREGA increased the frequency of cropland fires by approximately 10% with a range from 1% to 18% at 95% confidence. While the estimated impact when I do not control for weather is smaller than my preferred specification, the expected impact is still meaningful and there is reason to believe that this estimate may be biased downwards.

1.5.2 Mean impacts using the RCT sample

When I turn to the RCT subsample the results in table 1.4 show that the improvement in the implementation of NREGA did not have a large average impact on cropland fires in the subdistricts of Andhra Pradesh where the experiment was conducted. While the estimates in table 1.4 are imprecisely estimated - I cannot rule out an increase in fires of roughly 28% with 95% confidence - the estimates for both specifications (controlling for weather and not) are close to zero. There are several explanations for the difference in results. The first is that fires do not appear to have been widely used in Andhra Pradesh prior to the implementation of NREGA. As figure 1.2 showed, the average number of monthly fires prior to NREGA in Andhra Pradesh is far below the average in states known for using fires. To the extent that there is learning by doing (i.e. I learn how to use fires from my neighbor who uses fires) or protection from legal ramifications of burning (i.e. I can blame sparks from my neighbors field for the fire on mine if the authorities attempt to fine me) we may expect areas with higher levels of fires prior to NREGA to have seen larger effects. Second, as table 1.1 shows, the cropland fires are associated with certain crops and certain agricultural practices. In particular, places that practice

[42] Because the measurement error is not classical the standard result that measurement error in the dependent variable only reduces precision, and does not introduce attenuation bias (Hausman, 2001), does not apply. See the appendix section A.4 for a discussion of this measurement error.

coupled rice-wheat production are more likely to use fires and areas that have higher cropping levels in October have more fires (Jain *et al.*, 2014; Bhuvaneshwari *et al.*, 2019; Shyamsundar *et al.*, 2019). Andhra Pradesh has relatively little area in agriculture in October and, on average, plants less area in coupled rice-wheat production than states in which fires are widespread (see appendix table AA.2). Further, the level of combines, an important determinant of the share of residue in a given field that is burned (Yang *et al.*, 2008), is lower in Andhra Pradesh than in states with more fires.

1.6 Implications for air pollution

The primary concern regarding the use of fires to clear crop residue stems from concern about that this practice may increase air pollution. Previous estimates of the contribution of crop burning to pollution in Delhi suggest that substantial amounts, from 17% to 60% of particulate emissions in Delhi in the winter months are the result of upwind crop burning Liu *et al.* (2018); Bikkina *et al.* (2019). In a direct examination of the impact of upwind crop burning on infant mortality in India, Pullabhotla (2018) suggests that an increase of five upwind fires in a given year increases the infant mortality rate by approximately 10%.

The availability of satellite measures of emissions enables me to directly examine how the increase in cropland fires caused by NREGA translates to an increase in the emissions of three precursor pollutants of PM_{10} and $PM_{2.5}$: black carbon, organic carbon, and sulfur dioxide (SO_2). I use data from the MERRA-2 satellite platform that measures the monthly emissions rates from biomass burning of these three pollutants by district to determine the impact of NREGA on their emissions. To do so I replace cropland fires (F_{imy}) as the outcome of equation 1.3 with the monthly average emissions rate of each of these three pollutants and re-estimate the same difference-in-differences model on the national sample.

1.6.1 Direct effect of NREGA on pollutant emissions

To begin I present the event study for the emissions rate of each of the three pollutants in figure 1.4. Each panel shows the trend in emissions rates for black carbon, organic carbon, and SO_2 before and after the implementation NREGA. In all three panels there is a clear increase in emissions rates in the year of NREGA implementation that continues to grow in the year after implementation before leveling off and remaining above pre-NREGA levels. For both black carbon and organic carbon the pre-trends are relatively flat and not-distinguishable from zero. SO_2 shows more of an increasing trend but this is driven primarily by the estimates on the earliest years in the sample.[43]

[43] I pool the early and late years of the sample because I do not have a balanced sample in event time. In other words, the estimates of the coefficient on event time -4 is only identified by districts in Phase 2 or Phase 3. As a result, I pool event time -5 and -4 with -3 in figure 1.4. If I do not do this pooling, and drop event time -4 and -5 instead, the slight trend disappears.

TABLE 1.1: WHAT PREDICTS AVERAGE MONTHLY FIRES OVER 2003-2005?

Outcome: Average cropland fires by district	(1)	(2)
Avg. Per Capita GDP	.099	.124
	(.115)	(.132)
Avg. Per Capita GDP, Ag	.323***	.13
	(.078)	(.093)
Avg. Area Planted in Rice (000s HA)	.44***	.237*
	(.099)	(.121)
Avg. Area Planted in Wheat (000s HA)	.389***	.189***
	(.034)	(.07)
Avg. Area Planted in Sugarcane (000s HA)	.238***	.23***
	(.046)	(.014)
Avg. Area Planted in Other Crops (000s HA)	-.025	-.122
	(.062)	(.105)
Combines in 2006 (000s)	.119***	.018
	(9.0e-03)	(.015)
Share of holdings >4 HA	.5***	.471***
	(.064)	(.091)
Covariates regressed separately	X	
State FE	X	X

NOTES: In all columns the outcome is the average number of monthly fires in a district averaged over the years 2003-2005. In column 1 each row is a seperate regression. In column 2 all covariates listed down the left are included in the same regression. In all cases the specification is a fixed effects poisson regression. All independent variables are measured as Z scores so that units are comparable. Per Capita GDP (Ag) reports the average per capita value of (agricultural) GDP in the district in from 2003 to 2005 measured in Lakh Rs. Average Area Planted in Wheat, Rice, and Sugarcane measures the district average area in each crop from 2003-2005 in 1,000s of hectares as reported in the ICRISAT Meso data (Rao *et al.*, 2012). Total area reports the total area in all other crops also in 000s of HA from the same data. Combines reports the number (in 000s) of self-propelled combines in the district as recorded in the 2006 Agricultural Input Survey. The Share of Holdings in each size class reports the share of acreage in the district in holdings within each size class in 2005 as reported in the ICRISAT Meso data. All columns include state fixed effects. Standard errors clustered at the state level are reported in parentheses.

TABLE 1.2: SUMMARY OF PRE-NREGA ECONOMIC, AGRICULTURAL AND FIRE DATA

	Phase 1 (1)	Phase 2 (2)	Phase 3 (3)
Data from 2001 Census			
Total Population (000s)	1,712.12	1,772.15	1,693.61
Total Households (000s)	320.12	326.19	320.43
% Rural	0.84	0.83	0.73
% Urban	0.16	0.17	0.27
% Scheduled Castes	0.15	0.15	0.15
% Literate	0.48	0.52	0.57
% with domestic electricity	0.88	0.87	0.91
% with agricultural electricity	0.74	0.59	0.72
% with electricty	0.75	0.65	0.74
% Paved road	0.49	0.56	0.68
% Mud road	0.82	0.82	0.76
Data from 2011 Census, 2012 SECC and VCF Data			
Distance to nearest place with >100k population	50.78	41.87	40.74
Per capita consumption, rural	15.64	16.59	19.02
% HH with cultivation as main income	0.35	0.37	0.38
Avg. Forest Cover, 2002-05	11.31	15.15	13.34
Avg. Night Lights	29.50	27.93	61.63
Data from NASA FIRMS			
Monthly Fires prior to NREGA	3.61	3.73	7.33
Data from ICRISAT, Planning Commission and 2006 Input Survey			
Avg. Per Capita GDP, Ag	3.70	4.22	4.90
Avg. Per Capita GDP	14.39	15.13	21.26
Avg. Area Planted in Sugarcane (000s HA)	5.86	5.71	10.36
Avg. Area Planted in Rice (000s HA)	102.43	117.19	58.77
Avg. Area Planted in Wheat (000s HA)	39.76	57.85	61.63
Avg. Area Planted in Other Crops (000s HA)	172.67	165.39	207.96
Share of holdings >4 HA	0.26	0.28	0.33
Combines in 2006 (000s)	1.41	1.12	1.41

NOTES: Columns 1-3 report the mean of the named variable by district according to the NREGA phase of that district. Data from the 2001 Census, the 2011 Census, the 2012 SECC and the VCF data come from the SHRUG dataset (Asher *et al.*, 2019). Fire data is downloaded and assembled from the NASA FIRMS data and is derived from imagery from the MODIS satellite. ICRISAT data comes from the ICRISAT meso dataset (Rao *et al.*, 2012). GDP data is scraped from the Indian Planning Commission website and covers the years 2003-2005 for most districts. Information on combines is scraped from the Indian Ministry of Agriculture website and comes from the 2006 Agricultural Input survey.

TABLE 1.3: NATIONWIDE IMPACT OF NREGA ON MONTHLY FIRES

	Cropland Fires	
Post-NREGA	0.096**	0.213***
	(0.044)	(0.051)
Districts	558	558
Months	144	144
N	80,352	80,352
Avg. monthly fires pre-NREGA	5.46	5.46
District FE	X	X
Year × Month FE	X	X
Weather Controls		X

NOTES: Each column represents separate regressions. In all columns the outcome is monthly cropland fires. In all columns the coefficient can be interpreted as the approximate percentage change in fires after NREGA was statutorily implemented in a district. In all columns the base regression is a Poisson of the form $log\left(\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right]\right) = \beta Post_{imy} + \gamma_i + \psi_{my}$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. In column 2 I include controls for the monthly average cloud cover, precipitation and minimum and maximum temperature in district i in month, year my. N refers to the number of districts × months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2012. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

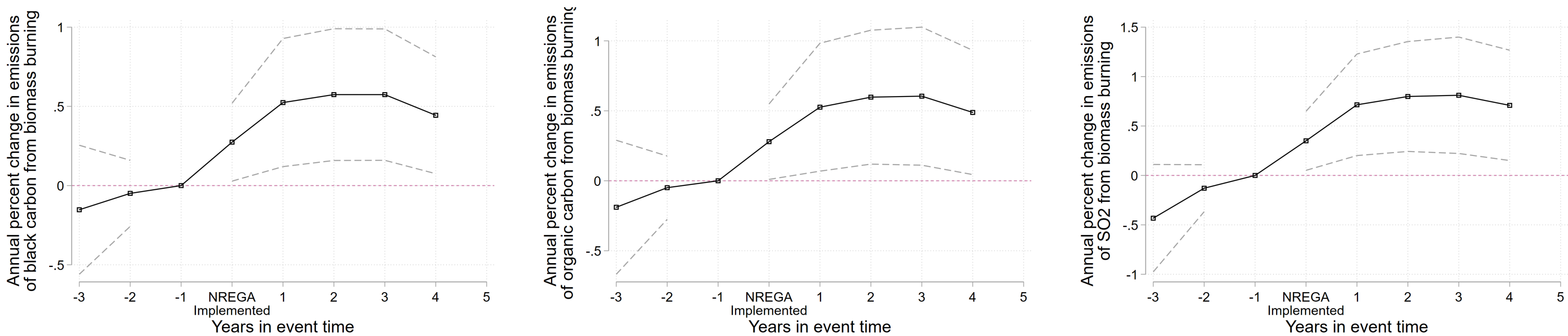
TABLE 1.4: IMPACT OF RANDOMIZED IMPROVEMENTS IN NREGA IMPLEMENTATION IN ANDHRA PRADESH ON MONTHLY FIRES

	Cropland Fires	
$\mathbb{1}$[Treated x Post]	0.027	0.018
	(0.137)	(0.143)
Subdistricts	145	145
Months	104	104
N	15,080	15,080
Avg. monthly fires pre-NREGA	.2	.2
Subdistrict FE	X	X
Year × Month FE	X	X
Weather Controls		X

NOTES: Each column represents separate regressions. In all columns the outcome is monthly cropland fires. In all columns the coefficient can be interpreted as the approximate percentage change in fires after treatment in the RCT in Muralidharan *et al.* (2016) (MNS) occurs in a subdistrict. In all columns the base regression is a Poisson of the form $log\bigg(\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right]\bigg) = exp\big(\beta[Post_{imy} \times Treated_i] + \gamma_i + \delta_{my}\big)$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after MNS treatment and Treated$_i$ is a dummy indicating that subdistrict i was among the treated subdistricts. γ_i are subdistrict fixed effects while δ_{my} is a year by month fixed effect. In column 2 I include controls for the monthly average cloud cover, precipitation and minimum and maximum temperature in subdistrict i in month, year my. N refers to the number of subdistricts × months included in each regression. The sample is a balanced, monthly panel of subdistricts in the MNS sample from 2003 to 2012. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

FIGURE 1.4: IMPACT OF NREGA ON EMISSIONS FROM BIOMASS BURNING OF VARIOUS POLLUTANTS

(A) *Black Carbon* (B) *Organic Carbon* (C) SO_2



NOTES: *Each point is the estimated ω_τ coefficient from the regression* $log\left(\mathbb{E}\left[F_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau\in T}\omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}$*, where Y_τ is an indicator for event-time year τ in the set $T = \{-3,-2,-1,0,1,2,3,4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\boldsymbol{W}_{imy}$ are weather controls. E_{imy} is the average monthly rate of emissions of the named pollutant in ng/m^{2}s in month m in year y in district i. 95% CIs are shown in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year prior to the implementation of NREGA.*

TABLE 1.5: EFFECT OF NREGA ON EMISSION RATES OF POLLUTANTS FROM BIOMASS BURNING AND MONTHS EXCEEDING AMBIENT $PM_{2.5}$ STANDARD

	Emissions Rates			Ambient $PM_{2.5}$ Standard
	Black Carbon	Organic Carbon	SO_2	Share of months > standard
Post-NREGA	0.384***	0.378**	0.491***	0.014***
	(0.143)	(0.163)	(0.176)	(0.003)
Districts	558	558	558	560
Months	144	144	144	144
N	80,352	80,352	73,656	80,640
Pre-NREGA Mean	45.81	536.09	52.09	.12
District FE	X	X	X	X
Year × Month FE	X	X	X	X
Weather Controls	X	X	X	X

NOTES: Each column represents separate regressions. In columns 1-3 the outcome is the monthly emissions rate of the pollutant named at the top of the column measured in ng/m^2s. In column 4 the outcome is the share of months in which the measured $PM_{2.5}$ concentration exceeds the annual ambient air quality standard for India set by the Air Prevention and Control of Pollution Act (1981). Concentrations are measured in $\mu g/m^3$. All data comes from the MERRA-2 satellite system. In columns 1-3 the coefficient can be interpreted as the approximate percentage change in the outcome after NREGA was statutorily implemented in a district. In column 4 the coefficient is the change in percentage points in the percent of months that exceed the Indian Ambient standard of $40\mu g/m^3$. In columns 1-3 the base regression is a fixed effects Poisson of the form $log\left(\mathbb{E}\left[E_{imy}|\mathbf{X}_{imy}\right]\right) = \beta Post_{imy} + \gamma_i + \delta_{my}$ where E_{imy} is the outcome in district i in month m in year y. In column 4 I use a linear fixed effects specification of the form $T_{imy} = \beta Post_{imy} + \gamma_i + \delta_{my}$ where T_{imy} is an indicator for whether district i had $PM_{2.5}$ levels that exceed ambient standards in month m in year y. $Post_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. In all columns I include controls for the monthly average cloud cover, precipitation and minimum and maximum temperature in district i in month t. N refers to the number of districts × months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2012. The mean of the outcomes prior to NREGA for each is presented. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

For all three pollutants that I examine, the implementation of NREGA significantly increases the emissions rate from biomass burning. Beginning with black carbon, column 1 of table 1.5 indicates that NREGA increased the emissions rate from biomass burning by 38%. Organic carbon emissions also increase by 37% while SO_2 emissions increase by an estimated 49%.

These estimated effects indicate substantial increases in the average monthly emissions rates as a result of the implementation of NREGA. They provide some confirmatory evidence that NREGA increased the number of fires; if fires increase one would expect to see a corresponding increasing in emissions from biomass burning. The size of the change in emissions rates is substantially larger than the estimated change in the number of fires. There are two explanations for this. The first is that the quantity of biomass burned per fire may have increased. This would be consistent with the model explored below of increased fires being driven by increased combine use. Previous estimates suggest fields cleared with combines have roughly 100% more biomass than fields cleared manually (Yang *et al.*, 2008). To the extent that fires that consume more biomass have higher emissions (Smil, 1999) this increase in biomass is comparable to the difference between the estimated increase in emissions rates and the increase in fires.

Fires being used more intensively following NREGA would also explain a larger increase in emissions than raw fires. While the MODIS satellite can detect fires as small as $100m^2$ it does not distinguish between pixels that have one fire and those that have many fires (Korontzi *et al.*, 2006). As a result, if the implementation of NREGA induces the use of fires in areas that already had frequent fires, MODIS will underestimate the increase in fires. Using more finely resolved data on fires in the period after 2014 from the VIIRS satellite platform I show in the appendix that MODIS does indeed under-count fires relative to VIIRS (appendix figure A.6a and A.6b).

Translating these increases in emissions rates into precise changes in pollutant concentrations is difficult and would require a model of pollution dispersion. I do not create such a model here. Rather, I estimate the impact of the change in emissions on pollution concentrations in two ways. First, I estimate the correlation between emissions rates of black carbon, organic carbon, and SO_2 using data on emissions concentrations from MERRA at the district $\times$ month level (appendix table A.3.[44] These estimated correlations suggest that in months with emissions rates that are higher than average $PM_{2.5}$ concentrations are also higher than average.

Consistent with this correlational evidence I show that the implementation of NREGA is associated with an increase of approximately 11% in the number of months in which $PM_{2.5}$ levels exceed the annual ambient standard set by the Indian government.[45] I look specifically at $PM_{2.5}$ as it has the

[44] MERRA provides estimates of the monthly concentration of black carbon, organic carbon and SO_4. I convert these into a measure of $PM_{2.5}$ using the formula described in He *et al.* (2019).

[45] The Air (Prevention and Control of Pollution) Act 1981 set standards for Annual and 24-hour concentrations in the ambient air of a number of pollutants. They are detailed here: http://www.arthapedia.in/index.php?title=Ambient_Air_Quality_Standards_in_India

largest negative impacts on health (Behrer and Mauter, 2017; Muller and Mendelsohn, 2007; Nel, 2005; Chen *et al.*, 2016). The annual threshold for ambient concentrations is set by the Central Pollution Control Board at 40 $\mu g/m^3$. I calculate the share of district $\times$ months that have an average $PM_{2.5}$ level that exceeds this threshold and use a linear fixed effects regression in the difference-in-differences framework to determine if the implementation of NREGA increased the share of months in which the threshold was exceeded.

Column 4 of table 1.5 shows that implementation of NREGA increases share of months that exceed the threshold by 0.014 percentage points. This represents an 11% increase relative to the pre-NREGA baseline rate.[46] Figure 1.5 shows the event study of the share of district$\times$months that exceed the standard. There are no obvious pre-trends and a clear increase after the implementation of NREGA.

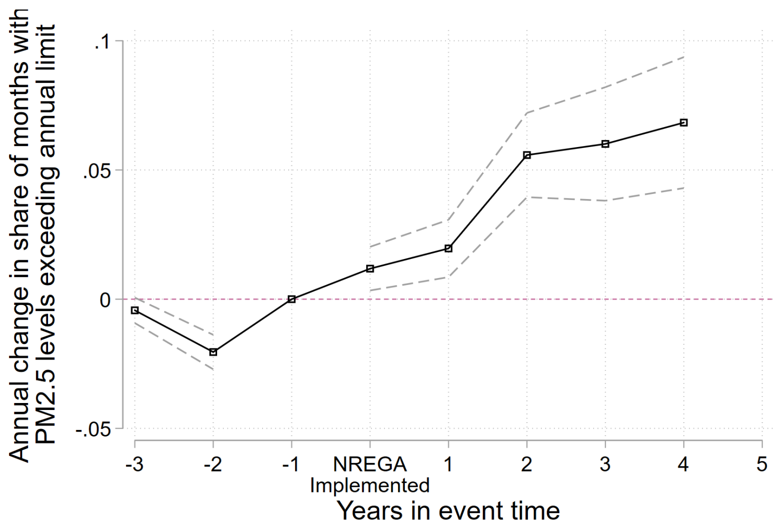
This discussion of pollution concentration has focused on the level of pollutant concentrations in the districts in which NREGA is implemented. That leaves out the change in concentration levels that are downwind of the implementing district. These downwind effects may be more severe than the impacts in implementing districts (Behrer and Mauter, 2017). Even if they are less severe considering pollution only in the implementing district presents only a partial picture of the impact of NREGA on pollutant concentrations. In on-going work I am collecting the necessary wind data to more closely examine the impact on pollutant concentrations in downwind districts.

There are potential negative health effects from an increase in pollution resulting from the observed increases in fires. While existing studies of NREGA's impact on health find generally positive effects, it is possible that these effects would have been larger absent the increase in pollution. Previous work has estimated substantial infant mortality and reductions in birth weight from exposure to cropland fire smoke (Pullabhotla, 2018; Rangel and Vogl, 2016; Singh *et al.*, 2019) and forest fire smoke Jayachandran (2009).[47] Rangel and Vogl (2016) in particular find meaningful negative impacts on infant health based on changes in concentrations similar to what I observe after the implementation

[46]When I use the 24-hour threshold of 60 $\mu g/m^3$ I observe a similar increase, 0.007 percentage points or roughly 10% on the pre-NREGA baseline rate.

[47]Additional work suggests that those same forest fires reduced the later in life wages of those children who were exposed but survived and that the pollution exposure rates they experienced may be comparable to the exposure rates residents of Delhi experience from crop burning (Tan-Soo and Pattanayak, 2019).

FIGURE 1.5: EVENT STUDIES OF THE SHARE OF DISTRICTS WHERE $PM_{2.5}$ EXCEEDS ANNUAL INDIAN AMBIENT STANDARD



NOTES: *Each point is the estimated ω_τ coefficient from the linear regression $T_{imy} = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \mathbf{W}_{imy} + \psi_i + \delta_{my}$, where Y_τ is an indicator for event-time year τ in the set $T = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\mathbf{W}_{imy}$ are weather controls. T_{imy} is an indicator for whether district i had $PM_{2.5}$ levels that exceed ambient standards in month m in year y. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year prior to NREGA implementation.*

of NREGA. An important dimension of these potential health effects is their potentially unequal distribution. Previous work (Behrer and Mauter, 2017) suggests that while the region containing the source of emissions suffers negative consequences from those emissions the majority of the damages may occur outside of the region containing the polluting source. This is particularly true if there are major cities downwind of the emitting region. Rangel and Vogl (2016) find clear evidence that crop burning has negative effects in downwind areas but, because crop burning is driven by economic activity in the burning areas, increased burning may be associated with slight improvements in health.

1.7 Heterogeneity of NREGA's impact on cropland fires

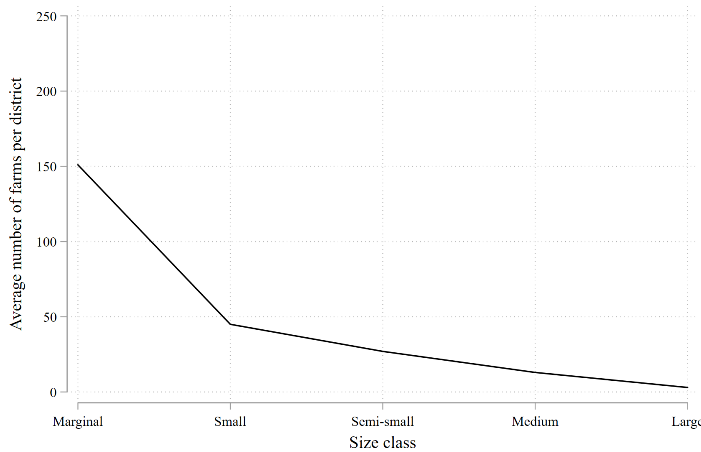
Proposition 2 of my model indicates the impact of NREGA on cropland fires will vary across districts based on A_D. Specifically, given a distribution of A that is single-peaked and where the threshold farm size for adoption of mechanization ($\hat{A}$) is above the median (A_D) of A, districts with initially higher A_D will see a larger increase in fires after the implementation of NREGA. Districts with more large farms will see a larger increase in fires after the implementation of NREGA. In other words, districts with more large farms are likely to have more fires prior to NREGA because more farmers will find themselves above a fixed $\hat{A}$ threshold. This implication is confirmed in table 1.1, which shows a strong correlation between the frequency of pre-NREGA fires and the number of large farms in a district. Districts with more large farms are also likely to be more mechanized prior to NREGA for the same reason that there will be more farms above a fixed $\hat{A}$.

Figure 1.6 indicates that the average distribution of farm sizes across India does appear to be single-peaked. Further, Foster and Rosenzweig (2011) report that in the pre-NREGA period (the late 1990s) fewer than 10% of farms were mechanized, suggesting a level of $\hat{A}$ substantially above A_D.

1.7.1 Heterogeneity of impact by pre-NREGA fire use

Based on the predictions of my model, the first heterogenity analysis I conduct is by the level of pre-NREGA fires. These districts should have higher A_D and therefore see a larger increase in fires after the implementation of NREGA. To determine whether the impact of NREGA was larger in areas that had more frequent fires prior to the implementation of NREGA I divide both samples, the

FIGURE 1.6: AVERAGE DISTRIBUTION OF FARM SIZE BY DISTRICT, 2005



NOTES*: I plot here the distribution of farms by size class across all districts in my sample in India in 2005. The vast majority of farms are marginal or small with holdings of no more than 1 hectare. Data comes from the Agricultural Input Survey in 2005. Marginal farms are those less than 1 hectare (HA), small farms are between 1 and 2 HA, semi-small are between 2 and 4, medium are between 4 and 10 and large are greater than 10.*

full national sample and the RCT sample, into quartiles based on their level of pre-NREGA fires. I then run the specifications described in equation 1.3 on the districts and subdistricts in each quartile of pre-NREGA fires.

The results in table 1.6 suggest that areas with the highest number of monthly fires prior to NREGA saw substantially larger increases in the use of fire after NREGA than areas that had lower monthly fire frequency. Column 1 reports the impact by quartile of pre-NREGA monthly fires, where higher quartiles had more fires, for the national sample. I find that districts in the fourth quartile saw an estimated increase in fires of 27% after the implementation of NREGA. This is substantially larger than the estimated effect for any other quartile and the estimates in every other quartile do not approach significance at standard levels. I can reject a null effect in the fourth quartile at $p < 0.001$. The fourth quartile estimate suggests that districts that had many fires prior to NREGA saw an additional six fires per month after the implementation of NREGA. Previous work on the relationship between fires and pollution suggests that six additional fires have a meaningful impact on downwind air pollution

(Pullabhotla, 2018).[48]

Column 2 of table 1.6 broadly confirms these results using the RCT sample. Within the RCT subdistricts those that had the highest number of pre-NREGA monthly fires see the largest increase in fires after NREGA. These districts see an estimated increase in monthly fires of roughly 45%, although from a low base. The estimated effect in the other quartiles is either not significantly different from zero or strongly negative. There are essentially no monthly fires in the lowest quartile of the RCT sample however, so the large percentage change has little economic meaning.

Because the number of subdistricts in the RCT sample is small and this heterogeneity exercise necessarily divides them further, in addition to providing the standard asymptotic estimates of statistical inference I also use a randomization inference approach. Randomization inference is an approach to inference in the style of a Fisher test that is recommended as an alternative to asymptotic based inference for RCTs, particularly those with small samples (Athey and Imbens, 2017; Young, 2018). In the randomization inference approach, treatment and control status is randomly reassigned and parameters are re-estimated providing a distribution of estimated effects under the sharp null of zero effect. The actual effect estimate can then be compared to this distribution to get an implied p-value.

I report p-values from this randomization inference approach in brackets in column 2. This approach reduces the significance of all of my estimates using the RCT sample but the estimated effect in the fourth quartile remains significant at the 5% level.

Overall, the results from the primary analysis suggest that the implementation of NREGA increased the number of cropland fires across all of India. This effect seems to have been concentrated within the areas that had a higher than average frequency of monthly fires prior to NREGA. The concentration results appear to be confirmed in the smaller RCT sample in Andhra Pradesh, an area in which fires are not frequently used relative to other parts of India. These results are consistent with the prediction of my model.

[48] I show in appendix table A.10 that the effects of NREGA do not vary across states that are more expeditious in implementing a separate government program.

TABLE 1.6: HETEROGENEITY OF TREATMENT IMPACT BY FREQUENCY OF FIRE USE PRE-NREGA

	All of India	Andhra Pradesh
(A) Quartile 1 of pre-treatment fires		
Post-NREGA (Treatment)	0.17	-0.85**
	(0.12)	(0.42)
RI p Value		[0.12]
Districts	138	T:5 C:3
Months	144	47
N	19,872	1,692
Avg. monthly fires pre-NREGA	.18	0.01
(B) Quartile 2 of pre-treatment fires		
Post-NREGA (Treatment)	0.14	-0.32
	(0.12)	(0.32)
RI p Value		[0.18]
Districts	140	T:11 C:3
Months	144	73
N	20,160	2,263
Avg. monthly fires pre-NREGA	.96	.95
(C) Quartile 3 of pre-treatment fires		
Post-NREGA (Treatment)	0.04	0.09
	(0.11)	(0.18)
RI p Value		[0.47]
Districts	140	T:25 C:13
Months	144	88
N	20,160	3,608
Avg. monthly fires pre-NREGA	2.59	2.66
(D) Quartile 4 of pre-treatment fires		
Post-NREGA (Treatment)	0.27***	0.45***
	(0.06)	(0.17)
RI p Value		[0.04]
Districts	140	T:25 C:5
Months	144	89
N	20,160	3,293
Avg. monthly fires pre-NREGA	18.12	12.16
Subdistrict FE	X	X
Month × Year FE	X	X
Weather Controls	X	X

NOTES: Each column represents seperate regressions. In all columns the outcome is monthly cropland fires. In column one the coefficient can be interpreted as the approximate percentage change in fires after NREGA was statutorily implemented in a district. In column two the coefficient can be interpreted as the approximate percentage change in fires after treatment in the RCT in Muralidharan *et al.* (2016) (MNS) occurs in a subdistrict. In column one the sample is all districts in India that were part of the NREGA program. In column two the sample is the subdistricts in Andhra Pradesh included in the MNS RCT. In column one the specification is a Poisson of the form $\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] = exp(\beta \sum_{f=1}^{4}\left[Post_{imy} \times Pre - Fires_{if}\right] + \gamma_i + \delta_{my})$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Pre-Fires$_{if}$ is an indicator for where district falls in the distribution of total pre-NREGA fires. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. In column two the specification is a fixed effects Poisson regression of the form $\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] = exp(\beta \sum_{f=1}^{4}\left[[Post_{imy} \times Treated_i \times Pre - Fires_{if}\right] + \gamma_i + \delta_{my})$ where F_{imy} is the outcome in subdistrict i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after MNS treatment and Treated$_i$ is a dummy indicating that subdistrict i was among the treated subdistricts. Pre-Fires$_{if}$ is an indicator for where subdistrict falls in the distribution of total pre-NREGA fires where the distribution is calculated within Anhdra Pradesh. γ_i are subdistrict fixed effects while δ_{my} is a year by month fixed effect. Each panel is a different quartile of pre-NREGA fires with Q4 corresponding to the largest number of pre-NREGA fires. N refers to the number of (sub)district × months included in each regression. Districts reports the treated and control (sub)districts in each sample. In the NREGA sample all districts are eventually treated. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The samples are a balanced, monthly panel of subdistricts in Andhra Pradesh from 2003 to 2012 and districts in India from 2003 to 2014. All columns include controls for weather in the month the outcome number of fires is measured. Heteroskedasticity robust standard errors clustered at the (sub)district level are in parentheses. In column two I also report the p-values from a randomization inference test on the MNS sample in which treatment is randomly assigned 1,000 times and the specification above is re-estimated for each assigned treatment. (* p<.10 ** p<.05 *** p<.01).

1.7.2 Heterogeneity in fire response by mechanization levels

The second heterogeneity test I conduct is by the level of pre-NREGA mechanization. To test whether the impact of NREGA varies by the level of pre-NREGA mechanization I construct an index of mechanization for each district. The construction of this index is driven by the features of the model described above and the relationship between fire use and mechanization in Indian agriculture as described in the existing literature (e.g. among others Jain *et al.* (2014) and Bhuvaneshwari *et al.* (2019)). I follow Asher and Novosad (2018) and use an index rather than testing multiple measures of mechanization to limit concerns of multiple hypothesis testing. [49]

To construct the index I consider the average level of the following variables in the years prior to 2006: the share of agricultural land in both marginal and medium or large holdings (less than 1 HA or more than 4 HA), the number of combines in 2006, and the production areas of wheat, rice, and sugarcane (the ability to mechanize varies across crops in India (Solomon, 2016)). To make the index I turn these averages into Z scores. Each district receives a Z score for each individual measure. I then add them together to determine an index measuring the cost of mechanization in a given district. The higher the index score the higher the predicted level of mechanization. In the appendix I show that this score is positively correlated with the number of combines in a district in 2011. I then divide districts into quartiles based on their mechanization index score, as I did when examining heterogeneity by pre-NREGA fire frequency, and run my difference-in-differences regression from equation 1.3 on districts in each quartile.[50]

I find in the full, nationwide sample, districts with the highest mechanization scores have substantially larger increases in the frequency of fire after the implementation of NREGA. Districts in the highest quartile of the index see a 27% increase in monthly fires after the implementation of NREGA (table 1.7). This is compared to an effect that is not different from zero in all other quartiles of the

[49]For details of each component of the index, as well as heterogeneity results by the individual components, see appendix section A.14.

[50]Li (2017) suggests that areas with higher land concentrations, an important piece of my mechanization score, had lower pre-NREGA agricultural wages than other districts, which suggests my mechanization score is not simply capturing districts with high wages prior to NREGA. Using, imperfect, wage data from ICRISAT I find that the correlation between the Z score of district wages in 2005 and the mechanization index is -0.11 ($p = 0.71$). I directly measure the correlation between the Z score of per capita GDP and my mechanization score and find that it is -0.07 ($p = 0.45$) which suggests that I am also not simply identifying wealthier districts.

TABLE 1.7: HETEROGENEITY OF TREATMENT IMPACT BY EASE OF MECHANIZATION INDEX

	All of India
(A)Quartile 1 of Ease of Mechanization Index	
Post-NREGA	-0.013
	(0.087)
Districts	140
Months	144
N	20,160
Avg. monthly fires pre-NREGA	2.82
(B)Quartile 2 of Ease of Mechanization Index	
Post-NREGA	-0.008
	(0.075)
Districts	143
Months	144
N	20,592
Avg. monthly fires pre-NREGA	2.7
(C)Quartile 3 of Ease of Mechanization Index	
Post-NREGA	0.144
	(0.112)
Districts	135
Months	144
N	19,440
Avg. monthly fires pre-NREGA	3.06
(D)Quartile 4 of Ease of Mechanization Index	
Post-NREGA	0.265^{***}
	(0.084)
Districts	140
Months	144
N	20,160
Avg. monthly fires pre-NREGA	13.28
District FE	X
Month $\times$ Year FE	X
Weather Controls	X

NOTES: The outcome is monthly cropland fires. The coefficient can be interpreted as the approximate percentage change in fires after NREGA was statutorily implemented in a district. The sample is all districts in India that were part of the NREGA program. The specification is a fixed effects Poisson of the form $\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] = exp(\beta \sum_{z=1}^{4}\left[Post_{imy} \times Mech_{iz}\right] + \gamma_i + \delta_{my})$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Mech$_{iz}$ is an indicator for where district i falls in the distribution of the ease of mechanization index. The ease of mechanization index is the sum of a district's Z score across measures of land concentration, combine presence and crop types. The mechanization index is calculated based on levels of each component variable in the district prior to 2006. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. Each panel is a different quartile of the mechanization index with quartile 4 corresponding to the places mechanization is predicted to be easiest. N refers to the number of district $\times$ months included in each regression. Districts reports districts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel of districts in India from 2003 to 2014. All columns include controls for weather in the month the outcome number of fires is measured. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

index.[51]

This suggests that districts in which more farmers could feasibly respond to the agricultural labor market shock caused by NREGA by mechanizing their harvests saw larger increases in fires. Larger increases in the use of fire in these districts is consistent with the predictions of my model and suggests that it was an increase in mechanization in response to the labor shock imposed by NREGA that drove the increase in fires. However, it is possible that NREGA impacted the use of fires through other channels. I explore two of these in the next section.

1.8 NREGA's impact on agricultural output

NREGA's shock to agricultural labor markets may have done more than simply changing low skill labor wages. Changing the cost and availability of labor may have caused farmers to change which crops they planted as well. Alternatively, by increasing incomes in local markets, NREGA may have increased demand, and prices, for agricultural products and incentivized farmers to increase their production of existing crops. If this increased production occurred in crops that used fire as part of the production process the increase in fires may have been driven by the increased crop production as opposed to being driven directly by changes in the labor market. I call this a consumption effect.

NREGA may have also acted as an implicit insurance program for farmers Sukhtankar (2016). By guaranteeing the availability of outside employment in the event of a crop failure NREGA may have encouraged farmers to plant higher value but higher risk crops. If these crops are associated with more fire use than the previously planted, lower value crops it could be that the insurance aspects of the NREGA program are the driving force behind the change in fire frequency. I explore both of these potential explanations now.

[51] I am limited in my ability to construct the same mechanization index for the RCT sample because a lack of data on inputs by subdistrict. However, I construct a similar index based only on plot size and crop type for the RCT sample and conduct the same exercise on the RCT sample. I show the results in the appendix. They are broadly consistent with those presented here; the largest effects are in the subdistricts that score highest on the mechanization index.

1.8.1 NREGA had little impact on production

I find that NREGA had little impact on the area planted (in 000s of HA) or total tonnage produced of crops most associated with fire production (wheat, rice, and sugarcane) Jain *et al.* (2014). When I use the full, national sample I find no effect on total area planted or total tonnage of fire associated crops or in total area planted and tonnage in non-fire associated crops. In the appendix I show event studies for each of these crop outcomes. In all cases the trend after the implementation of NREGA is flat or declining and in no case is it different from zero.

Using the same difference-in-differences approach as described by equation 1.3, replacing cropland fires (F_{imy}) with the area in each crop (A_{imy}) and tonnage produced (T_{imy}) in each crop as the outcome, I confirm the lack of impact suggested by the event studies (see table 1.8). In all crops I examine (total crops, other non-fire associated crops, wheat, and sugarcane), except rice, I estimate an effect of NREGA that is not statistically different from zero and is, in most cases, estimated to be close to zero with relatively high precision. The only crop I estimate a statistically significant effect for at standard levels is rice, which sees a small (approximately 3%-4%) increase in area under production and total tons produced. This increase is small relative to the estimated change in fires and relatively imprecisely estimated.

Based on these estimates I can reject an increase in fires of more than 0.40 % due to increasing area in wheat production at 95% confidence. Similarly I can reject changes greater than 1.60 % and 0.27 % for changes in the area in rice and sugarcane production. To determine these bounds I take the estimated change in area planted due to NREGA from table 1.8 and convert that estimated change in area into standard deviation units based on the distribution of area planted in each crop across all of India. This makes the units comparable to the units in table 1.1 where I estimate the correlation between area planted in wheat, rice, and sugarcane and the frequency of fires. Based on the correlations in table 1.1, I estimate the predicted change in fires for each crop based on the predicted change in area planted from table 1.8. I use the delta method to calculate standard errors for these estimates. I report the upper end of the 95% confidence interval for the estimated change in fires.

NREGA may have led to small increases in the area planted and total tons produced of rice. I find no evidence that it led to meaningful increases in area planted or tonnage of other crops. The estimated increase in fires due to the changes in area produced is small both in absolute terms and relative to the estimated overall impact of NREGA on the frequency of fires. No more than 30% of the smallest estimated increase in fires is estimated to be due to changes in area under production and the estimated changes in production would account for no more than approximately 10% of the estimated change in fires in the preferred specification.

TABLE 1.8: EFFECT OF NREGA ON CROP PRODUCTION

	All Crops		Wheat		Rice		Sugarcane		Other Crops	
	Area	Tons	Area	Tons	Area	Tons	Area	Tons	Area	Tons
Post-NREGA	0.011	0.016	-0.001	0.014	0.025**	0.039**	0.034	-0.119	0.006	0.032
	(0.007)	(0.022)	(0.019)	(0.025)	(0.010)	(0.019)	(0.040)	(0.102)	(0.008)	(0.060)
Districts	492	492	441	429	473	473	442	441	492	491
N	56,376	56,376	50,388	48,948	54,180	54,180	51,264	51,132	56,376	56,256
District FE	X	X	X	X	X	X	X	X	X	X
Year $\times$ Month FE	X	X	X	X	X	X	X	X	X	X
Weather Controls	X	X	X	X	X	X	X	X	X	X

NOTES: Each column represents separate regressions. In all columns the outcome is identified in the column heading. Area is measured in 000s of hectares while quantity produced is measured in 000s of tons. In all columns the coefficient can be interpreted as the approximate percentage change in the outcome after NREGA was statutorily implemented in a district. In all columns the base regression is a fixed effects Poisson of the form $\mathbb{E}\left[C_{imy}|\mathbf{X}_{imy}\right] = exp\left(\beta Post_{imy} + \gamma_i + \delta_{my}\right)$ where C_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. All columns include controls for the monthly average cloud cover, precipitation and minimum and maximum temperature in district i in month t. N refers to the number of districts $\times$ months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2012. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

1.8.2 Changes in crop choice induced by NREGA's role as insurance

Table 1.8 provides no evidence that the overall production of non-fire associated crops, which includes higher value, higher variance crops, increases as measured either by area in production or total tons. Existing work (Gehrke, 2013; Raghunathan and Hari, 2014) has suggested that farmers do in fact shift into higher variance, higher value crops such as cotton after the implementation of NREGA. I find no evidence of such a change but no evidence that production of these crops decline. More importantly, the production process of the crops that farmers may have shifted into (Gehrke, 2013) does not typically include fire (Jain *et al.*, 2014; Bhuvaneshwari *et al.*, 2019). It is therefore difficult to explain the observed increase in fires as being driven by NREGA's implicit insurance provision leading to transitions in crop type.

1.9 Conclusion

In this paper, I tested for the casual link between incomes and environmental quality by studying the impacts of India's anti-poverty program, NREGA. NREGA led to increases in the frequency of cropland fires of between 9% and 21%. It also led to large increases, between 30% and 50%, in the emissions rate of black carbon, organic carbon, and SO_2 from biomass burning. These pollutants are important contributors to both $PM_{2.5}$ and PM_{10} pollution. I also find evidence that the increase in cropland fires is concentrated in districts that had a higher number of cropland fires prior to the implementation of NREGA. In districts that had higher levels of a number of indicators of mechanization prior to NREGA I find an increase in cropland fires of 28%. Empirically ruling out alternative explanations for the increase in cropland fires, the findings are consistent with a model that suggests districts with more mechanization prior to NREGA saw larger effects on fires.

I cannot rule out that the increase in fires is driven by some other aspect of NREGA or a feature of districts that is correlated with my mechanization index. However, higher pre-NREGA wages are not highly correlated with higher scores in the mechanization index nor is district level wealth. I also observe low correlation between the mechanization index and pre-NREGA fires, suggesting that the mechanization index is not simply identifying areas with high levels of fire prior to NREGA. Despite this, it remains possible that the observed increase is driven by some factor other than changes in mechanization and these results should be interpreted with caution.

With that in mind, it is important to remember that this is not an analysis of NREGA in its entirety. Estimating the welfare consequences of NREGA inclusive of the measured increase in cropland fires is beyond the scope of this paper. Others have found meaningful increases in income, consumption, and health as a result of NREGA that may offset any negative effects of increased emissions from increases in the number of fires. How the distribution of these benefits compares to the distribution of negative impacts from the emissions increase I measure deserves the attention of future work.

The results presented here do not call for, nor justify, wholesale changes to NREGA. Nor should they be interpreted as casting doubt on the value of anti-poverty programs generally. Rather, they suggest that policy-makers should be cognizant of the potential consequences of large policy changes that raise wages and incomes on environmental quality and consider ways to mitigate negative impacts.

In the context of NREGA one possible policy response to the increase in fires in the short-run is to allow NREGA labor to be used to clear residue. There is some precedent for allowing NREGA labor to be used on private plots (GOI, 2009). Allowing the use of NREGA labor to collect residue may make the use of fire to clear it less appealing to farmers.

Another potential policy response is to expand existing programs to encourage the adoption of the agricultural practices that do not require residue removal or mechanize such removal (Shyamsundar *et al.*, 2019). Burning has declined as an agricultural practice in much of North America and Europe as the result of the adoption of “no till" agricultural practices (Korontzi *et al.*, 2006; Marlon *et al.*, 2008). Recent work has suggested that re-allocation of existing subsidies to more promising technology might encourage more widespread adoption of capital that reduces the need to burn (Shyamsundar *et al.*, 2019).

There are lessons here for policy-makers beyond India as well. Many countries around the world have a goal of raising incomes with some form of a work guarantee scheme. The results here highlight that these kinds of interventions in labor markets, and programs to raise incomes generally, can have environmental impacts by changing production decisions. Governments may be able to reduce these environmental impacts by incorporating Pigouvian tax and subsidy policies, which encourage firms to internalize environmental externalities, into income raising policies.

Chapter 2

The Impact of Historic Air Pollution Exposure on COVID19 Deaths[1]

2.1 Introduction

The coronavirus disease that emerged in late 2019 (COVID19) in Wuhan China has rapidly spread around the world to become the first global pandemic since 2009. As COVID19 has spread, a growing body of research has focused on understanding the environmental conditions that may modulate its spread, including the role of temperature (Carleton and Meng, 2020), weather (Liu *et al.*, 2020), and air pollution (Wu *et al.*, 2020; Setti *et al.*, 2020). Much of this work has been motivated by the observation that the spread of other respiratory diseases, most notably seasonal influenza, are modulated by both temperature and particulate air pollution (Singer *et al.*, 2020). We add to this by examining the impact of chronic exposure to particulate air pollution with an instrumental variables approach that suggests that particulate air pollution increases COVID19 deaths but this impact is approximately 30% smaller than that estimated using cross-sectional data (Wu *et al.*, 2020).

COVID19 is a severe acute respiratory syndrome (SARS) caused by a novel coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-COV-2) (Gorbalenya *et al.*, 2020; Zhu *et al.*, 2020; Zhou *et al.*, 2020). It endangers human health by damaging the lungs and inducing respiratory

[1] Co-authored with Christopher R. Behrer (Duke University)

distress. In severe cases, patients are unable to breathe effectively on their own and require mechanical ventilation. Roughly half of U.S. deaths from COVID19 co-occurr with pneumonia or hypertension (CDC, 2020). In these respects, it resembles previous epidemics caused by coronaviruses (e.g. SARS or MERS) but SARS-COV-2 appears to be both more virulent and more transmissible than past coronaviruses (Wu and McGoogan; Sanche1 *et al.*, 2020; WHO, 2003). Further, it appears to have more negative non-respiratory health effects (Mehta *et al.*, 2020).

Because COVID19 endangers human health primarily by attacking the lungs, the CDC has suggested that individuals who suffer comorbidities related to the respiratory system (e.g. asthma) may be more susceptible to COVID19. Individuals with lungs compromised by chronic exposure to particulate air pollution may be less able to fight off the effects of this new virus. Although it is not a coronavirus, the seasonal flu poses greater risks to patients with previously compromised respiratory systems (Singer *et al.*, 2020).

Rates of respiratory problems are higher in areas that experience higher levels of particulate air pollution. This is true both within and across countries (Jiang *et al.*, 2016; Mo *et al.*, 2018; Cohen *et al.*, 2017). Higher rates of respiratory problems in areas with higher particulate air pollution suggest that these areas may also experience more severe cases of COVID19. Cases that would manifest as mild symptoms in healthy individuals may manifest more severely in individuals compromised by perennial exposure to particulate air pollution. The association between exposure to particulate air pollution and severity of COVID19 has been proposed in the clinical literature (Misra *et al.*, 2020).

Specifically, pathophysiology of COVID19 suggests that particulate air pollution may increase COVID19 mortality by increasing its spread and by increasing its virulence. COVID19 enters the body by binding to a specific protein receptor (Hoffmann *et al.*, 2020). Chronic particulate air pollution has been associated with the up-regulation of the process that generates those proteins (Chen *et al.*, 2016). Therefore, chronic exposure to particulate air pollution may increase the ability of COVID19 to enter the body. Further, mortality from COVID19 is frequently associated with the occurrence of cytokine storms that trigger hyperinflammation (Mehta *et al.*, 2020; Chen *et al.*, 2020a).[2] Clinical literature

[2]A simple definition of a cytokine storm is a runaway reaction in the body that causes the immune system to turn on the body's own cells.

has shown that exposure to particulate pollution is associated with higher levels of cytokines in the lungs even after the exposure event (van Eeden *et al.*, 2001).

That areas with higher levels of particulate air pollution may suffer more severe cases of COVID19 is supported by more than physiology. In both Italy and China, two early hot-spots of the disease, the first and most severely effected areas were also areas with unusually high levels of particulate air pollution. The annual average $PM_{2.5}$ level in Wuhan is $\approx 150\mu g/m^3$, six times the WHO recommended limit.[3] In Italy the first cases, and the largest number of deaths to date, have occurred in the Po river valley. This is the industrial center of Italy and regularly has the highest levels of particulate air pollution in the country, and often in all of Europe. $PM_{2.5}$ levels in Milan, the largest city in the Po river valley, average between $100\mu g/m^3$ and $150\mu g/m^3$ from December to February.

In this paper we use data on the spread of COVID19 through the United States and satellite measures of particulate air pollution to examine the relationship between COVID19 mortality and historic exposure to particulate air pollution. To overcome concerns about confounding by unobservable determinants of COVID19 mortality due to the residential sorting process driven by particulate air pollution, we use an instrumental variables approach to estimate the impact of pollution levels of COVID19 mortality. We leverage the fact that roughly 50% of local particulate air pollution in the United States comes from distant power generation combined with wind direction and the hydraulic fracturing boom to instrument for local particulate air pollution with the unexpected closure of coal fired power plants 100 to 150 miles away from the county of interest.

We show that a $1\mu g/m^3$ increase in a county's ten year average $PM_2.5$ levels increases COVID19 deaths by 9%. A $1\mu g/m^3$ increase in the ten year annual average of $PM_{2.5}$ represents a 12% increase for the mean county in our sample, implying a 1% increase in particulate pollution leads to a 0.75% increase in mortality. We can compare this to recent cross-sectional estimates form the U.S. (Wu *et al.*, 2020) that suggest a $1\mu g/m^3$ increase in the ten year annual average of $PM_{2.5}$ leads to a 15% increase in mortality from COVID19. Consistent with the cross-sectional estimate being confounded by unobserved determinants of mortality that are correlated with particulate air pollution levels, our estimates for the causal impact of particulate air pollution are 30% smaller.

[3]Beijing, famous for its poor air quality, has an average of $\approx 100\mu g/m^3$

Our paper joins a long line of literature documenting the negative health consequences of particulate air pollution (Manisalidis *et al.*, 2020; Deschenes *et al.*, 2017; Chay and Greenstone, 2003a,b; Chay *et al.*, 2003). There is less work on the role that chronic exposure to particulate air pollution may play in exacerbating the spread of a pandemic. The most notable study in this space examines how particulate air pollution in U.S. cities prior to the 1918 Spanish Flu may have lead to substantial excess mortality in some cities (Clay *et al.*, 2018). Using variation in the share of electricity provided by coal fired power they find that “high coal" cities experienced all-age mortality rates that were ten percent higher than “low coal" capacity cities. They find that accounting for variation in city characteristics (e.g. poverty rates) and the public health response of the city to the pandemic does not change the impact that particulate air pollution appears to have had on excess mortality.

We briefly discuss the literature on particulate air pollution and health as well as our current understanding of the pathophysiology of COVID19 in section 2. In section 3 we outline our empirical approach and describe the IV strategy we utilize. In section 4 we describe our data and discuss results in section 5. Section 6 concludes.

2.2 Particulate air pollution and COVID19 mortality

Chronic exposure to particulate air pollution may increase COVID19 mortality by increasing transmission of the virus and by increasing its virulence once infected. We briefly highlight findings from three relevant literatures. First, we discuss existing evidence on the health effects of chronic exposure to particulate air pollution. Second, we review the limited existing evidence on the relationship between chronic exposure to particulate air pollution and the progression of epidemic and pandemic infectious disease. Finally, we discuss the emerging scientific and clinical literatures which describe the pathophysiology of COVID19, specifically the role of the angiotension-converting enzyme 2 (ACE2), and risk factors for mortality including chronic respiratory, cardiovascular, metabolic, and immunologic disease.

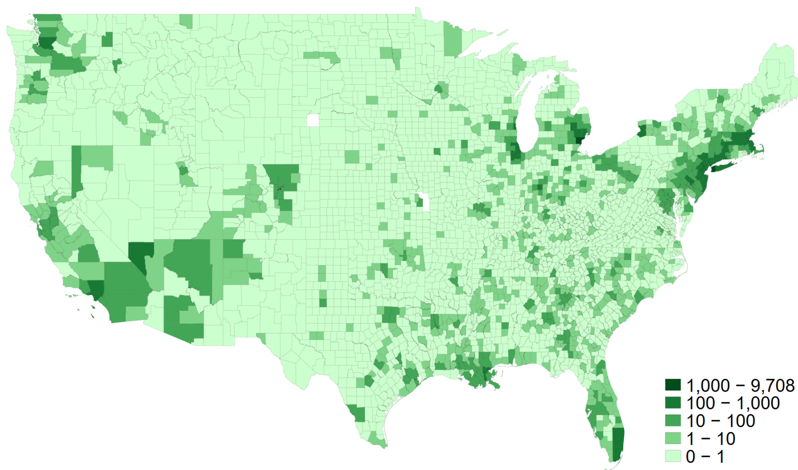
2.2.1 Health effects of exposure to particulate air pollution

It is well known that particulate air pollution exposure increases mortality from a range of causes. This is particularly true of exposure to $PM_{2.5}$ for the elderly, a population that appears disproportionately impacted by COVID19. Anderson (2015) finds a 3-7% increase in over-75 mortality for a doubling of time spent downwind of an LA highway, an increase he argues is due to elevated exposure to a range of pollutants but especially particulate pollution. More recently, Deryugina *et al.* (2019) use variation in the amount of out-of-county pollution carried into a county based on changes in the direction of the wind to show substantial increases in mortality due to $PM_{2.5}$ pollution among the Medicare population. Coal fired power plays a large role in the generation of out-of-state pollution and has substantial impacts on mortality among the very young and the elderly from respiratory disease (Clay *et al.*, 2016; Gupta and Spears, 2017). Chronic exposure to high levels of particulate pollution can have more subtle health effects as well. Bishop *et al.* (2018) estimate that a 1 $\mu g/m^3$ increase in decadal $PM_{2.5}$ levels increases dementia diagnosis by 1.68 percentage points, or roughly 8% of the mean number of dementia diagnoses among 80-year-olds.

2.2.2 Air pollution and pandemics

Work on the relationship between particulate air pollution and respiratory pandemics has been limited by the infrequency of pandemic events. However, using air pollution index (API) data, Cui *et al.* (2003) find that moving from a low API area to a high API area was associated with a doubling of the mortality rate during the 2002 SARS outbreak in China. This is consistent with more geographically resolved evidence from Beijing during the same outbreak that finds short-term increases in air pollution during the outbreak are associated with an increase in the relative risk of SARS mortality (Kan *et al.*, 2005). While they lack data on the specific pollutant loads Clay *et al.* (2018) and Clay *et al.* (2019) find that areas with higher levels of air pollution, proxied by coal power generation, experienced more excess mortality from the 1918 flu and that this effect was roughly 50% as important as pre-pandemic infant mortality rates in predicting excess mortality.

FIGURE 2.1: COVID19 DEATHS BY COUNTY



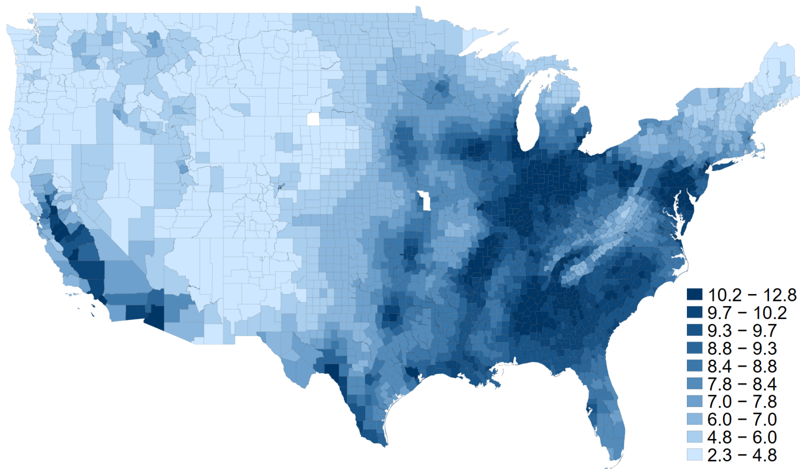
NOTES: *Darker green counties have higher death counts. Deaths come from the NYT database on COVID19 deaths as of April 19, 2020.*

2.2.3 Pathophysiology and epidemiology of COVID19

COVID19 is caused by SARS-COV-2, a betacoronavirus that enters human cells by binding to the angiotension-converting enzyme 2 (ACE2) receptor (Hoffmann *et al.*, 2020). The ACE2 receptor is expressed on type II pneumocytes of human lung tissue, but also in cells of the gut, heart, kidney, and blood vessels (Hamming *et al.*, 2004). It regulates blood pressure, heart function, and ongoing research suggests it may also play a role in regulation of insulin secretion (Li *et al.*, 2020b). Hypertension and pulmonary events are two of the most common causes of mortality among those exposed to high levels of air pollution (Manisalidis *et al.*, 2020). As a result, ACE2 has been suggested as a potential channel between air pollution and mortality (Aztatzi-Aguilar *et al.*, 2015). Importantly, diabetes, pulmonary conditions and hypertension are all frequent comorbidities of COVID19 (CDC, 2020; Shi *et al.*, 2020; Muniyappa and Gubbi, 2020).

Evidence from epidemiological studies indicate that ACE2 production may be up-regulated by chronic exposure to particulate air pollution (Aztatzi-Aguilar *et al.*, 2015). This results in more ACE2 receptors in the lungs of those who have been exposed to chronically high particulate air pollution. A higher number of ACE2 receptors may increase the probability that COVID19 viruses can successfully

FIGURE 2.2: 10 YEAR ANNUAL AVERAGE $PM_{2.5}$ ($\mu g/m^3$)



NOTES: *Darker blue counties have higher annual average levels of $PM_{2.5}$ pollution. Pollution is generally concentrated in urban areas but non-urban areas with high coal fired electricity production, the deep south, midwest and Ohio River Valley, in particular also have high pollution loads in non-urban areas. Data come from Van Donkelaar et al. (2019).*

bind and enter the cells of an individual who has breathed in the virus. While the specific biologic mechanisms of SARS-COV-2, the ACE-2 receptor are still active areas of research, there is a clear channel through which particulate air pollution may increase the probability of infection conditional on breathing in the virus.

When infection with SARS-COV-2 reaches the lungs, it directly damages lung cells and induces an inflammatory immune response (Luks *et al.*, 2020). These processes can lead to impaired surfactant production, fluid accumulation (edema), impaired gas exchange, and alveolar collapse causing cough, shortness of breath, and in severe cases, respiratory failure requiring mechanical ventilation (Force *et al.*, 2012). In severe cases, cardiovascular, metabolic, and immunologic processes appear to be implicated (Mehta *et al.*, 2020). Specifically cardiac injury, leukocytosis (elevated white blood cell count), elevations in other inflammatory markers (IL-2R, IL-6, IL-10, TNA-α), and hyperglycemia were associated with mortality (Li *et al.*, 2020a). This pattern of effects, as well as profound lab abnormalities of inflammatory markers, suggests immunologic involvement and, in some severe cases, cytokine storms (Mehta *et al.*, 2020; Chen *et al.*, 2020a,b).

Both lab and field epidemiological studies have shown that particulate pollution stimulates cytokine generation (van Eeden *et al.*, 2001; Tan *et al.*, 2000; Chen *et al.*, 2018). This has been suggested as the biological mechanism linking particulate air pollution and pulmonary and respiratory mortality (Seaton *et al.*, 1995). Further, there is evidence that elevated expression of cytokines can persist for many years after initial exposure (Gruzieva *et al.*, 2017). Aberrant regulation of cytokine levels in the lungs has been suggested as a potential cause of cytokine storms in response to viral loading (Tisoncik *et al.*, 2012; Cillóniz *et al.*, 2009). As a result, to the extent that chronic exposure to particulate air pollution leads to up-regulation of cytokines in the lungs it may contribute to increased COVID19 mortality by increasing the likelihood that patients experience the extreme inflammatory consequences of a cytokine storm.

2.3 Empirical approach

A challenge in assessing the impact of chronic pollution exposure on the mortality rates from COVID19 is isolating the direct impact of exposure to air pollution. Because areas that suffer consistently higher air pollution may have lower housing costs (Sullivan, 2016; Chay and Greenstone, 2005) residential sorting may lead to systematic differences between populations that are exposed to chronically higher and lower levels of air pollution. These differences may be correlated with COVID19 mortality rates in ways that confound simple cross-sectional estimates of the impact of chronic air pollution on mortality rates (e.g. Wu *et al.* (2020)).

To separately identify the direct impact of chronic exposure to air pollution on mortality rates, we use an instrumental variables approach. We rely on plausibly exogenous changes in the make-up of the power generation fleet in areas upwind and at substantial distance from the county of interest to instrument for changes in the long-run average level of pollution in a given county. Our IV approach is similar to that used in Johnsen *et al.* (2019) to measure the impact of hydraulic fracturing on air pollution and Bishop *et al.* (2018) using variation in power plant emissions to estimate the impacts of particulate air pollution on dementia.

Specifically, we instrument for the long-run average level of $PM_{2.5}$ pollution in a county from 2008 to 2017 with the number of coal fired power plants and the number of gas fired power plants that are opened or retired within a band 100 miles to 150 miles outside of the county of interest. Displacement of coal by natural gas reduces pollution levels in areas downwind of the plants either by removing generating capacity, and the associated pollution, and/or by displacing relatively dirty coal generation with relatively clean natural gas generation (Johnsen *et al.*, 2019). As a result, we use the opening and closing of coal and natural gas plants to predict pollution levels over the subsequent years in downwind counties. We then examine whether areas with larger predicted exposures experience more deaths from COVID19. We estimate the following system of equations:

$$\begin{aligned}\text{Avg. PM}_{2.5}\text{ Conc.}_i = \alpha_j + \psi_1 \text{Plant Retirement}_{ikq} \times \text{Wind Share}_{iq} + \psi_2 \\ \text{New Plants}_{ikq} \times \text{Wind Share}_{iq} + \delta \mathbf{X}_i + \eta \mathbf{P}_i + \epsilon_i\end{aligned} \tag{2.1}$$

$$\text{Log(COVID19 Mortality)}_i = \alpha_j + \beta \text{Avg. } \widehat{\text{PM}_{2.5}} \text{ Conc.}_i + \phi \mathbf{X}_i + \kappa \mathbf{P}_i + \mu_i \tag{2.2}$$

where we instrument for the ten year annual average $PM_{2.5}$ concentration in county i in state j from 2008-2017 with the interaction of the number of coal plants retired or opened and the number of new natural gas plants from 2006-2016 within distance band k of county i in compass quadrant q and the share of time that the wind blew into county i from quadrant q (Wind Share$_{iq}$).[4] In our primary specification k is a band from 100 to 150 miles around county i.[5] Equation 2.2 takes the predicted values of the average $PM_{2.5}$ concentration and relates them to COVID19 deaths. β describes how a change in our predicted level of $PM_{2.5}$ changes the number of COVID19 deaths in county i. $\mathbf{X}_i$ are vectors of county specific controls related to health quality, NAAQS attainment status, contemporaneous air quality, average wind direction, and the spread of COVID19. We select non-NAAQS controls using a cross-validation LASSO procedure where the starting data includes county average mortality rates of common COVID19 comorbidities, 5-year ACS county average socio-demographic variables, and cell phone based measures of the severity of lockdown implementation by county. Our cell phone based measures of lockdown severity come from anonymized data that tracks how often cell phones within

[4]Compass quadrants are Northeast, Southeast, Southwest, and Northwest and correspond to angles of degree 0°-90°, 90.1°-180°, 180.1°-270°, and 270.1°-360°.

[5]In figures B.2 and B.3 we present first and second stage estimates for a range of distance bands.

a county at the same location over time. We mandate that NAAQS non-attainment status for $PM_{2.5}$ in 2019 and the total number of years in our sample the county was in non-attainment, the number of days since the first COVID19 case in the county, the total population and population density of the county, and the number of hospital beds in the county remain in the LASSO. In table 2.1 we detail the variables that are included as controls as well as the set that the LASSO procedure selected over. $\mathbf{P}_{ir}$ is a vector of controls for the number of existing coal and gas plants in the radius r of county i where r is the outer envelope of the k distance band. These include those within county i. We include state fixed effects (α_j) in all regressions.

In order for these to be valid instruments for chronic $PM_{2.5}$ pollution exposure, they must satisfy two conditions: (1) there is a meaningful relationship between the number of new or retiring plants in our distance band and the average annual $PM_{2.5}$ pollution exposure in the county of interest and (2) the opening or closing of a plant in our distance band is plausibly exogenous with regard to correlated determinants of COVID19 mortality within the county of interest. That is, the exclusion restriction must hold. We offer an argument for why each of these conditions are satisfied here and present empirical evidence to support the first condition in the results section.

In the United States at the beginning of our sample, on average, more than 50% of the air pollution mortality in a given state was due to pollution generated out of state and transported by prevailing winds (Dedoussi *et al.*, 2020). Of this cross-state pollution, the power generation sector accounts for the largest share at more than 70%. To highlight the role of emissions in distant counties in determining local air pollution, we show in figure 2.3 the counties from which a marginal increase in emissions contributes the most to $PM_{2.5}$ pollution in Allegheny County in Pennsylvania. Data comes from the AP2 air quality model (Fowlie and Muller, 2019) and shows the top quartile of counties by contribution to pollution levels in Allegheny County.[6] While many of these counties are located in close proximity to Allegheny County, it is clear that the the counties in the band from 100 to 150 miles away, particularly those in the Ohio River Valley, are a meaningful source of pollution.

Over the time period we study, 2005-2018, there has been a decline in both the absolute and relative

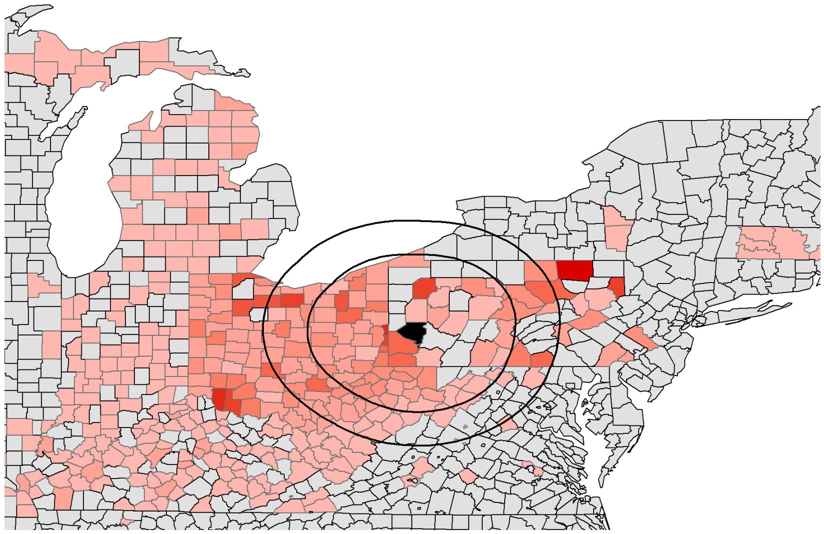
[6]Note that this is not a map of which counties contribute the most pollution to Allegheny County in absolute terms. That is, it is not a map of which counties have many power plants. Rather, it shows the counties in which the emission of 1 additional ton of pollution would generate the most pollution in Allegheny County.

Table 2.1: LASSO variables

Mandated Inclusion	Initial List	Selected List
Total population	Average smoking rate	Average smoking rate
Population density	Mortality rate from pulmonary causes	Mortality rate from pulmonary causes
Hospital beds	Difference from normal activity since activity $<$ 50%	Difference from normal activity since activity $<$ 50%
Days since first case	Days since activity $<$ 75%	Days since activity $<$ 50%
Years in non-attainment, 2006-2017	Days since activity $<$ 50%	Days since activity $<$ 25%
	Days since activity $<$ 25%	Native American population
Attainment status, 2019	White population	Asian population
	Black population	Other race population
	Hispanic population	Median gross rent
	Native American population	Total public transit use
	Asian population	Total without health insurance
	Other race population	
	Population over 55	
	Population over 65	
	Population over 75	
	Population in group quarters	
	Population over 25, less than high school	
	Population over 25, BA degree	
	Per capita income	
	Families in poverty	
	Owner occupied housing	
	Renter occupied housing	
	Median rent	
	Vacant housing units	
	Total male population	
	Mortality rate from ischaemic causes	
	All cause mortality rate	

NOTES: “Mandated Inclusion" are the variables we require the LASSO proceedure to always include. “Initial List" are the variables that the LASSO selects over. We use 10-fold cross-validation LASSO to select controls. “Selected List" are the variables chosen by this procedure that we include in our regressions in addition to the mandated list.

FIGURE 2.3: RANGE OF $PM_{2.5}$ DAMAGES



NOTES*: Darker red counties are those from which a larger share of the marginal unit of pollution ultimately reaches Allegheny county Pennsylvania (marked in black). All colored counties contribute a level of pollution that puts them in the top 25% of contributing counties according to the AP2 model (Fowlie and Muller, 2019). We show counties in the top 25% because damages are continuous and all counties contribute some non-zero level of pollution. The black rings denote 100 and 150 miles from Allegheny county.*

contribution of out-of-state electrical power generation on pollution levels (Dedoussi *et al.*, 2020). This has been driven in part by generation transitioning from old coal to natural gas or more modern coal generation (Burney, 2020; Holland *et al.*, 2018). These changes in the electrical generation fleet, and its consequent impact on pollution levels, have had a meaningful impact on the average level of pollution exposure over our time period, even for counties at a large geographic remove from where the generating plants may be located. Within our sample, the average county has experienced the closure of slightly more than 4 coal fired power plants in the band 100 to 150 miles away from the county (see table 2.2).

The location of power plants is clearly not exogenous. The count of power plants located in the band from 100 to 150 miles from a county would not be a valid instrument because, to the extent that

TABLE 2.2: SUMMARY STATISTICS

	Mean	SD	Min	Max
Outcomes				
Avg. $PM_{2.5}$, 2008-2017	7.96	1.96	2.60	12.78
Total Deaths	11.52	183.29	0	9,708
Total Deaths (w/o NYC or KC)	8.39	57.07	0	1,577
Power Plants				
Retiring Coal, Q1	1.16	2.69	0	19
Retiring Coal, Q2	0.98	2.59	0	26
Retiring Coal, Q3	1.04	2.72	0	24
Retiring Coal, Q4	1.00	2.55	0	32
New Gas, Q1	0.76	2.59	0	31
New Gas, Q2	0.94	3.11	0	31
New Gas, Q3	0.64	2.38	0	40
New Gas, Q4	0.96	3.16	0	35
Wind Direction				
Northeast	33.70	8.92	1	76
Southeast	25.01	9.79	4	75
Southwest	16.93	4.65	1	45
Northwest	24.36	8.56	4	78
CDC Data				
Influenza deaths/100K	23.76	13.14	0.0	135.1
Obesity deaths/100K	0.67	1.20	0.0	10.2
Diabetes deaths/100K	29.89	14.15	0.0	135.9
Avg. smoking rate	25.89	3.71	9.6	39.2
Lockdown controls				
Difference from normal since $<50\%$	26.69	15.42	2.3	221.1
Days since activity $<75\%$	32.94	4.95	0.0	42.0
Days since activity $<50\%$	20.05	7.47	0.0	42.0
Days since activity $<25\%$	4.33	6.97	0.0	35.0

NOTES: "Deaths w/o NYC or KC exclude deaths from New York and Kansas City because the *New York Times* data on COVID19 deaths groups all the counties in these cities together. Power plant statistics are for the distance band 100 to 150 miles. CDC data reports county average death rates from 1999 to 2016. Smoking rates are calculated from 2005 to 2016. Wind direction reports the share of time by county the wind blows from the named compass quadrant. Lockdown data is based on cell phone location data from Couture *et al.* (2020).

they influence pollution levels in distant locations, patterns of residential sorting should reflect their influence. Instead, we rely on the opening and closing of old and new power plants for our identification. We assume that the process of residential sorting happens with a time lag such that pollution levels fall immediately after the closing of a power plant but the process by which neighborhoods change in response to the new levels of pollution occurs slowly over time.[7]

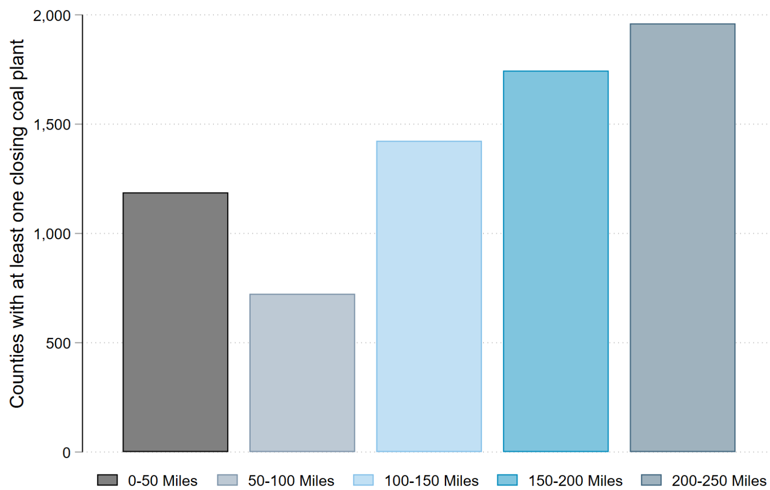
The validity of our instrument thus relies on the assumption that retirement or construction of a plant in our distance band, conditional on the number of plants that were operating at the beginning of our study period, is done for reasons that are exogenous to activity in our county of interest. In particular, we assume that plants are not more likely to be retired or constructed in areas that are contributing relatively more pollution to downwind areas in 2005.[8] We believe this assumption is justified because during our time period there was substantial turnover in the power generating fleet in the United States. More than 300 coal fired generators were retired and more than 600 natural gas generators were brought online.

In many cases the choice to retire coal and construct natural gas generators was driven by the negative shock to natural gas price from the revolution in hydraulic fracturing (Hausman and Kellogg, 2015). Hydraulic fracturing substantially, and unexpectedly, reduced the price of natural gas by around 70% from 2008 to 2012 (Knittel *et al.*, 2015). There was substantial geographic variation in where hydraulic fracturing lowered gas prices for generating firms (Johnsen *et al.*, 2016; Linn and Muehlenbachs, 2018). This variation was driven in large part by the location of the tight shale formations that held the gas which was made accessible by new hydraulic fracturing techniques and was exogenous to the location of existing power plants (EIA, 2016). Johnsen *et al.* (2019) use the variation in a similar IV strategy to show that decreases in the cost of natural gas led to an average 28% decline in coal usage in power generation and a 35% increase in air quality in the areas of highest

[7] Wind direction is commonly used as a source of exogenous variation in studies of air pollution. We do not use it as such here. We include the interaction with wind to increase the precision of our instrument. Because we are interested in long-term effects, and wind patterns are relatively consistent over long time periods, they do not provide us with exogenous variation.

[8] 97.5% of the counties in our sample have at least one coal fired power generation unit within 150 miles as of 2005 and roughly 50% have a coal fired power plant that closes in the band 100 to 150 miles away during that time period (figure 2.4). An additional 27% have had a new natural gas plant open in that same band. We do not find any evidence that pollution levels in 2008 predict the number of plants that close in our sample.

FIGURE 2.4: COUNTIES IN EACH BANDWIDTH



NOTES*: The count of counties that have at least one coal fired unit closing between 2006 and 2016 in each of the distance bandwidths 0-50 miles, 50-100 miles, 100-150 miles, 150-200 miles, and 200-250 miles. Closing plants within a county are grouped in the 0-50 miles bin.*

displacement.

An important assumption for the validity of our instrument is that there is a mismatch in the timing between when counties experience lower pollution levels and the attainment of a new equilibrium in the sorting process that determines the socioeconomic make-up of the county. Existing research has shown that areas become wealthier, whiter, and contain a greater share of home-owners relative to renters after the reduction in air pollution that comes from the closure of coal-fired power plants (Sullivan, 2016). This occurs as the reduction in air pollution makes the neighborhood more attractive and thus raises housing costs. However, this process takes time. Reforms of the power sector in California in 2000 substantially reduced air pollution by 2001 but these changes took nearly ten years, until 2009, to fully manifest as changes in the make-up of LA neighborhoods (Sullivan, 2016). This delay occurred despite the closure leading to the cleaner air occurring in LA and being reasonably salient to residents of LA as opposed to the variation we rely on that occurs hundreds of miles away. Intuitively, such a delay makes sense; it takes time for households to recognize the change in pollution and for real estate markets to adjust. Our identification relies on the condition that power plant closures lead to near immediate changes in pollution levels but neighborhood demographics change slowly in response.

One might reasonably be concerned that hydraulic fracturing has economic consequences other than shifting the mix between natural gas and coal in power generation. If these economic consequences spill over into neighboring counties it might confound our instrument. In particular, if hydraulic fracturing raised wages or incomes, or tax revenue, in neighboring counties, COVID19 mortality might be reduced through a channel other than air pollution that is correlated with our pollution instrument. Feyrer *et al.* (2017) examines exactly the question of spillovers and finds that hydraulic fracturing does increase wages, business income, and government royalties. However, these effects are concentrated in areas relatively near to the county in which hydraulic fracturing occurs. The increase in wages declines to near zero by 100 miles away from the site of hydraulic fracturing (see figure B.1). Further, the additional royalty revenue dissipates within approximately two years of the drilling of the well. To the extent that the largest drilling shock occurs from 2008 to 2012, that suggests the additional government revenue may not be affecting COVID19 mortality in 2020.

2.4 Data

2.4.1 Air pollution data

There are two main sources of air pollution data for the United States - monitor data from the EPA's system of ground based monitors and remotely sensed data that measures air pollution by processing satellite imagery to detect certain types of particles. There are advantages and disadvantages to each type of data. One major drawback of monitor based data is its incomplete geographic coverage that can lead to substantial underestimation of the true levels of air pollution across counties (Sullivan and Krupnick, 2018). This is exacerbated by the fact that regulators may strategically locate monitors to avoid detecting the highest levels of pollution in their jurisdiction (Grainger *et al.*, 2016).

To avoid the potential bias from ground-based monitor data, we use remotely sensed air pollution data. Specifically, we use pollution data from the North American Regional analysis conducted by Van Donkelaar *et al.* (2019). This remotely sensed data combines observed data on aerosol optical depth

(AOD) from NASA MODIS, MISR, and SeaWIFS satellite systems with a physics model of chemical transport to provide data at a fine resolution over long time periods. Remotely sensed pollution data is generally believed to provide a reasonable best estimate of pollution levels (Auffhammer *et al.*, 2013) and is widely used in both environmental economics (Schlenker and Lobell, 2010; Hsiang, 2016).

We focus on $PM_{2.5}$ because it has the most deleterious consequences for human health and collect data on $PM_{2.5}$ levels from 2006 to 2017 across the continental United States.[9] The satellite data is reported on a $0.01°$x$0.01°$ ($\approx$4mi x 4 mi) grid. We assign grid points to counties and calculate a county average annual exposure by averaging across all grid points in the county. For each county in the United States we then calculate the ten year average level of each of the pollutants. We end our average in 2017 because recent work from the Po river valley (Setti *et al.*, 2020) and from China (Yongjian *et al.*, 2020) suggests high contemporaneous levels of pollution may facilitate the spread of the virus by allowing the virus to bind to particulate matter in the air. Ending our average in 2017 avoids confounding our estimates with contemporaneous effects.[10]

2.4.2 COVID19 mortality data

Our COVID19 mortality data comes from the database on cases that has been assembled by the *New York Times*. Their database provides the current cumulative case and death count for all counties in the United States that have reported at least one case. From this data we calculate the date the first case was announced for each county and the date of the first death. [11]

In order to include cases in New York City and Kansas City we group all the counties that comprise each metropolitan area together in out analysis. This is because the NYT reports aggregate cases for both cities. For the pollution data we calculate averages across all five counties. For the Census ACS data (discussed below) we aggregate as total or median (as appropriate) across each the five counties. We do this because the *New York Times* reports a single aggregate case and death count for all the

[9]We also collect data on NH_4, SO_3, and NO_4 but as we show in table B.1 these pollutants do not appear to have a strong impact on COVID19 mortality.

[10]We use nearly contemporaneous data on NAAQS attainment status to control for short-term effects.

[11]The data can be accessed here: github.com/nytimes/COVID19-data

counties in New York City. We do the same for Kansas City for the same reason. We group all four counties that touch on Kansas City together and then add the cases and deaths that the New York Times reports for Kansas City to that total.

2.4.3 Other data

We combine our air pollution and mortality data with a number of other data sets in order to control for additional covariates of mortality.

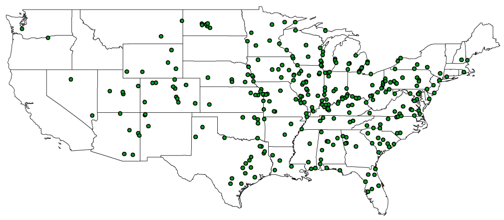
EIA power plant data The Energy Information Agency (EIA) requires power generation facilities in the U.S. with greater than 1MW nameplate capacity to provide information about the annual operation of their facilities on a form 860. This information includes the location, primary fuel types, gross load, operating hours, and retirement or planned retirement ages of the generators. We use the database of EIA-860 data assembled in Burney (2020) for coal and natural gas plants retired or opened from 2005 to 2016. This provides us with the latitude and longitude coordinates of each coal and natural gas plant operating in the U.S. during that time period as well as the opening or closing dates for plants that did not operate continuously for the whole time period. We map this data and count the number of new, retiring, and existing plants of each type within various distance bands from every county in the United States.[12]

ACS data In order to control as best as possible for the range of socioeconomic characteristics that may be correlated with chronic particulate pollution exposure, we download the comprehensive 5-year average ACS report from the U.S. Census at the county level for 2018. We match this to the COVID19 mortality data by FIPS code, subject to the changes described above.

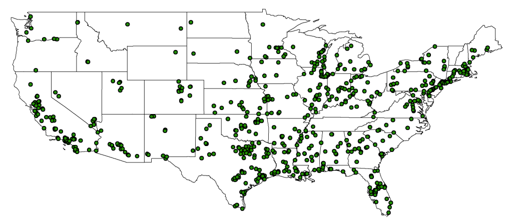
DEX data As the COVID19 pandemic has worsened across the United States, cities, counties, and states have responded with various containment measures and lockdowns in an attempt to limit the spread of the virus. In order to control for these measures we have included data measuring the average interaction of anonymized cell phone data at the county level from Couture *et al.* (2020). To measure interactions Couture *et al.* (2020) count the number of distinct devices that visited any commercial

[12] We map the location of all plants in the figure 2.5.

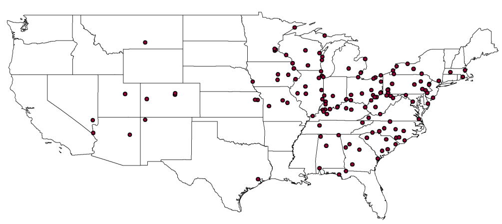
FIGURE 2.5: LOCATION OF EXISTING, NEW, AND RETIRING PLANTS, 2006-2016



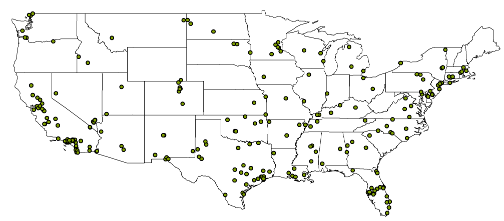
EXISTING COAL PLANTS



EXISTING NATURAL GAS PLANTS



RETIRED COAL PLANTS



NEW NATURAL GAS PLANTS

establishments in a single day. They report the county average across all cell phones in the county on that day. We use average over time to determine when activity in a given county falls below seventy five percent, fifty percent, and twenty five percent of the average level from the beginning of the sample (January 20, 2020) to March 1, 2020. For each county we then construct variables that count the days between crossing these thresholds and the most recent mortality report. We also calculate the difference in average activity levels after activity falls below the 50% threshold and the baseline activity level to account for the fact that lockdowns may become less severe or less obeyed over time.

CDC data We download the count of deaths and crude death rates from 1999-2016 for a range of comorbidities from the CDC WONDER system for all counties in our sample. The WONDER system reports mortality statistics for a range of causes of death across U.S. counties. The data is based on death certificates. We collect data on all cause mortality, pulmonary heart disease, ischaemic heart disease, hypertensive diseases, diabetes, obesity, influenza, and acute respiratory diseases.

EPA Greenbook data The EPA designates counties' non-attainment if they are not in compliance with the National Air Quality Standards (NAAQS) based on data from the EPA monitoring network. We collect data from the EPA Greenbook reporting historic compliance with the NAAQS standards on whether a county is in non-attainment with the $PM_{2.5}$ standard in 2019 as well as the number of years from 2006 to 2017 that the county was in non-attainment.

Wind direction data We collect data on the direction that the wind blows into a county on a daily level from 2005 to 2016 from Deryugina *et al.* (2019). We assign a quadrant to each angle recorded in this data and calculate the share of days over this time period the wind blows into a county from each quadrant. For counties in our sample that are not in the data collected by Deryugina *et al.* (2019) we use the average across available counties in the same state.

2.5 Results

We begin briefly with a discussion of the cross-sectional relationship between exposure to chronic particulate air pollution and COVID19 mortality. In table 2.3 we show the results of the simple cross-sectional, state-fixed effects regression of the log of one plus county deaths on the ten year annual average level of $PM_{2.5}$ pollution. In our preferred specification, in which we include controls selected by LASSO for the days since the first case was detected, a range of socioeconomic characteristics from the ACS, and the number of days since economic activity declined substantially as a result of lockdowns, a one unit increase in annual average $PM_{2.5}$ levels is associated with an increase of roughly 5% in the number of COVID19 deaths in a county.[13]

Our estimates differ substantially from those published in Wu *et al.* (2020) who find increases of 15% in mortality from COVID19 due to a $1\mu g/m^3$ increase in the ten year annual average of $PM_{2.5}$. There are several potential explanations for this discrepancy. The first is that we use more recent data on deaths than used in Wu *et al.* (2020). As the COVID19 epidemic grows, death counts are changing rapidly. The three week difference in the data used may explain the difference in results, particularly if the places that had high early death counts have higher-than-average levels of air pollution. We also control for more county features than in Wu *et al.* (2020). In particular we have richer controls for the change in activity after the advent of COVID19, which may be correlated with deaths and air pollution, and for contemporaneous air pollution.

[13]Our ACS controls include controls for the age and racial profile of a county, the population and population density, number of renters, common methods of commuting to work, health insurance status, and average incomes.

TABLE 2.3: CHRONIC $PM_{2.5}$ POLLUTION AND COVID19 MORTALITY

	LN1+	IHS
Avg. $PM_{2.5}$, 2008-2017	0.048***	0.063***
	(0.014)	(0.017)
N	3,096	3,096
LASSO Controls	X	X
State FE	X	X

NOTES: The outcome in column 1 is the log of deaths+1 in a county as of April 19^{th}. In column 2 it is the inverse hyperbolic sine. Coefficients should be interpreted as the percent change in deaths for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

2.5.1 First-stage IV results

We present results from our first stage estimates in table 2.4. To ease readability we present the total effect of plant retirements and openings in each quadrant accounting for wind share.[14] As hypothesized, the retirement of coal fired power plants reduces average $PM_{2.5}$ levels over 2008 to 2017. Our first stage results imply that the average county in our sample experienced a reduction of ten year annual average $PM_{2.5}$ levels of approximately 0.15 $\mu g/m^3$ (≈10% of 1 SD) due to changes in the power generation mix relative to the counter-factual of no change in the generation mix. For comparison's sake, scaling the estimates in Johnsen *et al.* (2019) to our distance bandwidths suggests that the closure of a large coal fired power plant would reduce ten year annual average $PM_{2.5}$ levels by 0.08 $\mu g/m^3$.

More surprisingly, the opening of new gas fired power plants also reduces average $PM_{2.5}$ levels in some cases. In a setting in which power generators are choosing how to meet power demand by utilizing a mix of generating sources, some dirtier than others, it is reasonable that adding more, relatively clean, natural gas sources will reduce overall pollution. Whether new gas generation capacity increases or reduces pollution will depend on the generation mix prevailing upwind of a given county

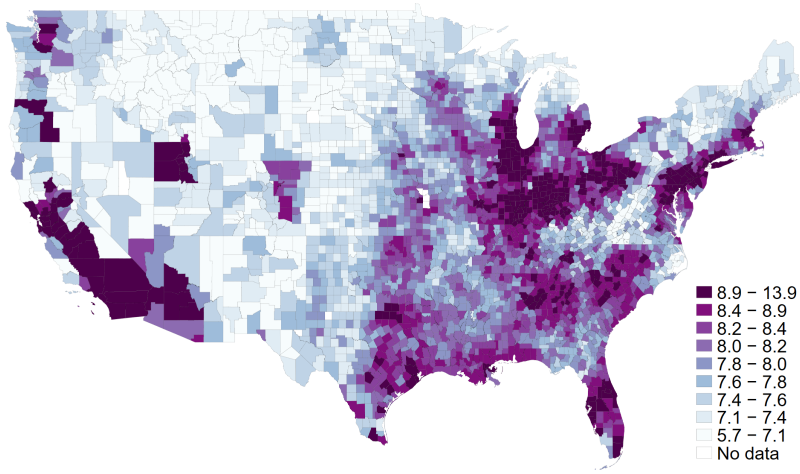
[14] Our fist stage equation estimates the impact of a plant retirement (ψ_1^A) and the interaction with wind share (ψ_1^B). The total effect of a plant closure in quadrant q is ψ_1^A+ψ_1^B × Wind Share$_{qi}$. We evaluate this total effect at the mean wind share for each quadrant and present these results in table 2.4.

TABLE 2.4: FIRST-STAGE RESULTS

	(1)	(2)
Retiring Coal, Q1	-0.041***	-0.041***
	(0.008)	(0.008)
Retiring Coal, Q2	-0.028**	-0.023*
	(0.012)	(0.012)
Retiring Coal, Q3	-0.045***	-0.044***
	(0.007)	(0.007)
Retiring Coal, Q4	-0.006	-0.008
	(0.009)	(0.008)
New Gas, Q1	0.005	0.003
	(0.006)	(0.006)
New Gas, Q2	-0.033***	-0.031***
	(0.006)	(0.006)
New Gas, Q3	0.029***	0.029***
	(0.010)	(0.009)
New Gas, Q4	-0.021***	-0.017***
	(0.006)	(0.005)
F-statistic	119.38	114.28
State FE	X	X
LASSO Controls	X	X
NAAQS Controls		X

NOTES: The outcome in all columns is the the percent change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* p<.10 ** p<.05 *** p<.01).

FIGURE 2.6: PREDICTED 10 YEAR ANNUAL AVERAGE $PM_{2.5}$ ($\mu g/m^3$)



NOTES: *Darker purple counties have higher predicted annual average levels of pollution. Pollution levels are predicted using the opening and closing of coal and natural gas power plants 100 to 150 miles away from a given county.*

prior to the new plant. Consistent with this idea, we observe that the gross load (a measure of plant utilization) of older coal fired power plants declines as new natural gas plants are brought online within a given area and there is a substitution of operating time from older coal fired power to newer gas fired power. In California this likely means new gas will increase local pollution because there is little coal to displace. In the Ohio River Valley, one of the most polluted areas in the country, new gas likely displaces coal and reduces pollution. The relative impacts of coal and natural gas on air pollution, as well as the signs we estimate, are consistent with the results in Johnsen *et al.* (2016). The F-statistics on the instruments in our first stage are substantially greater than 10, suggesting we do not have a weak instrument problem (Stock and Yogo, 2002).

2.5.2 IV results

Our IV estimates reported in table 2.5 are inflated relative to our OLS results. In our preferred specification, controlling for the LASSO controls as well as NAAQS status, we find that a one unit increase in ten year annual average $PM_{2.5}$ increases COVID19 deaths by approximately 9.5%. This implies a 1% increase in ten year annual average $PM_{2.5}$ increases COVID19 deaths by ≈0.75%.[15] The inflation of our IV results relative to our OLS results is relatively moderate. In our preferred specification the IV estimates are 100% larger than the OLS estimates. This is consistent with the differences in OLS and IV estimates in existing work using a similar instrumental strategy (Johnsen *et al.*, 2019).[16] Despite this inflation, our IV estimates are substantially smaller than the existing cross-sectional correlations.

TABLE 2.5: INSTRUMENTAL VARIABLE RESULTS

	LN1+	LN1+	IHS	IHS
Avg. $PM_{2.5}$, 2008-2017	0.107**	0.096**	0.133**	0.121**
	(0.047)	(0.049)	(0.058)	(0.059)
N	3,097	3,097	3,097	3,097
State FE	X	X	X	X
LASSO Controls	X	X	X	X
NAAQS Controls		X		X

NOTES: The outcome in columns 1-2 is the log of deaths+1 in a county as of April 19^{th}. In column 3-5 it is the inverse hyperbolic sine. Coefficients should be interpreted as the percent change in deaths for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

2.5.3 IV robustness checks

Our IV estimates are robust to specification choice. We present results of the LIML and GMM model in place of the 2SLS approach in tables B.2 and B.3 (Pischke, 2018). In both cases estimates are slightly larger than in the 2SLS case but are consistent with the 2SLS approach. In columns 3-6 of table 2.5 we present 2SLS results using the inverse hyperbolic sine instead of log 1+deaths. Our results using the IHS imply an increase in deaths of 1% for every 1% increase in ten year annual average $PM_{2.5}$ pollution. We also present results in table B.4 excluding NYC and KC because of the unusual

[15] In our data a one unit increase in $PM_{2.5}$ represents a ≈12% increase in 10 year annual average $PM_{2.5}$ levels from the mean.

[16] The inflation of their IV to OLS estimates ranges from 19% to 72% depending on the specification.

way in which cases are aggregated in those cities. Our estimates are somewhat smaller but similar and remain significant.

To test the robustness of our distance band choices we show the result of different 50 mile bandwidths in figures B.3 and B.4. The point estimates are stable across the first three distance bands (0-50 miles, 50-100 miles, and 100-150 miles) and begin to decline in the fourth (150-200 miles) and fifth (200-250 miles). This is consistent with the influence of power plants on pollutant levels waning beyond 150 miles as suggested by the results in figure B.2 testing the change in the first stage estimates for different distance bandwidths. Our results suggest that the influence of power plants on particulate pollution levels begins to decline steeply after 150 miles. The impacts in both bands from 150-200 miles and 200-250 are close to and not statistically different from zero.

2.5.4 Effect on COVID19 case counts

In addition to the impact on deaths we examine the impact of chronic exposure to particulate air pollution on the number of COVID19 cases. Recall the current understanding of the pathophysiology of COVID19 suggests particulate air pollution may increase both mortality and transmission, thus increasing cases. Our analysis of the impact on cases comes with the significant caveat that COVID19 case numbers appear to be substantially under-counted in the United States due to shortages of testing materials.[17] Despite this the estimates in table 2.6 suggest that particulate air pollution may also increase case counts. We estimate a roughly 13% increase in cases for a $1\mu g/m^3$ increase in the ten year annual average level of $PM_{2.5}$.

2.6 Conclusion

Understanding role of exposure to chronic particulate air pollution in modulating the impact of the COVID19 pandemic is important for understanding which areas may expect to suffer higher hospitalization and mortality rates from the pandemic. Early estimates of the role of chronic exposure to particulate air pollution may have overestimated its impact. Our estimates suggest that particulate

[17] Deaths may also be under-counted but it is generally thought that death counts are more accurate than case counts.

TABLE 2.6: INSTRUMENTAL VARIABLE RESULTS WITH CASES

	LN1+	LN1+	IHS	IHS
Avg. $PM_{2.5}$, 2008-2017	0.140^{**}	0.126^{**}	0.164^{**}	0.151^{**}
	(0.062)	(0.062)	(0.067)	(0.067)
N	3,097	3,097	3,097	3,097
State FE	X	X	X	X
LASSO Controls	X	X	X	X
NAAQS Controls		X		X

NOTES: The outcome in columns 1-2 is the log of cases+1 in a county as of April 19^{th}. In column 3-5 it is the inverse hyperbolic sine. Coefficients should be interpreted as the percent change in cases for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* p<.10 ** p<.05 *** p<.01).

air pollution may in fact play a causal role in increasing mortality from the virus but our estimates suggest that its impact is roughly 30% smaller than the existing cross-sectional estimates.

We estimate that a 1 $\mu g/m^3$ increase in the ten year annual average $PM_{2.5}$ level, which represents a 12% increase from the mean, increases mortality from COVID19 by 9.5%. There are multiple physiological pathways through which particulate air pollution may increase mortality. Chronic particulate air pollution exposure may up-regulate the production of protein receptors in the lungs that the virus binds to in order to enter the body. This may increase transmission of the virus and we find evidence that suggests particulate air pollution may increase case counts. Chronic particulate air pollution may also increase the production of cytokines in the lungs and increase the probability of patients experiencing cytokine storms, believed to be a source of COVID19 mortality.

More research on the relationship between COVID19 and air pollution is necessary. In particular, understanding the role of short-term air pollution in modulating the spread, and potentially the mortality, of the virus is critical. The relationship between short-term air pollution exposure may be important in guiding decisions about when to lift lock-downs and managing social interactions after they are lifted. Areas with higher levels of pollution may need to maintain stricter social distancing than areas with lower levels.

This work also underlines the important health consequences of air pollution in general. While pandemics are rare events, our evidence suggests that some areas experienced substantial excess deaths as a result of their existing exposure to air pollution. Minimizing the costs of future pandemics may be an important consideration for policy-makers choosing how to regulate air pollution.

Chapter 3

Learning is Inhibited by Heat Exposure, Both Internationally and Within the U.S.[1]

3.1 Introduction

Both across and within countries, people living in hotter climates complete less formal schooling, score lower on standardized tests, and exhibit worse economic outcomes than those living in cooler climates (Montesquieu, 1750; Dell *et al.*, 2012; Park *et al.*, 2020). Such associations are important given the growing role of cognitive skill in income mobility and economic growth (Goldin and Katz, 2009; Chetty *et al.*, 2014; David, 2014; Hanushek and Woessmann, 2016), and because of current and expected changes to the earth's climate, which appear to influence macroeconomic growth (Burke and Emerick, 2016). Whether and how climatic factors causally affect human capital development, however, remains debated, in part because so many other institutional and economic factors are correlated with a warmer historical climate. Some argue that initial conditions during colonization influenced the institutions created in hotter, more disease-prone climates, leading to lower levels of human capital today (Acemoglu *et al.*, 2001). Others emphasize the role of correlated impediments to

[1] Co-authored with R. Jisung Park (UCLA) and Joshua Goodman (Boston University)

agricultural productivity (Schlenker *et al.*, 2006) or child nutrition and health (Currie, 2009), which may in turn change the incentive to pursue schooling (Maccini and Yang, 2009; Shah and Steinberg, 2017).

We propose a more direct mechanism that may operate alongside institutional, agricultural, or other factors. Across a range of laboratory and field environments, temperature has been shown to affect working memory, stamina, and cognitive performance (Seppanen *et al.*, 2006; Park, forthcoming), and to lead individuals to reduce time spent engaging in labor activities (Graff Zivin and Neidell, 2014). This suggests that, in addition to the channels above, heat may directly affect students' capacity to learn or teachers' ability and willingness to teach. Given vast international differences in thermal conditions experienced by students (Table 3.1), even small marginal effects of heat on learning could result in large educational disparities over time. Students in Indonesia and Thailand, for instance, experience over 200 days above 80°F per school year, compared to approximately 40 such days in the United States and South Korea. Causal estimates of the returns to schooling suggest that small changes in educational achievement can result in persistent differences in lifetime earnings potential (Acemoglu and Autor, 2011). There is, however, limited evidence on how heat exposure affects the rate of learning and human capital accumulation in the context of formal schooling (Graff Zivin *et al.*, 2017; Park *et al.*, 2020).

We provide evidence that heat exposure during learning periods directly impacts human capital accumulation, suggesting another channel through which climate is linked to macroeconomic development. To do so, we provide two sets of analyses, each using quasi-experimental research designs and incorporating region-specific academic calendars to measure temperature shocks that occur on school days preceding cognitive testing. The empirical designs focus on heat exposure during the school year - as opposed to momentary reductions in cognitive performance due to temperature on the day of assessment - and exploit year-to-year variation in weather within a given region to isolate the causal impact of hotter school years on learning.

The first analysis uses test score data from 58 developed and developing countries participating in the Programme for International Student Assessment (PISA) between 2000 and 2015. PISA's tests are designed to measure formal learning in math, reading, and science in nationally representative

TABLE 3.1: HEAT EXPOSURE IN SELECTED PISA COUNTRIES

	School Days> 80°	Per Capita Income (USD)	Avg. PISA Score
Indonesia	240	2,180	-1.17
Thailand	204	3,937	-0.76
Brazil	119	7,043	-1.08
Mexico	145	8,160	-0.87
Vietnam	114	1,894	0.09
Israel	80	27,759	-0.40
United States	44	46,247	-0.08
South Korea	36	19,467	0.35
Spain	24	25,224	-0.14
Turkey	26	8,899	-0.59
France	12	34,616	-0.01
Netherlands	7	45,164	0.18

NOTES: The school day measures report average annual number of school days over 80° experienced by each country during our sample period from 1995-2015. Per capita income reports the average per capita income in constant USD over the same time period using data from the World Bank. Normalized PISA Score reports the average normalized overall PISA score within each country over our sample period.

samples of 15-year-olds. We find compelling evidence that students in school during hotter periods score worse on these exams than their peers in the same country who are schooled in cooler periods. The effects of years with more hot days (above 80°F) on subsequent performance persists even when adding controls for changes in economic conditions (e.g. per capita income) and possible spurious correlation between regional time trends in warming and educational performance. To isolate the causal impact of heat exposure on learning, we link within-country temperature fluctuations over time to within-country fluctuations in test scores, controlling for country- and time-varying confounds.

Exploiting variation in the timing of hot days within a given calendar year, we provide suggestive evidence on the potential mechanisms at play. Heat on school days prior to PISA exams lowers test scores while heat on non-school days (e.g. weekends, summer vacation) has little effect, consistent with our hypothesis that heat directly interferes with learning time. In addition, including controls for potential correlated shocks to agricultural yields does not affect the magnitude or significance of these findings. Specifically, the effects are robust to controlling for hot days during region-specific rice growing seasons as well as time-varying, country-level measures of agricultural employment, suggesting that the effects of hot temperature are not driven solely by correlated shocks to nutrition or time reallocation decisions in response to correlated changes in economic incentives to pursue schooling.

Even with a rich set of controls, the range of countries in our data implies these effects could be driven by other correlated mechanisms noted above, particularly in lower-income, agrarian economies. The second analysis therefore focuses on the United States, a highly developed, non-agrarian setting where nutrition and agricultural income-related channels seem less likely to be empirically first-order in explaining the impact of heat on achievement. We use district-level annual math and English Language Arts (ELA) test scores from over 12,000 U.S. school districts, from the Stanford Education Data Archive (SEDA). These tests are mandatory components of school accountability systems, so that the sample of test-takers represents the near-universe of American students. The tests are deliberately aligned with school curricula to measure learning that is meant to occur during formal schooling. Similarly to the international data, we link within-district temperature fluctuations over time to within-district fluctuations in test scores to isolate the causal effect of hotter temperature during the school year.

We find that US students in school during hotter years score worse than peers in the same district schooled in cooler periods. Consistent with the international evidence and the hypothesis that heat interferes with learning, we find that heat on school days entirely drives our results. These results are robust to the inclusion of controls for district-level changes in school funding and demographic composition, potential spurious correlation between regional warming patterns and trends in educational achievement, and controls for exam-day temperature.

Across both sets of analyses, we find that the marginal damage associated with hotter temperature appears to be larger for lower income populations, consistent with previous work on climate adaptation (Carleton *et al.*, 2018). These results suggests that the effects of hot temperature may be regressive not only across but also within countries, consistent with recent work (Hsiang *et al.*, 2018; Park *et al.*, 2020). In the U.S., heat's effects appear to be larger for racial minorities and students in lower income school districts, who likely have less access to potentially compensatory resources. We also present novel evidence suggesting that the effect of heat exposure during learning periods on achievement is larger for younger students. The effect of heat on children may be more pronounced if children rely more heavily than adults on well-functioning institutions to enable effective avoidance behaviors or carry out necessary protective investments. These and other reasons suggest that children may be more susceptible to hyperthermia and heat-exhaustion (Rowland, 2008), but so far there has been little evidence regarding the differential impact of heat exposure on learning across age groups.

We note three observations about these analyses. First, they study the impact of heat on learning, rather than momentary reductions in cognition that may arise from temperature stress. Existing evidence suggest many factors including temperature (Graff Zivin *et al.*, 2017; Park *et al.*, 2018), air pollution (Ebenstein *et al.*, 2016), sleep deprivation (Alhola and Polo-Kantola, 2007), and attentional capture (Mani *et al.*, 2013) can affect short-run cognition. The mechanism studied here does not operate through such short-term reductions in cognition during test-taking or in the immediate lead-up to test taking, and controls for the possibility of correlation between heat exposure during learning periods and hot temperature during a subsequent exam. The outcome measures are standardized assessments designed to capture cumulative learning throughout formal schooling, as opposed to tests of raw intelligence or cognitive capacity that are highly sensitive to test-taking conditions, in contrast to prior studies (Graff Zivin *et al.*, 2017).

Second, these results encompass students in both the developing and developed world, presumably with varying levels of adaptation investment. Previous studies find that the effect of climatic shocks on health and economic outcomes vary substantially by income or previous exposure (Dell *et al.*, 2012; Burke *et al.*, 2015; Carleton *et al.*, 2018), and that investments such as air conditioning may be effective at mitigating heat-related impacts (Barreca *et al.*, 2016). Given vast differences in the rate of air conditioning across countries, and notably between the US and most other countries, it is important to assess the external validity of existing US-based findings (Graff Zivin *et al.*, 2017; Park *et al.*, 2020). Recent survey evidence suggests that, whereas 90 percent of US households have some form of air conditioning, only 73 percent, 19 percent and 13 percent of households in Australia, Sweden and Mexico respectively have air conditioning (Davis and Gertler, 2015; Randazzo *et al.*, 2020). This study suggests that the smaller macro-level effects of temperature documented in developed economies(Burke *et al.*, 2015) may mask substantial heterogeneity within these countries .

Third, we suggest a seemingly universal physiological channel through which heat affects human capital accumulation, in contrast to an older and racially charged literature arguing that the association between climate and human capital is driven by genetic or cultural factors. Such literature claimed that those living in tropical countries were genetically and culturally “lazy" or otherwise disinclined to engage in cognitively intensive activities (Gilfillan, 1920; Huntington, 1922). The unfortunate implications of this work may have inhibited discussion of a simpler and more policy-relevant explanation for the observed associations between heat and human capital. We suggest that the universal physiological burden of heat reduces students' capacity to learn and teachers' capacity to teach, independent of intelligence or disposition. Hotter climates may thus interfere with economic development by reducing the human capital stock of nations, which implies that investments aimed at protecting students from heat exposure may confer important economic benefits, particularly in hotter, poorer countries.

3.2 Results

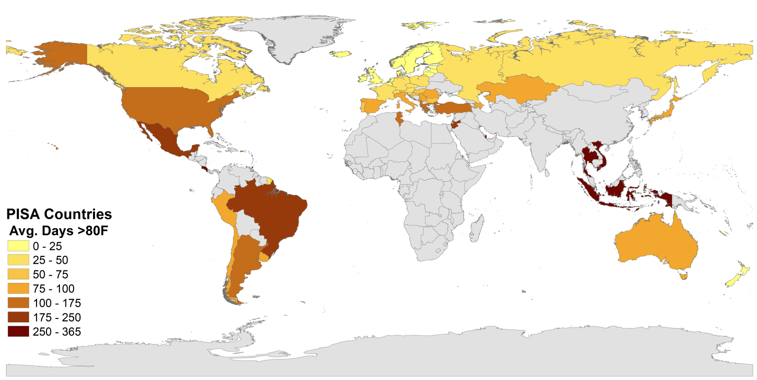
3.2.1 International analysis

The first analysis explores the relationship between heat exposure and standardized performance on the Programme for International Standardized Assessments (PISA). The sample comprises exam records from 58 countries who participated in PISA, which is administered by the OECD and provide internationally harmonized exams to nationally representative samples of 15 year olds every 3 years since 2000.

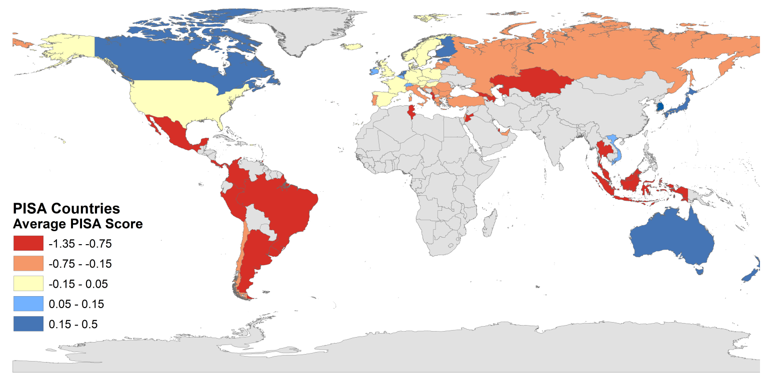
Our sample spans a wide range of incomes and average climates, including poor tropical countries such as Vietnam and Thailand as well as many richer temperate countries such as South Korea, France, and New Zealand. Average per capita income across the countries in our sample is $25,962 in current U.S. dollars (Table 3.2), with some as low as $662 per capita (Kyrgyz Republic) and some as high as $80,857 (Luxembourg). The countries in our sample are plotted in Figure 3.1, and represent approximately 144 million 15-19 year olds across the participating countries.

Our empirical design leverages random variation in temperature within a given country over

FIGURE 3.1: TEMPERATURE EXPOSURE AND PISA SCORES.



(A) *Temperatures*



(B) *PISA Scores*

NOTES: *In both figures the colored countries are the countries that take the PISA exam. Panel (a) shows the average number of days with a maximum temperature above 80°F over our sample period, 1995-2015. Panel (b) shows the average normalized PISA score in terms of standard deviations across all subjects and years that countries report scores during our sample period. There are 67 countries reporting at least one PISA score in our sample period.*

TABLE 3.2: SUMMARY STATISTICS FOR PISA SAMPLE

	Full Sample			Richer Sample			Poorer Sample		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD
Hot school days	58	114	177	24	67	125	34	156	205
Normalized PISA Score	58	-0.28	0.49	24	0.02	0.27	34	-0.54	0.48
Employment in Ag.	58	10	11	24	3	2	34	16	12
Per capita Income	58	25,962	21,704	24	44,320	18,076	34	9,807	6,311
Population, 15-19	58	3.0	5.2	24	2.4	4.5	34	3.7	5.7

NOTES: School days measures the total number of days over 80°F in the previous three years. Population is reported in millions. Employment in Agriculture is as a percentage of total employment. Employment, Income and Population data all come from the World Bank World Development Indicators data. We calculate hot school days based on NOAA's GHCN data. PISA scores come from the NCES and we standardize them as described in the supplementary materials. Rich countries are defined as countries whose per capita income in 1995 was greater than $14,000, the average per capita income in 1995 in countries in our sample. Poor countries are those countries with per capita income less than $14,000 in 1995.

multiple years. While unobserved determinants of student achievement may be correlated with average climate in the cross-section, year-to-year fluctuations in temperature within a country are plausibly random, particularly when adjusting for correlated global or regional trends in warming and development. Our strategy compares deviations from country-specific averages in PISA performance with deviations from country-specific average temperature, controlling flexibly for other time-varying factors including precipitation and share of labor force in agriculture. We focus on the impact of the number of days with temperatures above 80°F, noting that previous studies of heat on cognitive performance and other behavioral outcomes find adverse impacts beginning around 80°F (Graff Zivin *et al.*, 2017; Park *et al.*, 2018, 2020).

We find that hotter temperatures in years leading up to the PISA exam negatively impact student performance. Each additional day above 80°F during the 3 years preceding an exam lower scores by 0.18 percent of a standard deviation (p= 0.007, 95% CI = [-0.22, -0.04], Figure 3.2). We measure hot days over 3 years to maintain consistency with the periodicity of the PISA exams. A one standard deviation increase in hot days conditional on country and year fixed effects amounts to 14 school days. Cold days have statistically insignificant impacts on performance ($\beta = 0.07$, $p = 0.517$, 95% CI = [-0.14, 0.28]). These results are robust to the inclusion of continent-specific temperature trends, which suggests that they are not driven by spurious correlation between regional warming patterns

TABLE 3.3: PRIOR HEAT AND NATIONAL ACHIEVEMENT - PISA

	PISA Scores	PISA Scores	Richer Sample	Poorer Sample
(A) All Days				
Total hot days	-0.125	-0.126	-0.024	-0.143
	(0.051)	(0.045)	(0.071)	(0.038)
	[0.017]	[0.007]	[0.733]	[0.001]
	{-0.23,-0.02}	{-0.22,-0.04}	{-0.17,0.12}	{-0.22,-0.07}
N	282	282	132	150
(B) School Days				
Hot school days	-0.200	-0.224	-0.099	-0.256
	(0.077)	(0.070)	(0.143)	(0.053)
	[0.012]	[0.002]	[0.493]	[0.000]
	{-0.36,-0.04}	{-0.36,-0.08}	{-0.39,0.20}	{-0.36,-0.15}
Hot non-school days	0.037	0.038	0.127	0.088
	(0.099)	(0.091)	(0.146)	(0.092)
	[0.710]	[0.676]	[0.394]	[0.342]
	{-0.16,0.24}	{-0.14,0.22}	{-0.18,0.43}	{-0.10,0.28}
N	282	282	132	150
Continent Specific Linear Trend		X	X	X
Precipitation controls		X	X	X
Additional controls		X	X	X

NOTES: Heteroskedasticity robust standard errors clustered by country are in parentheses. *p*-values reported in brackets and 95% confidence intervals in curly brackets. Temperature is measured with the daily maximum temperature on days in each country in the three years prior to the year the exam was taken. All columns include country and year fixed effects, a continent-specific linear trend, and controls for temperature in the year of the exam and precipitation in both the year of the exam and the three years proceeding the exam. The outcome is the average across PISA scores available in a given year standardized according to PISA's methodology as described in the Materials and Methods. All regressions weight countries by the number of 15-19 year olds using data from the World Bank. "Additional controls refers to time-varying, country-specific indicators of economic development, including per capita income, share of male and female employment in agriculture, and total share of employment in agriculture, taken from the World Bank.

and long-run trends in educational achievement, as well as specifications that allow for different functional forms of temperature.

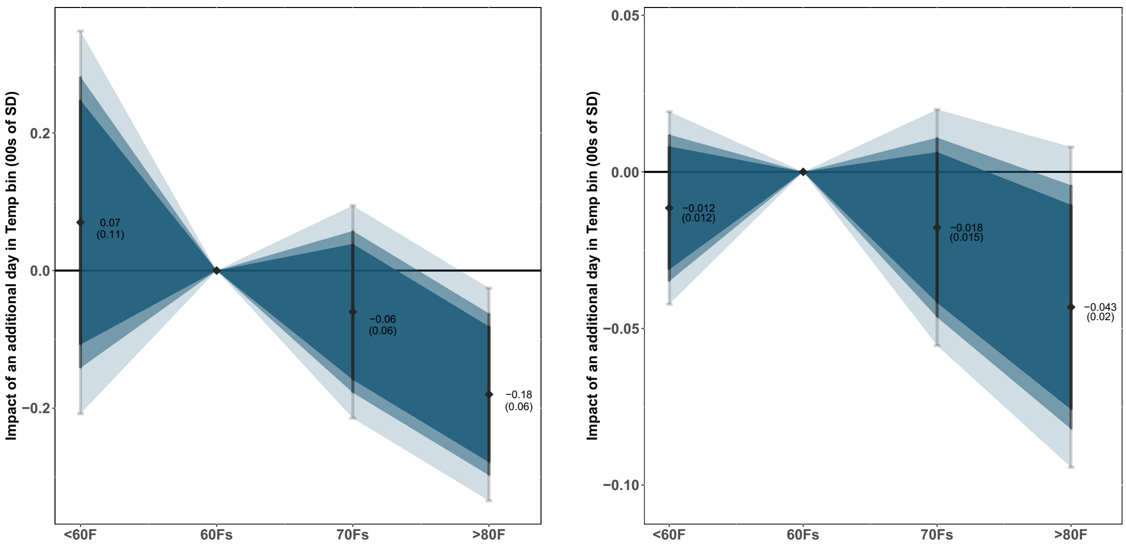
To provide evidence on potential mechanisms, we assess the impact of heat that occurs during three sets of mutually exclusive days of the year for each country in our sample: weekdays during the school year (henceforth “school days"), weekends during the school year, and summer vacation days. The effect of hot temperature on learning appears to be driven almost exclusively by hot school days (Figure 3.3 and Appendix Table C.1). Each additional hot school day lowers scores by 0.22 standard deviations (p= 0.002, 95% CI = [-0.36, -0.08], Table 3.3). A Wald test indicates a significant difference between the impact of hot school days and hot non-school days ($F_{1,57}$=3.41, p=0.07).

To further probe whether heat impacts learning through other correlated shocks, including the effects of heat on agricultural productivity, we run analyses that control for hot days during the rice growing season (Appendix Table C.2). In the countries for which we have data on growing seasons we find that hot school days, controlling for the number of hot days during the rice growing season, still appear to reduce student performance by 0.31 percent of a standard deviation (p= 0.013, 95% CI = [-0.55, -0.07]), whereas hot growing season days have statistically insignificant impacts (β=-0.297, $p = 0.322$, 95% CI = [-0.90, 0.31]). Furthermore, the findings are robust to including controls for changes in per capita income, share of labor force in agriculture, and female labor force participation, suggesting that they are likely not driven solely by correlated shocks to (gender-specific) economic incentives for educational investment (Shah and Steinberg, 2017).

Splitting the sample into “rich" and “poor" countries (above and below mean per capita income in 1995 in our sample, listed in Appendix Table C.3), we find that temperature exerts a significant impact in poorer countries (β=-0.14, $p = 0.001$, 95% CI = [-0.22, -0.07], Figure 3.3) but less so in richer ones (β=-0.024, $p = 0.733$, 95% CI = [-0.17, 0.12], Figure 3.3) , consistent with lower levels of adaptation and/or other channels (e.g. conflict (Hsiang *et al.*, 2013)) through which heat can affect student outcomes in developing countries.

Taken together, these results provide further evidence consistent with the claim that hotter temperature during learning periods exert a negative and casual impact on human capital accumulation. While these reduced form effects do not on their own demonstrate the mechanisms through which such

FIGURE 3.2: IMPACT OF TEMPERATURES ON PISA AND SEDA EXAM SCORES.



(A) *International (PISA)* (B) *US (SEDA)*

NOTES: *In both figures the shaded areas connect coefficients representing the effect of an additional school day in each temperature bin on subsequent achievement in hundredths of a standard deviation, with light to dark shading corresponding to 99 percent, 95 percent and 90 percent confidence intervals respectively. Sample sizes are N=281 and N=825,416 for panels (a) and (b) respectively. In panel (a) we show the impact of days below 60°F (+0.07 standard deviations,* $p = 0.517$*, 95% CI = [-0.14, 0.28]), days between 70°F and 80°F (-0.06 standard deviations,* $p = 0.316$*, 95% CI = [-0.17, 0.56]), and days greater than 80°F (-0.18 standard deviations,* $p = 0.007$*, 95% CI = [-0.31, -0.05]) on performance on the PISA exams. In panel (b) we show the impact of days below 60°F (-0.012 standard deviations,* $p = 0.335$*, 95% CI = [-0.035, 0.012]), days between 70°F and 80°F (-0.018 standard deviations,* $p = 0.226$*, 95% CI = [-0.05, 0.011]), and days greater than 80°F (-0.043 standard deviations,* $p = 0.03$*, 95% CI = [-0.08, -0.004]) on performance on the SEDA exams.*

impacts arise, they are consistent with the possibility that a portion of the effect is driven through heat's disruptive impact on learning.

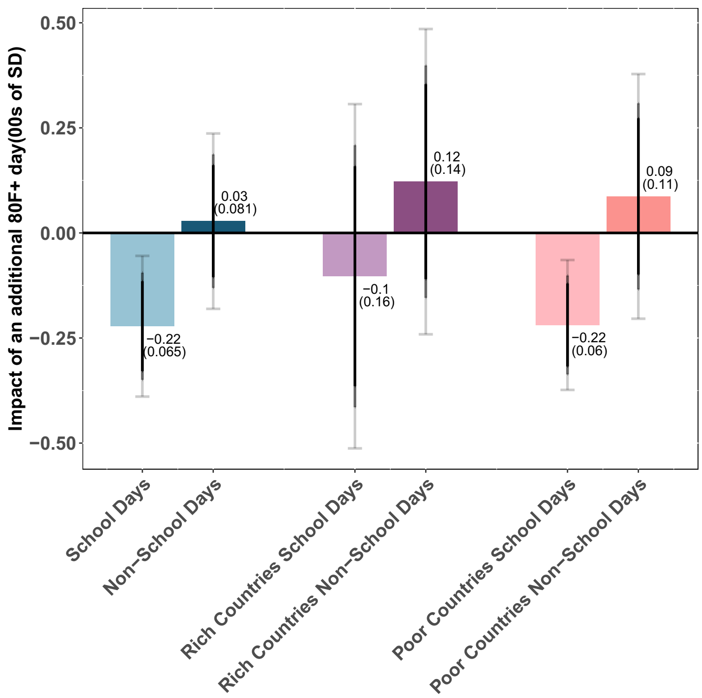
To better understand the extent to which our results are driven by physiological channels, we conduct a second set of analyses using more spatially resolved data from a highly developed, non-agrarian setting – where non-physiological factors are plausibly less influential – and with a richer set of demographic and location-specific characteristics.

3.2.2 U.S. analysis

Our second analysis examines data on standardized student achievement for over 12,000 U.S. school districts between 2009 and 2015 (Figure 3.4). Drawn from the Stanford Educational Data Archives (SEDA, (Reardon *et al.*, 2017)), these records comprise the near-universe of state-administered standardized math and verbal assessments for 3rd-8th graders, representing over 270 million test scores. These assessments, typically taken in March, April or May, vary across states but have been standardized by SEDA for national comparability. Similarly to PISA exams, these tests are meant to capture cumulative learning specific to each state-grade-subject. These data are thus uniquely suited for assessing the effect of heat during formal instructional periods, in contrast to tests used in other US studies(Graff Zivin *et al.*, 2017; Park *et al.*, 2020). Our unit of observation is at the level of district-by-grade-by-subject-by-year, resulting in approximately 825,000 observations, matched to district-level daily weather information using data from approximately 3,400 weather stations from the National Climatic Data Center (NCDC). To account for possible differences in school-year heat arising from regional differences in start/end dates, we use state-specific academic calendars as represented by the largest urban district in each state.

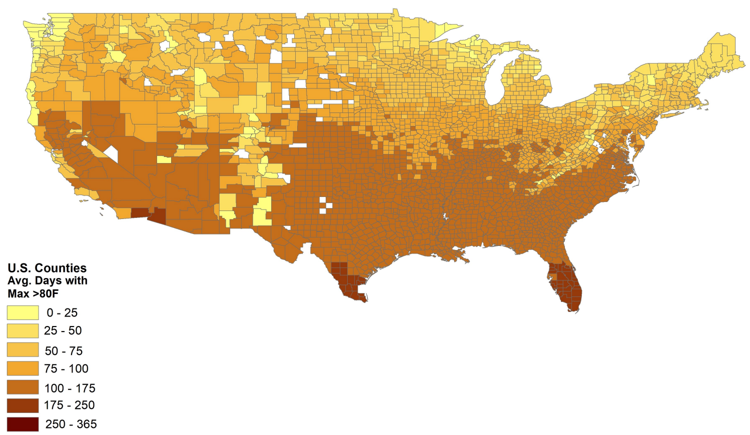
We again exploit random variation in year-to-year temperature within a given district over time to account for potential correlation between unobserved determinants of educational achievement and average climates across districts. For instance, schools in the American South typically perform worse than schools in the Northeast, but many factors other than climate including teacher quality and legacies of segregation may affect this cross-sectional relationship. The number of hot days during any given school year within a particular district, however, is plausibly exogenous, especially when

FIGURE 3.3: HETEROGENEITY OF HOT TEMPERATURE IMPACTS - PISA

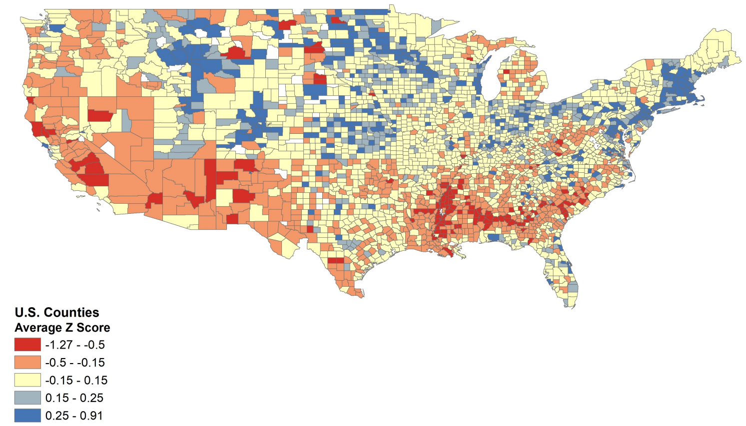


NOTES: *The first two columns show the impact on subsequent standardized achievement of hot (≥80°F) school days – i.e. weekdays during the school year – versus hot weekends, holidays, and summer vacation in the three years leading up to any given PISA assessment for all participating countries in our sample over the period 2000-2015 (n=271). We show the impact of hotter school days(-0.22 standard deviations,* $p = 0.002$*, 95% CI = [-0.36, -0.08]), hot non-school days (+0.03 standard deviations,* $p = 0.676$*, 95% CI = [-0.14, 0.22]) in the first two columns. Columns 3 through 6 show the corresponding effects for countries with below mean and above mean income in 1995 in our sample n=150, columns 5 and 6). We show the impact of hot school days (-0.099 standard deviations,* $p = 0.493$*, 95% CI = [-0.39, 0.2])and hot non-school days (+0.127 standard deviations,* $p = 0.394$*, 95% CI = [-0.18, 0.43]) in richer countries in columns 3 and 4 (n=132). Columns 5 and 5 (n=150) we show the impact in poorer countries of hot school days (-0.256 standard deviations,* $p < 0.001$*, 95% CI = [-0.36, -0.15]) and hot non-school days (+0.088 standard deviations,* $p = 0.342$*, 95% CI = [-0.10, 0.28]). Consistent with existing literature (e.g. (Barreca et al., 2015)) all coefficients can be interpreted as the effect relative to an additional day in the 60°'s F on combined math, verbal, and science scores. We provide our coefficient estimates and standard errors (in parenthesis) at the end of each bar.*

FIGURE 3.4: TEMPERATURE EXPOSURE AND SEDA SCORES.



(A) *Temperature Exposure*



(B) *SEDA Scores*

NOTES: *Panel A shows the average number of days with a maximum temperature above 80°F (2009-2015). Panel B shows the average Z score across all subjects and years within a county in our sample. Missing counties are those omitted from the SEDA data for confidentiality reasons.*

taking aggregate (regional) warming patterns into account.

TABLE 3.4: PRIOR YEAR TEMPERATURE AND TEST SCORES - SEDA

	Math and ELA scores		Math scores		ELA scores	
	(1)	(2)	(3)	(4)	(5)	(6)
(A) All hot days						
Total hot days	-0.043	-0.036	-0.066	-0.047	-0.031	-0.028
	(0.020)	(0.020)	(0.025)	(0.025)	(0.018)	(0.017)
	[0.030]	[0.071]	[0.008]	[0.057]	[0.079]	[0.104]
	{-0.08,-0.00}	{-0.07,0.00}	{-0.11,-0.02}	{-0.10,0.00}	{-0.07,0.00}	{-0.06,0.01}
N	825,416	825,416	400,953	400,953	424,198	424,198
(B) School vs. non-school days						
Hot school days	-0.065	-0.070	-0.111	-0.107	-0.036	-0.039
	(0.026)	(0.027)	(0.033)	(0.034)	(0.024)	(0.024)
	[0.014]	[0.010]	[0.001]	[0.002]	[0.131]	[0.099]
	{-0.12,-0.01}	{-0.12,-0.02}	{-0.18,-0.05}	{-0.17,-0.04}	{-0.08,0.01}	{-0.09,0.01}
Hot non-school days	-0.021	0.015	0.001	0.039	-0.034	-0.011
	(0.028)	(0.025)	(0.034)	(0.031)	(0.027)	(0.024)
	[0.447]	[0.561]	[0.974]	[0.215]	[0.204]	[0.635]
	{-0.08,0.03}	{-0.04,0.06}	{-0.06,0.07}	{-0.02,0.10}	{-0.09,0.02}	{-0.06,0.04}
N	825,416	825,416	400,953	400,953	424,198	424,198
Additional controls		X		X		X

NOTES: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses, with p-values in brackets and 95% confidence intervals in curly brackets. Coefficients in each column and panel come from a regression of hundredths of a standard deviation in test scores on the number of days above 80°F. Also included are controls for the number of days from 70-80°F and below 60°F, so that days from 60-70°F are the baseline category. Temperature is measured with the daily maximum temperature on school days from June to February prior to the test. All regressions include fixed effects for each school district and for each combination of test year, grade and subject (Mathematics or English Language Arts). Each observation is a district-year-grade-subject combination and all regressions are weighted by the number of test-takers per observation.

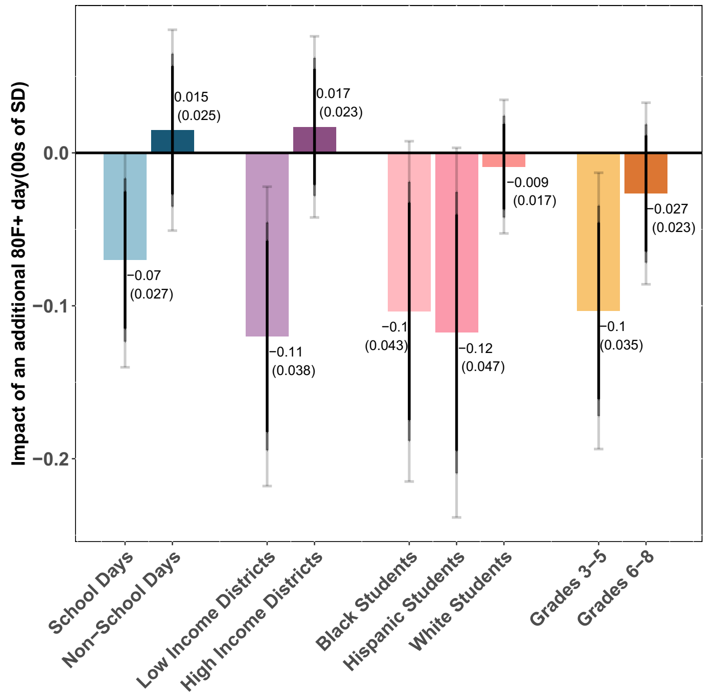
We find that students who experience hotter temperatures during the school year prior to their exams exhibit reduced learning. Each additional day of 80°F or hotter temperature reduces achievement by approximately 0.04 percent of a standard deviation (Figure 3.2 and Table 3.4, p=0.071, 95% CI = [-0.07, 0.00]). Our measures of significance are robust to correlation in error terms within any given state, which typically holds over 200 school districts. This effect is concentrated among school days, with each additional hot school day lowering achievement by 0.07 percent of a standard deviation (Figure 3.5, p=0.01, 95% CI = [-0.12, -0.02]). Similarly to the international analysis, heat on non-school days, such as weekends and summers, has no statistically significant impact on achievement (β=0.015, p=0.561, 95% CI = [-0.04, 0.06]). A Wald test indicates a significant difference between the impact of hot school days and hot non-school days ($F_{1,3394}$=5.54, p=0.019). These estimates imply that a student who experiences an additional school week (five school days) with daily maximum temperatures above 80°F will learn 0.35 percent of a standard deviation less than she otherwise would have during that school-year, which is equivalent to reducing teacher quality by about 3-4 percent (Chetty *et al.*, 2011).

The impact of heat exposure on learning is not confounded by precipitation, exam-day weather shocks, changing demographic compositions or resource levels of school districts, or spurious correlation between regional warming patterns and trends in educational achievement. That only hot weekdays during the school year reduce learning suggests once again that the set of mechanisms likely includes a reduction in contemporaneous educational inputs – whether in terms of the amount or intensity of learning time.

In the US, the impact of heat on math achievement is about three times larger than on ELA achievement. Each additional hot school day lowers math scores by 0.11 percent of a standard deviation (panel B, columns 3-4 of Table 3.4, p=0.002, 95% CI = [-0.17, -0.04]) but lowers ELA scores by less than 0.04 percent of a standard deviation (panel B, columns 5-6 of Table 3.4, p=0.099, 95% CI = [-0.09, 0.01]). There is little evidence that heat on non-school days affects achievement in either subject.

Importantly, hot temperature affects disadvantaged students much more than advantaged ones. Heat has substantially larger impacts on the achievement of students in lower income school districts and little impact in higher income districts, defined respectively as those in the lower and upper thirds of the district-level income distribution. Each additional hot school day lowers achievement in lower income districts by 0.12 percent of a standard deviation but has little discernible effect on achievement in higher income districts (Figure 3.5 and Table 3.5, p=0.002, 95% CI = [-0.19, -0.04]). Each hot school day lowers the achievement of Black and Hispanic students by 0.10-0.12 percent of a standard deviation but has no statistically significant impact on White students (Table 3.6, Black students: p=0.017, 95% CI = [-0.19, -0.02]; Hispanic students: p=0.012, 95% CI = [-0.21, -0.03]; White students: β=-0.009, p=0.593, 95% CI = [-0.04, 0.02]). A week above 80°F for the average Black or Hispanic student reduces learning by an amount equivalent to reducing teacher value-added

FIGURE 3.5: HETEROGENEITY OF HOT TEMPERATURE IMPACT - SEDA



NOTES: *The first two columns show the impact on subsequent standardized achievement of hot (≥80°F) school days (-0.07 standard deviations,* $p = 0.01$*, 95% CI = [-0.12, -0.02]) – i.e. weekdays during the school year – versus hot weekends, holidays, and summer vacation (+0.015 standard deviations,* $p = 0.974$*, 95% CI = [-0.06, 0.07]) for all U.S. school districts over the period 2009-2015 (n=825,416). Columns 3 through 9 show the effect of hot school days for the bottom (-0.11 standard deviations,* $p = 0.002$*, 95% CI = [-0.19, -0.04]) and top terciles (+0.017 standard deviations,* $p = 0.456$*, 95% CI = [-0.03, 0.06]) of the district income distribution, for Black (-0.1 standard deviations,* $p = 0.017$*, 95% CI = [-0.19, 0.02]), Hispanic (-0.12 standard deviations,* $p = 0.012$*, 95% CI = [-0.21, -0.03]) and White (-0.009 standard deviations,* $p = 0.593$*, 95% CI = [-0.04, 0.02]) students within each district, and for elementary (-0.1 standard deviations,* $p = 0.003$*, 95% CI = [-0.17, -0.03]) and middle school students (-0.027 standard deviations,* $p = 0.249$*, 95% CI = [-0.07, 0.02]) in each district respectively (n=273,466; 273,266; 183,060; 222,042; 733,219; 425,301; 400,095 for columns 3-9 respectively). All coefficients can be interpreted as the effect relative to an additional day in the 60's°F on combined math and English Language and Arts scores.*

TABLE 3.5: HETEROGENEITY BY DISTRICT INCOME AND TEMPERATURE - SEDA

	All (1)	Lower income (2)	Higher income (3)	Grades 3-5 (4)	Grades 6-8 (5)
Math and ELA	-0.070	-0.118	0.017	-0.103	-0.027
	(0.027)	(0.038)	(0.023)	(0.035)	(0.023)
	[0.010]	[0.002]	[0.456]	[0.003]	[0.249]
	{-0.12,-0.02}	{-0.19,-0.04}	{-0.03,0.06}	{-0.17,-0.03}	{-0.07,0.02}
N	825,416	381,254	444,162	425,301	400,095
Test scores (MM)	270.9	149.5	121.4	141.8	129.1
Math	-0.107	-0.187	0.033	-0.162	-0.024
	(0.036)	(0.043)	(0.032)	(0.044)	(0.022)
	[0.003]	[0.000]	[0.296]	[0.000]	[0.277]
	{-0.18,-0.04}	{-0.27,-0.10}	{-0.03,0.10}	{-0.25,-0.08}	{-0.07,0.02}
N	400,953	183,547	217,406	211,442	189,326
Test scores (MM)	129.3	71.2	58.1	70.2	59.1
ELA	-0.049	-0.082	0.012	-0.057	-0.036
	(0.026)	(0.036)	(0.021)	(0.028)	(0.027)
	[0.065]	[0.022]	[0.567]	[0.043]	[0.179]
	{-0.10,0.00}	{-0.15,-0.01}	{-0.03,0.05}	{-0.11,-0.00}	{-0.09,0.02}
N	424,198	197,586	226,612	213,563	210,471
Test scores (MM)	141.7	78.3	63.3	71.6	70.0

NOTES: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses, with p-values in brackets and 95% confidence intervals in curly brackets. Coefficients in each column and panel come from a regression of hundredths of a standard deviation in test scores on the number of days above 80°F. Also included are controls for the number of days from 70-80°F and below 60°F, so that days from 60-70°F are the baseline category. Temperature is measured with the daily maximum temperature on school days from June to February prior to the test. All regressions include fixed effects for each school district and for each combination of test year, grade and subject (Mathematics or English Language Arts). Each observation is a district-year-grade-subject combination and all regressions are weighted by the number of test-takers per observation. Lower income and higher income districts are those with an average student poverty rate respectively above and below 50 percent.

by 5 to 6 percent of a standard deviation.

TABLE 3.6: HETEROGENEITY BY RACE AND GENDER - SEDA

	All (1)	White (2)	Black (3)	Hispanic (4)	Male (5)	Female (6)
Math and ELA	-0.070	-0.009	-0.104	-0.118	-0.079	-0.072
	(0.027)	(0.017)	(0.043)	(0.047)	(0.033)	(0.026)
	[0.010]	[0.593]	[0.017]	[0.012]	[0.018]	[0.007]
	{-0.12,-0.02}	{-0.04,0.02}	{-0.19,-0.02}	{-0.21,-0.03}	{-0.14,-0.01}	{-0.12,-0.02}
N	825,416	733,219	183,060	222,042	695,141	684,263
Test scores (MM)	270.9	137.2	41.0	56.1	135.4	129.0
Math	-0.107	-0.002	-0.138	-0.187	-0.099	-0.121
	(0.036)	(0.021)	(0.046)	(0.058)	(0.039)	(0.035)
	[0.003]	[0.937]	[0.003]	[0.001]	[0.011]	[0.001]
	{-0.18,-0.04}	{-0.04,0.04}	{-0.23,-0.05}	{-0.30,-0.07}	{-0.18,-0.02}	{-0.19,-0.05}
N	400,953	357,385	88,276	106,134	337,617	332,431
Test scores (MM)	129.3	66.4	19.8	26.1	64.7	61.6
ELA	-0.049	-0.010	-0.081	-0.090	-0.067	-0.037
	(0.026)	(0.016)	(0.047)	(0.039)	(0.031)	(0.024)
	[0.065]	[0.518]	[0.083]	[0.020]	[0.030]	[0.116]
	{-0.10,0.00}	{-0.04,0.02}	{-0.17,0.01}	{-0.17,-0.01}	{-0.13,-0.01}	{-0.08,0.01}
N	424,198	375,501	94,593	115,560	357,142	351,420
Test scores (MM)	141.7	70.8	21.2	30.0	70.7	67.4

NOTES: Heteroskedasticity robust standard errors clustered by weather sensor are in parentheses, with p-values in brackets and 95% confidence intervals in curly brackets. Coefficients in each column and panel come from a regression of hundredths of a standard deviation in test scores on the number of days above 80°F. Also included are controls for the number of days from 70-80°F and below 60°F, so that days from 60-70°F are the baseline category. Temperature is measured with the daily maximum temperature on school days from June to February prior to the test. All regressions include fixed effects for each school district and for each combination of test year, grade and subject (Mathematics or English Language Arts). Each observation is a district-year-grade-subject combination and all regressions are weighted by the number of test-takers per observation.

The effect of hot school days is also larger for younger students than for older students. Each additional such day lowers the achievement of third through fifth graders by 0.08-0.13 percent of a standard deviation but has a statistically insignificant impact on those in grades six through eight (Figure 3.5 and Table 3.5, p=0.003, 95% CI = [-0.17, -0.03])). This is consistent with previous evidence suggesting that younger children are likely to be more adversely affected by thermal stress, either due to physiology or behavior (Rowland, 2008). This could, however, be due to other factors such as the potentially lower prevalence of school air conditioning in elementary schools relative to middle schools.

3.3 Discussion

Taken together, these results suggest a different perspective on how climate shapes human cognitive capacity. Thermal conditions in the physical learning environment appear to causally influence cumulative learning: a fact not yet documented in the voluminous literature on cross-country comparisons in student achievement (Woessmann, 2016). It appears that heat exposure during the learning period, all else equal, directly slows the rate of human capital formation, in part through persistent disruptions to the learning process. As noted above, the realized temperature environments facing students across the world vary dramatically, suggesting important implications for our understanding of differences in educational achievement and human capital.

We find heat exposure to be a compelling mechanism. It matches emerging findings on the effects of temperature on labor capacity (Kjellstrom and Crowe, 2011; Graff-Zivin and Neidell, 2012), morbidity and mortality (Deschênes and Greenstone, 2011; Anderson *et al.*, 2013), and short-run cognition (Graff Zivin *et al.*, 2017; Park, forthcoming). We note, however, that this analysis does not imply that heat exposure is the only mechanism at play: many others are likely relevant in explaining the relationship between climate and levels of human capital across countries. Teasing apart the potential mechanisms in greater detail – for instance, whether hotter temperatures drive student/teacher absenteeism; and understanding the extent to which these mechanisms interact – for instance, whether poor nutrition and hunger exacerbate heat-induced cognitive impacts – are important questions for future work.

Importantly, the magnitude of these disruptions appear to vary greatly across socioeconomic groups – both across and within countries. As shown in Figure 3.5, the effect of an additional 80°F day in US school districts in the lower third of average income is approximately -0.12 (p=0.002, 95% CI = [-0.19, -0.04]) percent of a standard deviation, while the effect in the top third is statistically indistinguishable from zero. Impacts are also larger for some racial minorities, particularly Black and Hispanic students. This is consistent with evidence from the United States suggesting that school and home air conditioning status is correlated with student race and income (Park *et al.*, 2020), and suggest that climatic factors may contribute to longstanding racial achievement gaps.

How large are these effects? Suppose we take the US estimates as a lower bound for the rest of the world, given relatively high rates of air conditioning there. Education researchers have, for instance, examined the impact of improving teacher quality or reducing class sizes on learning outcomes. Our US analyses suggest that, even with relatively high levels of air conditioning, a school year with 30 additional days above 80°F reduces learning by approximately 2.1 percent of a standard deviation. This is large enough to offset the gains of reducing class sizes by approximately 3-4 percent, or to offset improving teacher quality by 20 percent of a standard deviation. For lower income students, the effect of the same temperature event appears to be nearly 3 times larger. These sizable magnitudes suggest the learning impacts of a hotter climate could result in large real consequences, especially given that students in many tropical economies regularly experience more than 100 such days per school year (Table 3.1). Put differently, greater heat exposure during the school year may lead students in Brazil to learn 6 percent less than their South Korean counterparts per year, which, over time, might explain around a third of the difference in their PISA performance.[2]

This perspective has important policy implications. It suggests that climate may have a more direct and persistent influence on economic growth than previously appreciated. Human capital accumulation is central to national economic growth and individual economic mobility (Goldin and Katz, 2009; Hanushek and Woessmann, 2016), and current climatic conditions appear to slow the rate of human capital accumulation for some more than others. This suggests that policies aimed at improving physical learning environments, whether in the form of electric infrastructure or low-income energy assistance, may pay larger dividends over time than previously appreciated. These pro-growth, pro-adaptation policies may or may not include school air conditioning, which may improve student cognition as well as teacher attendance/retention, but which may also exacerbate the climate externality. Making such investments to facilitate learning in hotter environments may be particularly important in light of evidence suggesting that education itself may be an important climate adaptation strategy (Lutz *et al.*, 2014).

It also suggests that current estimates of the social costs of carbon (SCC) may be understated.

[2]The gap in average PISA performance between South Korea and Brazil is approximately 1.43 standard deviations, while the difference in hot school days is approximately 85 per year. Assuming that PISA exams test knowledge that is accumulated over 9 years of formal schooling, and assuming for simplicity that effects accumulate linearly, this would amount to 0.53/1.43=0.37 of the PISA gap at age 15.

Existing integrated assessment models do not include direct impacts on human capital, and often model climate impacts as a non-accumulating reduction in the level of GDP as opposed to cumulative growth rate effects. Adding these arguments to the damage function would likely shift the entire distribution of estimates to be more negative (Greenstone *et al.*, 2013). Accounting for within-country regressivity of these impacts, as suggested by our findings, may also imply larger SCC estimates, regardless of one's choice of pure rate of time preference or discount rate (Anthoff and Emmerling, 2019).

3.4 Methods

3.4.1 Data description

Global Temperature Data

We use separate temperature data sets for our global and domestic analyses given varying geographic and temporal coverage. For the global analysis, which uses PISA test scores from many different countries, we start with data from NOAA's Global Historical Climatology Network (GHCN). This provides us with daily data from a network of more than 100,000 stations located in approximately 180 countries. The data provided includes daily max and min temperatures and total daily precipitation. We collect data starting in 1995 and pull all the available data for the countries that appear in our PISA sample.

In order to count school vs. non-school days we exclude weekend days from the school days and assign each country a dummy called "summer" on the days that students in that country are typically on summer vacation. When schools start on a range of days, for example "the first two weeks of September," we choose a date at or adjacent to the midpoint of the range. We separately identify weekends that occur during the school year and those that occur during the summer so that we can examine whether heat on non-school days during the school year has different effects than heat on non-school days outside of the school year.

To create our temperature bins we count the number of days with maximum temperature in 10°F bins from 0°F to 140°F by station. We group all days below 0°F into a single bin. Each country is then assigned the weighted average (across all stations) number of days in each temperature bin in each year. Weights are based on the population living within 15km of the station as measured by LandScan population data. We weight stations based on their population in 2000, at the beginning of our sample. We also create lagged variables that count the number of days in each bin in each of five lagged years as well as the cumulative days in each bin over the previous 1-5 years. The cumulative lag variable does not count the number of days in a given bin in the contemporaneous year.

Because we use a temperature binning approach where we average across the number of days in each bin by country-year - as opposed to averaging temperature across stations within a country and then binning - we need to impute the missing days. Most of the existing literature avoids this problem because it averages across stations in a geography to create a geography based average temperature for each day, and then counts the number of days in any given bin for each geography's average temperature. Over small geographies this may lead to relatively small measurement error. Over larger geographies such as the countries in our PISA sample, however, we believe that such an approach would introduce substantial measurement error. For example, in the US the average temperature for a given day in June might by 70°F but that masks the fact that much of the Southern U.S. might be experiencing 90°F+ temperatures. Creating station specific bins and then averaging within the bins accounts for this by allowing those 90°F+ days in June in the South to count as 90°F+ days. We do this through an iterative process that identifies days with a missing temperature reading and the closest days before and after that day with a reading.

We assign a weighted average of the nearest non-missing days where the weights are the number of days between the missing and non-missing day. For example, if April 5^{th} was missing but April 4^{th} and 9^{th} were not, the temperature on the 5^{th} would be imputed as the average of the temperature on the 4^{th} and 9^{th} with greater weight on the 4^{th}. The 6^{th}-8^{th} would be imputed iteratively using the same process. We limit the gaps we impute in this way to 21 days. In total, impute temperature on 2.2% of the station-days in our sample in this way.

If, after this iterative process, there are still days with missing data for a given station we impute those data as the average temperature at all stations in the country with non-missing data on that day. This imputation affects 3.8% of our data.

To calculated population weighted temperature averages we use data from LandScan on the population in each 1km x 1km pixel across all of the countries in our sample. We draw 15km buffers around each weather station and then assign the population within the buffer to that weather station. When we calculate the average within temperature bins across stations we weight each station by this count of population. We do this separately for each year in our sample. As a robustness check we also calculate Thiessen polygons around each station and assign the population within each Thiessen polygon to each station. The results are similar.

U.S. Temperature Data

Daily temperature data come from the National Oceanic and Atmospheric Administration's Daily Global Historical Climatology Network, which includes station-level data for thousands of weather stations across the United States. We focus on the subset of nearly 3,400 weather stations with daily temperature data available for at least 95 percent of the days from from July 1, 2004 through June 30, 2015, the time period covering potential test-taking dates of our sample. Doing so allows us to assign each school district a single, stable weather station over the entire time period, which avoids endogeneity concerns driven by the possibility that stations coming online or going offline are somehow correlated with local population growth, economic conditions or temperatures conditions in ways that might contaminate our estimates (Auffhammer and Mansur, 2014). We impute the small proportion of missing daily observations with those from the nearest stations with non-missing data.

We assign each school district to the weather station nearest to that district's centroid, resulting in an average distance of 9.6 miles between each district's centroid and the weather station being used to measure temperature at that district. We define our primary heat exposure variable as the number of days the average daily maximum temperature exceeded a given multiple of 10°F from June 1 to February 28 in the year prior to the test. We use daily maximum temperature because schooling occurs during the daytime when such temperatures usually occur. Of course, to the extent that daytime maximum and nighttime minimum temperature is correlated, some of our effect may be driven by disrupted sleep. We do not take a stand on whether sleep is a factor or not, as both in-class and at-home disruptions through learning that are brought about by the physiological effects of heat are of interest.

We use the June-February time period because the exact timing of SEDA's standardized exams varies by state and year but almost always occurs between March and May. We focus particularly on temperature experienced on school days, treating non-school days (weekends and all summer days between June 15 and August 15) as separate sources of variation. We also use the weather stations to construct test date temperature, rain and snowfall, as well as cumulative rain and snowfall exposure over the year prior to the test, which help account for potential independent effects of such precipitation.

PISA Data

PISA assessments are designed to capture cumulative skills developed during formal schooling (e.g. arithmetic, basic scientific concepts, reading comprehension), and to be comparable across countries. Our data on average PISA scores by country comes from the National Center for Education Statistics (NCES) International Data Explorer. The NCES assembles average country scores by year in math, science and reading from the PISA microdata provided by the OECD. We follow the advice of the NCES and do not compare math and science scores from 2000 or 2003 (for science) with later years because of changes in PISA methodology. We do not modify the raw PISA data from NCES except to drop countries from the sample for which we do not have temperature data and those with only one year of PISA data. We exclude PISA data from sub-national units (from individual states within the U.S. for example). A minimum of 5,000 students are sampled in each country that participates unless the total population of 15 year old students is less than 5,000, in which case all students are tested. Some large countries sample more students. In total, more than 500,000 students took a PISA exam across all participating countries in 2015.

PISA scores are designed to have a global average of 500 and student-level standard deviation of 100, which we use to compute standardized versions of each country's math, science, and reading scores. In any given year, there is wide variation in performance across countries. On the 2009 PISA exam, for example, South Korean students averaged 546 points in math while Indonesian students averaged 371 points. Our primary outcome measure is the average of each country's three subject scores in any given year, standardized so that effects can be interpreted in terms of student-level standard deviations (similar to SEDA).

Summary Statistics for PISA Sample

See Table 3.2 for summary statistics on the PISA sample. On average, countries in our sample are hotter than the United States - experiencing 114 school days over the previous 3 years above 80°F vs. only 97 such days in the US - and poorer (per capita income=$26,000 vs $42,000) with lower PISA scores (Normalized score = -0.28 vs. -0.08). We also split the sample into rich and poor countries based on where a country's per capita income in 1995 ranks in our sample. We define rich as countries that have a per capita income in 1995 above $14,000, roughly the average in our sample for that year. Splitting the sample into rich and poor indicates that the rich sample is substantially cooler, wealthier and has a lower population of test takers than the poor countries. PISA scores are substantially better in the rich sample on average, with lower variance within the sample.

SEDA Data

Data from the Stanford Education Data Archive (SEDA) are based on the standardized accountability tests in math and English Language Arts (ELA) administered annually by each state to all public-school students in grades 38. SEDA combines information on the test scores in each school district with information from the National Assessment of Educational Progress, creating scores that are nationally comparable across districts in different states.

Our version of SEDA's data spans the school years ending 2009-15 and contains elementary and middle school students from approximately 12,000 school districts across all 50 states. We observe a standardized measure of both math and ELA achievement at the district-by-grade-by-year level. We observe this measure averaged across all test-takers in a school district, as well as for some demographic subgroups. The particular standardization used implies that effect sizes can be interpreted in student-level standard deviations.

Other International Data

In addition to temperature and PISA performance data we collect data on a set of potentially relevant co-variates for the countries in our sample. All of these data come from the World Bank's World Development Indicators. We collect time-varying measures of the share of total employment in agriculture, per capita income, share of male and female employment in agriculture, total population, and the share of the population made up by 15-19 year olds. The only data we modify is the 15-19 year old population share, which we combine with the total population to estimate the absolute number of 15-19 year-olds in each country-year. We match all data to temperature and PISA country-years using country ISO codes.

3.4.2 Empirical approach

Our econometric approach exploits the quasi-random variation within a given geography's total exposure to days above 80°F in the years between test takes. The geographic unit in PISA is a country and in SEDA is a school district. The time between test takes is 3 years in PISA and 1 year in SEDA. To account for serial correlation in temperature shocks across geographies we cluster standard errors at the relevant geographic unit. In all statistical tests we assume normality but do not formally test for it. All tests of significance are two-tailed.

We estimate several versions of the base model:

$$\bar{Z}_{it} = \sum_{k=1}^{9} \beta_k TMAX_{ikg} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \tag{3.1}$$

where $\bar{Z}_{it}$ is the normalized PISA score in country i and year t. $TMAX$ is the total number of days with maximum temperature in each of k degree bins in geography i in the gap g between exam takes. X_{it} is a vector of geography-year specific controls, including total annual precipitation in the year of the exam as well as the gap year(s), the same set of k degree bins in the year of the exam and, in the case of the PISA data, the controls from the World Bank described above. Parameters δ_i and γ_t are geography and year fixed effects. ω_{ct} is a continent-specific time trend included in the PISA regressions. ϵ_{it} is the error term. We weight each geography by the total number of 15-19 year-olds in that country in the exam year in the PISA data, as calculated from the World Bank data, and by the students in each district taking the exam in the SEDA data.

Our variable of interest is β_9 for the bin representing days over 80°F. Because we omit the 60-70°F bin from our set of controls the coefficient β_k should be interpreted as exchanging one day over the relevant gap in the 60-70°F bin for one $>$ 80°F.

Identification rests on the assumption that the number of days in any given temperature bin, and therefore the 80°F+ bin we are interested in, varies randomly within a geography from year-to-year. This year-to-year variation results in random variation in the aggregate exposure that students experience in the lead up to their exams. To account for possible spurious correlation between regional warming trends and secular changes in educational outcomes, we include continent-specific trends in all regressions. Our approach is analogous to the now widely used binning of annual temperatures first described in (Deschênes and Greenstone, 2007).

School vs. Non-school days Estimation

In our primary specifications we bin all days in a year together. We also separately report results of the effect of school days above 80°F and non-school days above 80°F. There we estimate the following model:

$$\begin{aligned} \bar{Z}_{it} = \sum_{k=1}^{9} \beta_k TMAX_{ikg}^{School} + \sum_{k=1}^{9} \psi_k TMAX_{ikg}^{Non-school} + \sigma \mathbf{X}_{it} \\ + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \end{aligned} \tag{3.2}$$

where the variables are as before but β reports the estimates of the impact of days while school is in session while ψ reports the effects of non-school days.

Summer vs. Non-summer Estimation

We distinguish school year days further by separating school year weekend days from school year non-weekend days. We estimate:

$$\begin{aligned}\bar{Z_{it}} = \sum_{k=1}^{9} \beta_k TMAX_{ikg}^{School} + \sum_{k=1}^{9} \psi_k TMAX_{ikg}^{Summer} \\ + \sum_{k=1}^{9} \phi_k TMAX_{ikg}^{School\ weekend} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{kt} + \epsilon_{ijt}\end{aligned} \quad (3.3)$$

where the variables are as before but β reports the estimates of the impact of days while school is in session while ψ reports the effects of non-school days.

Subject Specific Estimation

Finally, we estimate subject specific effects. There we return to the original estimating equation:

$$\bar{Z_{its}} = \sum_{k=1}^{9} \beta_k TMAX_{ikg} + \sigma \mathbf{X}_{it} + \gamma_t + \delta_i + \omega_{ct} + \epsilon_{it} \quad (3.4)$$

However, we replace $\bar{Z_{it}}$ with the subject specific normalized score, $\bar{Z_{its}}$ for each of reading, science and math in PISA and ELA and math in SEDA. In each case we calculate the normalized score in the way described above. In the PISA data, for both math and science we use the shorter panel in order to avoid data comparability issues due to changes in the PISA methodology in those subjects. To estimate the subject specific effects of school and non-school days we substitute $\bar{Z_{its}}$ into equation 3.2.

References

ABADIE, A., ATHEY, S., IMBENS, G. W. and WOOLDRIDGE, J. (2017). *When should you adjust standard errors for clustering?* Tech. rep., National Bureau of Economic Research.

ABAY, K. A., ABATE, G. T., BARRETT, C. B. and BERNARD, T. (2019). Correlated non-classical measurement errors,second bestpolicy inference, and the inverse size-productivity relationship in agriculture. *Journal of Development Economics*, **139**, 171–184.

ABRAHAM, S. and SUN, L. (2018). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Available at SSRN 3158747*.

ACEMOGLU, D. and AUTOR, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, vol. 4, Elsevier, pp. 1043–1171.

—, JOHNSON, S. and ROBINSON, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *The American Economic Review*, **91** (5), 1369–1401.

ALHOLA, P. and POLO-KANTOLA, P. (2007). Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment*.

ALIX-GARCIA, J., MCINTOSH, C., SIMS, K. R. and WELCH, J. R. (2013). The ecological footprint of poverty alleviation: evidence from mexico's oportunidades program. *Review of Economics and Statistics*, **95** (2), 417–435.

ALLEN, R. C. (2011). Why the industrial revolution was british: commerce, induced invention, and the scientific revolution 1. *The Economic History Review*, **64** (2), 357–384.

ALMÅS, I., AUFFHAMMER, M., BOLD, T., BOLLIGER, I., DEMBO, A., HSIANG, S. M., KITAMURA, S., MIGUEL, E. and PICKMANS, R. (2019). *Destructive Behavior, Judgment, and Economic Decision-making under Thermal Stress*. Working Paper 25785, National Bureau of Economic Research.

ALPERT, P., SHVAINSHTEIN, O. and KISHCHA, P. (2012). Aod trends over megacities based on space monitoring using modis and misr. *American Journal of Climate Change*, **1** (03), 117.

AMBASTA, P., SHANKAR, P. V. and SHAH, M. (2008). Two years of nrega: The road ahead. *Economic and Political Weekly*, pp. 41–50.

ANAND, G. (2016). Farmers unchecked crop burning fuels indias air pollution. *The New York Times*.

ANDERSON, G. B., DOMINICI, F., WANG, Y., MCCORMACK, M. C., BELL, M. L. and PENG, R. D. (2013). Heat-related emergency hospitalizations for respiratory diseases in the medicare population. *American journal of respiratory and critical care medicine*, **187** (10), 1098–1103.

ANDERSON, M. L. (2015). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*.

ANDINI, A., BONNET, S., ROUSSET, P. and HASANUDIN, U. (2018). Impact of open burning of crop residues on air pollution and climate change in indonesia. *Current Science*, **115** (12), 2259–2266.

ANTHOFF, D. and EMMERLING, J. (2019). Inequality and the social cost of carbon. *Journal of the Association of Environmental and Resource Economists*, **6** (2), 29–59.

ARROW, K., BOLIN, B., COSTANZA, R., DASGUPTA, P., FOLKE, C., HOLLING, C. S., JANSSON, B.-O., LEVIN, S., MÄLER, K.-G., PERRINGS, C. *et al.* (1995). Economic growth, carrying capacity, and the environment. *Ecological economics*, **15** (2), 91–95.

ASHER, S., LUNT, T., MATSUURA, R. and NOVOSAD, P. (2019). The Socioeconomic High-resolution Rural-Urban Geographic Dataset on India (SHRUG), working paper.

— and NOVOSAD, P. (2018). *Rural roads and local economic development*. The World Bank.

ATHEY, S. (2018). The impact of machine learning on economics.

— and IMBENS, G. W. (2017). The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, vol. 1, Elsevier, pp. 73–140.

AUFFHAMMER, M. (2018). Quantifying economic damages from climate change. *Journal of Economic Perspectives*, **32** (4), 33–52.

—, HSIANG, S. M., SCHLENKER, W. and SOBEL, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*, **7** (2), 181–198.

— and MANSUR, E. T. (2014). Measuring climatic impacts on energy consumption: A review of the empirical literature. *Energy Economics*, **46**, 522–530.

AZAM, M. (2011). The impact of indian job guarantee scheme on labor market outcomes: Evidence from a natural experiment. *Available at SSRN 1941959*.

AZTATZI-AGUILAR, O. G., URIBE-RAMÍREZ, M., ARIAS-MONTAÑO, J. A., BARBIER, O. and DE VIZCAYA-RUIZ, A. (2015). Acute and subchronic exposure to air particulate matter induces expression of angiotensin and bradykinin-related genes in the lungs and heart: Angiotensin-ii type-i receptor as a molecular target of particulate matter exposure. *Particle and fibre toxicology*, **12** (1), 17.

BALAKRISHNAN, K., DEY, S., GUPTA, T., DHALIWAL, R., BRAUER, M., COHEN, A. J., STANAWAY, J. D., BEIG, G., JOSHI, T. K., AGGARWAL, A. N. *et al.* (2019). The impact of air pollution on deaths, disease burden, and life expectancy across the states of india: the global burden of disease study 2017. *The Lancet Planetary Health*, **3** (1), e26–e39.

BANERJEE, R. and MAHARAJ, R. (2019). Heat, infant mortality and adaptation: Evidence from india. *Journal of Development Economics*, p. 102378.

BANK, T. W. (2018). *The State of Social Safety Nets 2018*. Tech. rep., The World Bank, Washington, DC.

BARBIER, E. B. (2010). Poverty, development, and environment. *Environment and Development Economics*, **15** (6), 635–660.

BARRECA, A., CLAY, K., DESCHÊNES, O., GREENSTONE, M. and SHAPIRO, J. S. (2015). Convergence in adaptation to climate change: Evidence from high temperatures and mortality, 1900-2004. *American Economic Review*, **105** (5), 247–51.

—, —, DESCHENES, O., GREENSTONE, M. and SHAPIRO, J. S. (2016). Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, **124** (1), 105–159.

BARROWS, G., GARG, T. and JHA, A. (2018). The economic benefits versus environmental costs of india's coal fired power plants. *Available at SSRN 3281904*.

BEHRER, A. P. and MAUTER, M. (2017). Allocating damage compensation in a federalist system: Lessons from spatially resolved air emissions in the marcellus. *Environmental Science and Technology*, **51**.

BERG, E., BHATTACHARYYA, S., DURGAM, R. and RAMACHANDRA, M. (2012). Can rural public works affect agricultural wages? evidence from india.

BHARGAVA, A. K. (2014). The impact of india's rural employment guarantee on demand for agricultural technology.

BHUVANESHWARI, S., HETTIARACHCHI, H. and MEEGODA, J. N. (2019). Crop residue burning in india: Policy challenges and potential solutions. *International journal of environmental research and public health*, **16** (5), 832.

BIKKINA, S., ANDERSSON, A., KIRILLOVA, E. N., HOLMSTRAND, H., TIWARI, S., SRIVASTAVA, A., BISHT, D. and GUSTAFSSON, Ö. (2019). Air quality in megacity delhi affected by countryside biomass burning. *Nature Sustainability*, p. 1.

BISHOP, K. C., KETCHAM, J. D. and KUMINOFF, N. V. (2018). *Hazed and Confused: The Effect of Air Pollution on Dementia*. Working Paper 24970, National Bureau of Economic Research.

BLACKBURN, M. L. (2015). The relative performance of poisson and negative binomial regression estimators. *Oxford Bulletin of Economics and Statistics*, **77** (4), 605–616.

BLACKMAN, A. (2010). Alternative pollution control policies in developing countries.

BURKE, M. and EMERICK, K. (2016). Adaptation to climate change: Evidence from us agriculture. *American Economic Journal: Economic Policy*, **8** (3), 106–40.

—, HSIANG, S. M. and MIGUEL, E. (2015). Global non-linear effect of temperature on economic production. *Nature*, **527** (7577), 235–239.

BURLIG, F. and PREONAS, L. (2016). Out of the darkness and into the light? development effects of rural electrification. *Energy Institute at Haas WP*, **268**.

BURNEY, J. and RAMANATHAN, V. (2014). Recent climate and air pollution impacts on indian agriculture. *Proceedings of the National Academy of Sciences*, **111** (46), 16319–16324.

BURNEY, J. A. (2020). The downstream air pollution impacts of the transition from coal to natural gas in the united states. *Nature Sustainability*, pp. 1–9.

CAMERON, A. C. and TRIVEDI, P. K. (2001). Essentials of count data regression. *A companion to theoretical econometrics*, **331**.

CARLETON, T., DELGADO, M., GREENSTONE, M., HOUSER, T., HSIANG, S., HULTGREN, A., JINA, A., KOPP, R. E., MCCUSKER, K., NATH, I. *et al.* (2018). Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *NBER working paper*.

— and MENG, K. C. (2020). Causal empirical estimates suggest covid-19 transmission rates are highly seasonal. *medRxiv*.

CASSOU, E. (2018). *Field Burning*. Tech. rep., World Bank, Washington, DC.

CDC (2020). *Provisional death counts for coronavirus disease (COVID-19)*. Tech. rep., Centers for Disease Control.

CHANG, T., GRAFF ZIVIN, J., GROSS, T. and NEIDELL, M. (2016). Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy*, **8** (3), 141–69.

CHARBONNEAU, K. (2012). Multiple fixed effects in nonlinear panel data models. *Unpublished manuscript*.

CHAY, K., DOBKIN, C. and GREENSTONE, M. (2003). The clean air act of 1970 and adult mortality. *Journal of risk and uncertainty*, **27** (3), 279–300.

CHAY, K. Y. and GREENSTONE, M. (2003a). *Air quality, infant mortality, and the Clean Air Act of 1970*. Tech. rep., National Bureau of Economic Research.

— and — (2003b). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The quarterly journal of economics*, **118** (3), 1121–1167.

— and — (2005). Does air quality matter? evidence from the housing market. *Journal of political Economy*, **113** (2), 376–424.

CHEN, L., LIU, H., LIU, W., LIU, J., LIU, K., SHANG, J., DENG, Y. and WEI, S. (2020a). Analysis of clinical features of 29 patients with 2019 novel coronavirus pneumonia. *Zhonghua jie he he hu xi za zhi= Zhonghua jiehe he huxi zazhi= Chinese journal of tuberculosis and respiratory diseases*, **43**, E005–E005.

CHEN, R., HU, B., LIU, Y., XU, J., YANG, G., XU, D. and CHEN, C. (2016). Beyond pm2. 5: the role of ultrafine particles on adverse health effects of air pollution. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1860** (12), 2844–2855.

—, LI, H., CAI, J., WANG, C., LIN, Z., LIU, C., NIU, Y., ZHAO, Z., LI, W. and KAN, H. (2018). Fine particulate air pollution and the expression of micrornas and circulating cytokines relevant to inflammation, coagulation, and vasoconstriction. *Environmental health perspectives*, **126** (1), 017007.

CHEN, T., WU, D., CHEN, H., YAN, W., YANG, D., CHEN, G., MA, K., XU, D., YU, H., WANG, H. *et al.* (2020b). Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *Bmj*, **368**.

CHETTY, R., FRIEDMAN, J. N., HILGER, N., SAEZ, E., SCHANZENBACH, D. W. and YAGAN, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly Journal of Economics*, **126** (4), 1593–1660.

—, — and ROCKOFF, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *The American Economic Review*, **104** (9), 2633–2679.

CILLÓNIZ, C., SHINYA, K., PENG, X., KORTH, M. J., PROLL, S. C., AICHER, L. D., CARTER, V. S., CHANG, J. H., KOBASA, D., FELDMANN, F. *et al.* (2009). Lethal influenza virus infection in macaques is associated with early dysregulation of inflammatory related genes. *PLoS pathogens*, **5** (10).

CLAY, K., LEWIS, J. and SEVERNINI, E. (2016). *Canary in a coal mine: Infant mortality, property values, and tradeoffs associated with mid-20th century air pollution*. Tech. rep., National Bureau of Economic Research.

—, — and — (2018). Pollution, infectious disease, and mortality: evidence from the 1918 spanish influenza pandemic. *The Journal of Economic History*, **78** (4), 1179–1209.

—, — and — (2019). What explains cross-city variation in mortality during the 1918 influenza pandemic? evidence from 438 us cities. *Economics & Human Biology*, **35**, 42–50.

— and TROESKEN, W. (2010). *Did Frederick Brodie discover the world's first environmental Kuznets curve? Coal smoke and the rise and fall of the London fog*. Tech. rep., National Bureau of Economic Research.

CLEMENS, M. A., LEWIS, E. G. and POSTEL, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. *American Economic Review*, **108** (6), 1468–87.

COHEN, A. J., BRAUER, M., BURNETT, R., ANDERSON, H. R., FROSTAD, J., ESTEP, K., BALAKRISHNAN, K., BRUNEKREEF, B., DANDONA, L., DANDONA, R. *et al.* (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *The Lancet*, **389** (10082), 1907–1918.

COLE, M., RAYNER, A. and BATES, J. (1997). The environmental kuznets curve: an empirical analysis. *Environment and Development Economics*, **2** (4), 401416.

COMMISSION, P. *et al.* (2003). Report of the task force on. *Identification of Districts for Wage and Self employment programmes, Technical Report New Delhi*.

COOK, J. and SHAH, M. (2019). Aggregate Effects from Public Works: Evidence from India, working paper.

CORREIA, S., GUIMARÃES, P. and ZYLKIN, T. (2019). Ppmlhdfe: Fast poisson estimation with high-dimensional fixed effects. *arXiv preprint arXiv:1903.01690*.

COUTURE, V., DINGEL, J. I., GREEN, A., HANDBURY, J. and WILLIAMS, K. (2020). *Location exposure indeex based on PlaceIQ data*. Tech. rep.

CROPPER, M., GAMKHAR, S., MALIK, K., LIMONOV, A. and PARTRIDGE, I. (2012). The health effects of coal electricity generation in india. *Resources for the Future Discussion Paper*, (12-25).

CUARESMA, J. C. and HEGER, M. (2019). Deforestation and economic development: Evidence from national borders. *Land Use Policy*, **84**, e347 – e353.

CUI, Y., ZHANG, Z.-F., FROINES, J., ZHAO, J., WANG, H., YU, S.-Z. and DETELS, R. (2003). Air pollution and case fatality of sars in the people's republic of china: an ecologic study. *Environmental Health*, **2** (1), 15.

CURRIE, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, **47** (1), 87–122.

DASGUPTA, A. (2017). Can the major public works policy buffer negative shocks in early childhood? evidence from andhra pradesh, india. *Economic Development and Cultural Change*, **65** (4), 767–804.

DASGUPTA, S., LAPLANTE, B., WANG, H. and WHEELER, D. (2002). Confronting the environmental kuznets curve. *Journal of Economic Perspectives*, **16** (1), 147–168.

DAVID, H. (2014). Skills, education, and the rise of earnings inequality among the other 99 percent. *Science*, **344** (6186), 843–851.

DAVIS, L. W. and GERTLER, P. J. (2015). Contribution of air conditioning adoption to future energy use under global warming. *Proceedings of the National Academy of Sciences*, **112** (19), 5962–5967.

DE CHAISEMARTIN, C. and DHAULTFŒUILLE, X. (2017). Fuzzy differences-in-differences. *The Review of Economic Studies*, **85** (2), 999–1028.

DEDOUSSI, I. C., EASTHAM, S. D., MONIER, E. and BARRETT, S. R. (2020). Premature mortality related to united states cross-state air pollution. *Nature*, **578** (7794), 261–265.

DEININGER, K. and LIU, Y. (2013). *Welfare and poverty impacts of India's national rural employment guarantee scheme: evidence from Andhra Pradesh*. The World Bank.

DELL, M., JONES, B. F. and OLKEN, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, **4** (3), 66–95.

DERYUGINA, T., HEUTEL, G., MILLER, N. H., MOLITOR, D. and REIF, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, **109** (12), 4178–4219.

DESCHÊNES, O. and GREENSTONE, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review*, **97** (1), 354–385.

— and GREENSTONE, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. *American Economic Journal: Applied Economics*, **3** (4), 152–185.

DESCHENES, O., GREENSTONE, M. and SHAPIRO, J. S. (2017). Defensive investments and the demand for air quality: Evidence from the nox budget program. *American Economic Review*, **107** (10), 2958–2989.

DESHINGKAR, P., JOHNSON, C. and FARRINGTON, J. (2005). State transfers to the poor and back: The case of the food-for-work program in india. *World Development*, **33** (4), 575–591.

DINDA, S. (2004). Environmental kuznets curve hypothesis: A survey. *Ecological Economics*, **49** (4), 431 – 455.

DOGO, H., BRANDON, C., HEGER, M., CHONABAYASHI, S., GASKELL, J., BANGALORE, M., NORMAN, T., LEE, J. J., SPENCER, P. and ELIZONDO, A. C. (2017). *Exploring Hidden Dimensions: Environmental and Natural Resource Aspects of Poverty*. Tech. rep.

DONALDSON, D. and STOREYGARD, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, **30** (4), 171–98.

EBENSTEIN, A., FAN, M., GREENSTONE, M., HE, G., YIN, P. and ZHOU, M. (2015). Growth, pollution, and life expectancy: China from 1991-2012. *American Economic Review*, **105** (5), 226–31.

—, —, —, — and ZHOU, M. (2017). New evidence on the impact of sustained exposure to air pollution on life expectancy from chinas huai river policy. *Proceedings of the National Academy of Sciences*, **114** (39), 10384–10389.

—, LAVY, V. and ROTH, S. (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.

EIA (2016). *Many natural gas-fired power plants under construction are near major shale plays*. Tech. rep., U.S. Energy Information Administration.

EMERICK, K. (2018). Agricultural productivity and the sectoral reallocation of labor in rural india. *Journal of Development Economics*, **135**, 488–503.

FEYRER, J., MANSUR, E. T. and SACERDOTE, B. (2017). Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review*, **107** (4), 1313–34.

FLA (2012). *TASK AND RISK MAPPING OF SUGARCANE PRODUCTION IN INDIA*. Tech. rep., The Fair Labor Association.

FORCE, A. D. T., RANIERI, V., RUBENFELD, G., THOMPSON, B., FERGUSON, N., CALDWELL, E. *et al.* (2012). Acute respiratory distress syndrome. *Jama*, **307** (23), 2526–2533.

FOSTER, A. D. and ROSENZWEIG, M. R. (2011). Are indian farms too small? mechanization, agency costs, and farm efficiency.

FOWLIE, M. and MULLER, N. (2019). Market-based emissions regulation when damages vary across sources: What are the gains from differentiation? *Journal of the Association of Environmental and Resource Economists*, **6** (3), 593–632.

FRANKEL, J. A. (2003). *The environment and globalization*. Tech. rep., National Bureau of Economic Research.

FROMM, M., LINDSEY, D. T., SERVRANCKX, R., YUE, G., TRICKL, T., SICA, R., DOUCET, P. and GODIN-BEEKMANN, S. (2010). The untold story of pyrocumulonimbus. *Bulletin of the American Meteorological Society*, **91** (9), 1193–1210.

GAIHA, R. (1997). Do rural public works influence agricultural wages? the case of the employment guarantee scheme in india. *Oxford Development Studies*, **25** (3), 301–314.

GARG, T., JAGNANI, M. and TARAZ, V. (2018). Temperature and human capital in india. *Available at SSRN 2941049*.

GATEBE, C., VARNAI, T., POUDYAL, R., ICHOKU, C. and KING, M. (2012). Taking the pulse of pyrocumulus clouds. *Atmospheric environment*, **52**, 121–130.

GEHRKE, E. (2013). Does the indian employment guarantee reduce households risk exposure? assessing the effects of the nrega on crop choice. *Unpublished manuscript, IZA–Institute of Labor Economics. Retrieved from http://www. iza. org/conference_files/worldb2013/gehrke_e8393. pdf*.

GERTLER, P., MARTINEZ, S. and RUBIO-CODINA, M. (2006). *Investing cash transfers to raise long term living standards*. The World Bank.

GILFILLAN, S. C. (1920). The coldward course of progress. *Political Science Quarterly*, **35** (3), 393–410.

GOI (2007). *NATIONAL RURAL EMPLOYMENT GUARANTEE ACT 2005 (NREGA): Report of the Second Year April 2006 March 2007*. Tech. rep., Ministry of Rural Development, Government of India.

— (2009). *GUIDELINES FOR IMPLEMENTATION OF WORKS ON INDIVIDUAL LAND UNDER NREGA*. Tech. rep., Ministry of Rural Development, Government of India.

GOLDIN, C. D. and KATZ, L. F. (2009). *The race between education and technology*. Harvard University Press.

GOODMAN-BACON, A. (2018). *Difference-in-differences with variation in treatment timing*. Tech. rep., National Bureau of Economic Research.

GORBALENYA, A., BAKER, S., BARIC, R. *et al.* (2020). The species severe acute respiratory syndrome-related coronavirus: classifying 2019-ncov and naming it sars-cov-2. *Nature Microbiology*, p. 1.

GRAFF ZIVIN, J., HSIANG, S. M. and NEIDELL, M. (2017). Temperature and human capital in the short-and long-run. *Journal of the Association of Environmental and Resource Economists*.

GRAFF-ZIVIN, J. and NEIDELL, M. (2012). The impact of pollution on worker productivity. *The American Economic Review*, **102** (7), 3652–3673.

GRAFF ZIVIN, J. and NEIDELL, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, **32** (1), 1–26.

GRAINGER, C., SCHREIBER, A. and CHANG, W. (2016). *How states comply with federal regulations: strategic ambient pollution monitoring*. Tech. rep., Working paper, University of Wisconsin-Madison.

GREENSTONE, M. and HANNA, R. (2014). Environmental regulations, air and water pollution, and infant mortality in india. *American Economic Review*, **104** (10), 3038–72.

— and JACK, B. K. (2013). Envirodevonomics: A research agenda for a young field. *National Bureau of Economic Research*.

—, KOPITS, E. and WOLVERTON, A. (2013). Developing a social cost of carbon for us regulatory analysis: A methodology and interpretation. *Review of Environmental Economics and Policy*, **7** (1), 23–46.

GROSSMAN, G. M. and KRUEGER, A. B. (1995). Economic growth and the environment. *The quarterly journal of economics*, **110** (2), 353–377.

GRUZIEVA, O., MERID, S. K., GREF, A., GAJULAPURI, A., LEMONNIER, N., BALLEREAU, S., GIGANTE, B., KERE, J., AUFFRAY, C., MELÉN, E. *et al.* (2017). Exposure to traffic-related air pollution and serum inflammatory cytokines in children. *Environmental health perspectives*, **125** (6), 067007.

GUPTA, A. and SPEARS, D. (2017). Health externalities of india's expansion of coal plants: Evidence from a national panel of 40,000 households. *Journal of environmental economics and management*, **86**, 262–276.

HAMMING, I., TIMENS, W., BULTHUIS, M., LELY, A., NAVIS, G. and VAN GOOR, H. (2004). Tissue distribution of ace2 protein, the functional receptor for sars coronavirus. a first step in understanding sars pathogenesis. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, **203** (2), 631–637.

HANNA, R. and OLIVA, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. *Journal of Public Economics*, **122**, 68–79.

HANUSHEK, E. A. and WOESSMANN, L. (2016). Knowledge capital, growth, and the east asian miracle. *Science*, **351** (6271), 344–345.

HARBAUGH, W. T., LEVINSON, A. and WILSON, D. M. (2002). Reexamining the empirical evidence for an environmental kuznets curve. *Review of Economics and Statistics*, **84** (3), 541–551.

HAUSMAN, C. and KELLOGG, R. (2015). *Welfare and distributional implications of shale gas*. Tech. rep., National Bureau of Economic Research.

HAUSMAN, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic perspectives*, **15** (4), 57–67.

HAUSMAN, J. A., HALL, B. H. and GRILICHES, Z. (1984). *Econometric models for count data with an application to the patents-R&D relationship*. Tech. Rep. 17, National Bureau of Economic Research Cambridge.

HE, L., LIN, A., CHEN, X., ZHOU, H., ZHOU, Z. and HE, P. (2019). Assessment of merra-2 surface pm2. 5 over the yangtze river basin: Ground-based verification, spatiotemporal distribution and meteorological dependence. *Remote Sensing*, **11** (4), 460.

HERRENDORF, B., ROGERSON, R. and VALENTINYI, K. (2013). *Growth and Structural Transformation*. Working Paper 18996, National Bureau of Economic Research.

HERSBACH, H. (2016). The era5 atmospheric reanalysis. In *AGU Fall Meeting Abstracts*.

HOFFMANN, M., KLEINE-WEBER, H., SCHROEDER, S., KRÜGER, N., HERRLER, T., ERICHSEN, S., SCHIERGENS, T. S., HERRLER, G., WU, N.-H., NITSCHE, A. *et al.* (2020). Sars-cov-2 cell entry depends on ace2 and tmprss2 and is blocked by a clinically proven protease inhibitor. *Cell*.

HOLLAND, S. P., MANSUR, E. T., MULLER, N. and YATES, A. J. (2018). *Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation*. Tech. rep., National Bureau of Economic Research.

HORNBECK, R. and NAIDU, S. (2014). When the levee breaks: black migration and economic development in the american south. *American Economic Review*, **104** (3), 963–90.

HSIANG, S. (2016). Climate econometrics. *Annual Review of Resource Economics*, **8**, 43–75.

—, OLIVA, P. and WALKER, R. (2018). The distribution of environmental damages. *Review of Environmental Economics and Policy*.

HSIANG, S. M., BURKE, M. and MIGUEL, E. (2013). Quantifying the influence of climate on human conflict. *Science*, **341** (6151), 1235367.

HUMPHRIES, J. (2013). The lure of aggregates and the pitfalls of the patriarchal perspective: a critique of the high wage economy interpretation of the british industrial revolution. *The Economic History Review*, **66** (3), 693–714.

HUNTINGTON, E. (1922). *Civilization and climate*. Yale University Press.

IMBERT, C. and PAPP, J. (2014). Short-term migration and rural workfare programs: Evidence from india. *Journal of the European Economic Association*.

— and — (2015). Labor market effects of social programs: Evidence from india's employment guarantee. *American Economic Journal: Applied Economics*, **7** (2), 233–63.

JAIN, M., MONDAL, P., GALFORD, G., FISKE, G. and DEFRIES, R. (2017). An automated approach to map winter cropped area of smallholder farms across large scales using modis imagery. *Remote Sensing*, **9** (6), 566.

JAIN, N., BHATIA, A. and PATHAK, H. (2014). Emission of air pollutants from crop residue burning in india. *Aerosol and Air Quality Research*, **14** (1), 422–430.

JAYACHANDRAN, S. (2009). Air quality and early-life mortality evidence from indonesias wildfires. *Journal of Human resources*, **44** (4), 916–954.

JHA, R., BHATTACHARYYA, S., GAIHA, R. and SHANKAR, S. (2009). capture of anti-poverty programs: An analysis of the national rural employment guarantee program in india. *Journal of Asian Economics*, **20** (4), 456–464.

JIANG, X.-Q., MEI, X.-D. and FENG, D. (2016). Air pollution and chronic airway diseases: what should people know and do? *Journal of thoracic disease*, **8** (1), E31.

JITENDRA, S. V., KUKRETI, I., PANDEY, K., NIYOGI, D. G. and MUKERJEE, P. (2017). India's burning issue of crop burning takes a new turn. *Down to Earth*.

JOHNSEN, R., LARIVIERE, J. and WOLFF, H. (2019). Fracking, coal, and air quality. *Journal of the Association of Environmental and Resource Economists*, **6** (5), 1001–1037.

—, LARIVIERE, J. S. and WOLFF, H. (2016). Estimating indirect benefits: fracking, coal and air pollution.

KAN, H., CHEN, B.-H., FU, C., YU, S.-Z. and MU, L. (2005). Relationship between ambient air pollution and daily mortality of sars in beijing. *Biomedical and Environmental Sciences*, **18** (1), 1–4.

KELLY, M., MOKYR, J. and GRÁDA, C. Ó. (2014). Precocious albion: a new interpretation of the british industrial revolution. *Annu. Rev. Econ.*, **6** (1), 363–389.

KJELLSTROM, T. and CROWE, J. (2011). Climate change, workplace heat exposure, and occupational health and productivity in central america. *International Journal of Occupational and Environmental Health*, **17** (3), 270–281.

KNITTEL, C. R., METAXOGLOU, K. and TRINDADE, A. (2015). *Natural gas prices and coal displacement: Evidence from electricity markets*. Tech. rep., National Bureau of Economic Research.

KORONTZI, S., MCCARTY, J., LOBODA, T., KUMAR, S. and JUSTICE, C. (2006). Global distribution of agricultural fires in croplands from 3 years of moderate resolution imaging spectroradiometer (modis) data. *Global Biogeochemical Cycles*, **20** (2).

KREMER, M., LEINO, J., MIGUEL, E. and ZWANE, A. P. (2011). Spring cleaning: Rural water impacts, valuation, and property rights institutions. *The Quarterly Journal of Economics*, **126** (1), 145–205.

— and WILLIS, J. (2016). Guns, latrines, and land reform: Dynamic pigouvian taxation. *American Economic Review*, **106** (5), 83–88.

KUZNETS, S. (1973). Modern economic growth: findings and reflections. *The American economic review*, **63** (3), 247–258.

LEE, J., PERI, G. and YASENOV, V. (2017). *The Employment Effects of Mexican Repatriations: Evidence from the 1930's*. Tech. rep., National Bureau of Economic Research.

LI, X., XU, S., YU, M., WANG, K., TAO, Y., ZHOU, Y., SHI, J., ZHOU, M., WU, B., YANG, Z. *et al.* (2020a). Risk factors for severity and mortality in adult covid-19 inpatients in wuhan. *Journal of Allergy and Clinical Immunology*.

LI, Y. (2017). *Land Inequality and the Provision of Public Works-Evidence from National Rural Employment Guarantee Scheme*. Tech. rep., Unpublished Job Market Paper.

—, ZHOU, W., YANG, L. and YOU, R. (2020b). Physiological and pathological regulation of ace2, the sars-cov-2 receptor. *Pharmacological Research*, p. 104833.

LIN, W. and WOOLDRIDGE, J. M. (2019). Testing and correcting for endogeneity in nonlinear unobserved effects models. In *Panel Data Econometrics*, Elsevier, pp. 21–43.

LINN, J. and MUEHLENBACHS, L. (2018). The heterogeneous impacts of low natural gas prices on consumers and the environment. *Journal of Environmental Economics and Management*, **89**, 1–28.

LIU, J., ZHOU, J., YAO, J., ZHANG, X., LI, L., XU, X., HE, X., WANG, B., FU, S., NIU, T. *et al.* (2020). Impact of meteorological factors on the covid-19 transmission: A multi-city study in china. *Science of The Total Environment*, p. 138513.

LIU, T., MARLIER, M. E., DEFRIES, R. S., WESTERVELT, D. M., XIA, K. R., FIORE, A. M., MICKLEY, L. J., CUSWORTH, D. H. and MILLY, G. (2018). Seasonal impact of regional outdoor biomass burning on air pollution in three indian cities: Delhi, bengaluru, and pune. *Atmospheric environment*, **172**, 83–92.

LOHAN, S. K., JAT, H., YADAV, A. K., SIDHU, H., JAT, M., CHOUDHARY, M., PETER, J. K. and SHARMA, P. (2018). Burning issues of paddy residue management in north-west states of india. *Renewable and Sustainable Energy Reviews*, **81**, 693–706.

LONGHURST, J. and CONLAN, D. (1970). Changing air quality in the greater manchester conurbation. *WIT Transactions on Ecology and the Environment*, **3**.

LUKS, A. M., FREER, L., GRISSOM, C. K., MCINTOSH, S. E., SCHOENE, R. B., SWENSON, E. R. and HACKETT, P. H. (2020). Covid-19 lung injury is not high altitude pulmonary edema. *High Altitude Medicine & Biology*.

LUTZ, W., MUTTARAK, R. and STRIESSNIG, E. (2014). Universal education is key to enhanced climate adaptation. *Science*, **346** (6213), 1061–1062.

MACCINI, S. and YANG, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, **99** (3), 1006–26.

MANDAL, K. G., MISRA, A. K., HATI, K. M., BANDYOPADHYAY, K. K., GHOSH, P. K. and MOHANTY, M. (2004). Rice residue-management options and effects on soil properties and crop productivity. *Journal of Food Agriculture and Environment*, **2**, 224–231.

MANI, A., MULLAINATHAN, S., SHAFIR, E. and ZHAO, J. (2013). Poverty impedes cognitive function. *Science*, **341** (6149), 976–980.

MANISALIDIS, I., STAVROPOULOU, E., STAVROPOULOS, A. and BEZIRTZOGLOU, E. (2020). Environmental and health impacts of air pollution: A review. *Frontiers in Public Health*, **8**.

MARLON, J. R., BARTLEIN, P. J., CARCAILLET, C., GAVIN, D. G., HARRISON, S. P., HIGUERA, P. E., JOOS, F., POWER, M. and PRENTICE, I. (2008). Climate and human influences on global biomass burning over the past two millennia. *Nature Geoscience*, **1** (10), 697.

MEHTA, P., MCAULEY, D. F., BROWN, M., SANCHEZ, E., TATTERSALL, R. S. and MANSON, J. J. (2020). Covid-19: consider cytokine storm syndromes and immunosuppression. *The Lancet*, **395** (10229), 1033–1034.

MINISTRY OF FINANCE, G. (2018). *Economic Survey 2017-18 (Volume I and Volume II)*. OUP India.

MISRA, D. P., AGARWAL, V., GASPARYAN, A. Y. and ZIMBA, O. (2020). Rheumatologists perspective on coronavirus disease 19 (covid-19) and potential therapeutic targets. *Clinical Rheumatology*, pp. 1–8.

MO, Z., FU, Q., ZHANG, L., LYU, D., MAO, G., WU, L., XU, P., WANG, Z., PAN, X., CHEN, Z. *et al.* (2018). Acute effects of air pollution on respiratory disease mortalities and outpatients in southeastern china. *Scientific reports*, **8** (1), 1–9.

MONTESQUIEU, C. S. (1750). The spirit of the laws.

MORRISON, K. D. (1994). Monitoring regional fire history through size-specific analysis of microscopic charcoal: The last 600 years in south india. *Journal of Archaeological Science*, **21** (5), 675–685.

MULLER, N. Z. and MENDELSOHN, R. (2007). Measuring the damages of air pollution in the united states. *Journal of Environmental Economics and Management*, **54** (1), 1–14.

MUNIYAPPA, R. and GUBBI, S. (2020). Covid-19 pandemic, corona viruses, and diabetes mellitus.

MURALIDHARAN, K., NIEHAUS, P. and SUKHTANKAR, S. (2016). Building state capacity: Evidence from biometric smartcards in india. *American Economic Review*, **106** (10), 2895–2929.

—, — and — (2017). *General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India*. Working Paper 23838, National Bureau of Economic Research.

NAAS (2017). *Innovative Viable Solution to Rice Residue Burning in Rice-Wheat Cropping System through Concurrent Use of Super Straw Management System-fitted Combines and Turbo Happy Seeder.* Tech. rep., National Academy of Agricultural Sciences, New Delhi.

NAIR, M., ARIANA, P., OHUMA, E. O., GRAY, R., DE STAVOLA, B. and WEBSTER, P. (2013). Effect of the mahatma gandhi national rural employment guarantee act (mgnrega) on malnutrition of infants in rajasthan, india: a mixed methods study. *PloS one*, **8** (9), e75089.

NASA (2017a). *Global Fire Emissions Indicators, Country-Level Tabular Data: 1997-2015*. Tech. rep., Center for International Earth Science Information Network (CIESIN), Columbia University.

— (2017b). *India Annual Winter Cropped Area, 2001-2016*. Tech. rep., Center for International Earth Science Information Network (CIESIN), Columbia University.

NEL, A. (2005). Air pollution-related illness: effects of particles. *Science*, **308** (5723), 804–806.

NICHOLS, A. *et al.* (2010). Regression for nonnegative skewed dependent variables. In *BOS10 Stata Conference*, Stata Users Group, vol. 2, pp. 15–16.

NIEHAUS, P. and SUKHTANKAR, S. (2013). Corruption dynamics: The golden goose effect. *American Economic Journal: Economic Policy*, **5** (4), 230–69.

OKEEFE, P. (2005). Workfare programs in india and internationally: Note on issues and experience. *The World Bank Office, New Delhi, India*.

PARK, J., BANGALORE, M., HALLEGATTE, S. and SANDHOEFNER, E. (2018). Households and heat stress: estimating the distributional consequences of climate change. *Environment and Development Economics*, **23** (3), 349–368.

PARK, R. J. (forthcoming). Hot temperature and high stakes performance. *Journal of Human Resources*.

—, GOODMAN, J., HURWITZ, M. and SMITH, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*.

PISCHKE, J.-S. (2018). Weak instruments.

PRASAD, R., GANGAIAH, B. and AIPE, K. (1999). Effect of crop residue management in a rice–wheat cropping system on growth and yield of crops and on soil fertility. *Experimental Agriculture*, **35** (4), 427–435.

PRASANN, A. (2016). *The Spillover Effects of Public Works on Labor Allocation and Wages: Evidence from National Rural Employment Guarantee Scheme, India*. Tech. rep.

PULLABHOTLA, H. (2018). Fires, wind, and smoke: Air pollution and infant mortality. *Job Market Paper. Available here*.

PYNE, S. J. (2019). *Fire: a brief history*. University of Washington Press.

— and GOLDAMMER, J. G. (1997). The culture of fire: an introduction to anthropogenic fire history. In *Sediment Records of Biomass Burning and Global Change*, Springer, pp. 71–114.

RABOTYAGOV, S. S., CAMPBELL, T. D., WHITE, M., ARNOLD, J. G., ATWOOD, J., NORFLEET, M. L., KLING, C. L., GASSMAN, P. W., VALCU, A., RICHARDSON, J. *et al.* (2014). Cost-effective targeting of conservation investments to reduce the northern gulf of mexico hypoxic zone. *Proceedings of the National Academy of Sciences*, **111** (52), 18530–18535.

RAGHUNATHAN, K. and HARI, S. (2014). Providing more than just employment? evidence from the nrega in india. *Unpublished Working Paper*.

RANDAZZO, T., DE CIAN, E., EURO-MEDITERRANEO, C. and MISTRY, M. (2020). Climate change impacts on household electricity expenditure: the contribution of air conditioning in oecd countries. *Working Paper*.

RANGEL, M. A. and VOGL, T. (2016). *Agricultural fires and infant health*. Tech. rep., National Bureau of Economic Research.

RANSON, M. (2014). Crime, weather, and climate change. *Journal of environmental economics and management*, **67** (3), 274–302.

RAO, P. P., PANDEY, L., JAGADEESH, E., DEB, U. K., JAIN, R. and BASU, K. (2012). Meso-level database coverage and insights village dynamics in south asia.

REARDON, S. F., KALOGRIDES, D. and SHORES, K. (2017). The geography of racial/ethnic test score gaps. cepa working paper no. 16-10. *Stanford Center for Education Policy Analysis*.

RIENECKER, M. M., SUAREZ, M. J., GELARO, R., TODLING, R., BACMEISTER, J., LIU, E., BOSILOVICH, M. G., SCHUBERT, S. D., TAKACS, L., KIM, G.-K. *et al.* (2011). Merra: Nasas modern-era retrospective analysis for research and applications. *Journal of climate*, **24** (14), 3624–3648.

RODGERS, H. (1960). The lancashire cotton industry in 1840. *Transactions and Papers (Institute of British Geographers)*, (28), 135–153.

ROTH, J. (2018). *Pre-test with Caution Event-study Estimates After Testing for Parallel Trends*. Tech. rep., Working Paper.

ROWLAND, T. (2008). Thermoregulation during exercise in the heat in children: old concepts revisited. *Journal of Applied Physiology*, **105** (2), 718–724.

RUE, D., WEBSTER, D. and TRAVERSE, A. (2002). Late holocene fire and agriculture in the copan valley, honduras. *Ancient Mesoamerica*, **13** (2), 267–272.

SACKS, W. J., DERYNG, D., FOLEY, J. A. and RAMANKUTTY, N. (2010). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, **19** (5), 607–620.

SANCHE1, S., LIN, Y. T., XU, C., ROMERO-SEVERSON, E., HENGARTNER, N. and KE, R. (2020). High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*.

SANCHEZ, F. M. A., RESTUCCIA, D. and RUD, J. P. (2019). *Are Small Farms Really more Productive than Large Farms?* Tech. rep., National Bureau of Economic Research.

SAWYER, J. (2019). *Estimating losses when cornstalk fields are accidentally burnt*. Tech. rep., Iowa State University Extension.

SCHLENKER, W., HANEMANN, W. M. and FISHER, A. C. (2006). The impact of global warming on us agriculture: an econometric analysis of optimal growing conditions. *The Review of Economics and Statistics*, **88** (1), 113–125.

— and LOBELL, D. B. (2010). Robust negative impacts of climate change on african agriculture. *Environmental Research Letters*, **5** (1), 014010.

SEATON, A., GODDEN, D., MACNEE, W. and DONALDSON, K. (1995). Particulate air pollution and acute health effects. *The lancet*, **345** (8943), 176–178.

SEPPANEN, O., FISK, W. J. and LEI, Q. (2006). Effect of temperature on task performance in office environment. *Lawrence Berkeley National Laboratory*.

SETTI, L., PASSARINI, F., DE GENNARO, G., DI GILIO, A., PALMISANI, J., BUONO, P., FORNARI, G., PERRONE, M. G., PIAZZALUNGA, A., BARBIERI, P., RIZZO, E. and MIANI, A. (2020). *Evaluation of the potential relationship between Particulate Matter (PM) pollution and COVID-19 infection spread in Italy*. Tech. rep., mimeo.

SHAH, A. and MEHTA, A. K. (2008). Experience of the maharashtra employment guarantee scheme: are there lessons for nregs? *Available at SSRN 1538914*.

SHAH, M. and STEINBERG, B. M. (2015). *Workfare and human capital investment: Evidence from India*. Tech. rep., National Bureau of Economic Research.

— and — (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy*, **125** (2), 527–561.

SHARMA, A. (2015). Employment guarantee scheme and child health outcome: Evidence from mgnrega using propensity score matching analysis. *Available at SSRN 2708549*.

SHI, S., QIN, M., SHEN, B., CAI, Y., LIU, T., YANG, F., GONG, W., LIU, X., LIANG, J., ZHAO, Q. *et al.* (2020). Association of cardiac injury with mortality in hospitalized patients with covid-19 in wuhan, china. *JAMA cardiology*.

SHYAMSUNDAR, P., SPRINGER, N., TALLIS, H., POLASKY, S., JAT, M., SIDHU, H., KRISHNAPRIYA, P., SKIBA, N., GINN, W., AHUJA, V. *et al.* (2019). Fields on fire: Alternatives to crop residue burning in india. *Science*, **365** (6453), 536–538.

SINGER, G., ZIVIN, J. G., NEIDELL, M. and SANDERS, N. (2020). Air pollution increases influenza hospitalizations. *medRxiv*.

SINGH, P., DEY, S., CHOWDHURY, S. and BALI, K. (2019). Early life exposure to outdoor air pollution: Effect on child health in india.

SMIL, V. (1999). Crop residues: Agriculture's largest harvest: Crop residues incorporate more than half of the world's agricultural phytomass. *Bioscience*, **49** (4), 299–308.

SMITH, P., BUSTAMANTE, M., AHAMMAD, H., CLARK, H., DONG, H., ELSIDDIG, E. A., HABERL, H., HARPER, R., HOUSE, J., JAFARI, M. *et al.* (2014). Agriculture, forestry and other land use (afolu). In *Climate change 2014: mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press.

SOLOMON, S. (2016). Sugarcane production and development of sugar industry in india. *Sugar Tech*, **18** (6), 588–602.

SOLOW, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, pp. 312–320.

STAVINS, R. N. (1996). Correlated uncertainty and policy instrument choice. *Journal of environmental economics and management*, **30** (2), 218–232.

STERN, D. I. (2018). The environmental kuznets curve. *Companion to Environmental Studies*, pp. 49–54.

STOCK, J. H. and YOGO, M. (2002). *Testing for weak instruments in linear IV regression*. Tech. rep., National Bureau of Economic Research.

SUKHTANKAR, S. (2016). Indias national rural employment guarantee scheme: What do we really know about the worlds largest workfare program?

SULLIVAN, D. M. (2016). The true cost of air pollution: Evidence from house prices and migration. *Harvard University*.

— and KRUPNICK, A. (2018). Using satellite data to fill the gaps in the us air pollution monitoring network. *Resources for the Future Working Paper*, pp. 18–21.

TAN, W. C., QIU, D., LIAM, B. L., NG, T. P., LEE, S. H., VAN EEDEN, S. F., D'YACHKOVA, Y. and HOGG, J. C. (2000). The human bone marrow response to acute air pollution caused by forest fires. *American journal of respiratory and critical care medicine*, **161** (4), 1213–1217.

TAN-SOO, J.-S. and PATTANAYAK, S. K. (2019). Seeking natural capital projects: Forest fires, haze, and early-life exposure in indonesia. *Proceedings of the National Academy of Sciences*, **116** (12), 5239–5245.

THOMAS, M. (2015). Guaranteeing children a better life? the impact of the mahatma gandhi national rural employment guarantee act on child health and learning in rural india. *University of Texas Austin Working Paper*.

TISONCIK, J. R., KORTH, M. J., SIMMONS, C. P., FARRAR, J., MARTIN, T. R. and KATZE, M. G. (2012). Into the eye of the cytokine storm. *Microbiol. Mol. Biol. Rev.*, **76** (1), 16–32.

UPPAL, V. (2009). Is the nregs a safety net for children. *Young Lives Student Paper, Oxford: Young Lives*.

VAN DER WERF, G. R., RANDERSON, J. T., GIGLIO, L., COLLATZ, G. J., KASIBHATLA, P. S. and ARELLANO JR, A. F. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmospheric Chemistry and Physics*, **6** (11), 3423–3441.

VAN DONKELAAR, A., MARTIN, R. V., LI, C. and BURNETT, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental science & technology*, **53** (5), 2595–2611.

VAN EEDEN, S. F., TAN, W. C., SUWA, T., MUKAE, H., TERASHIMA, T., FUJII, T., QUI, D., VINCENT, R. and HOGG, J. C. (2001). Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (pm10). *American journal of respiratory and critical care medicine*, **164** (5), 826–830.

VASILICA, S., FÎNTÎNERU, G. and MIHALACHE, M. (2014). Multicriteria analysis of the effects of field burning crop residues. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, **42** (1), 255–262.

VOORHEIS, J. *et al.* (2017). *Air quality, human capital formation and the long-term effects of environmental inequality at birth*. Tech. rep., Center for Economic Studies, US Census Bureau.

WANG, X., YAMAUCHI, F., HUANG, J. and ROZELLE, S. (2018). What constrains mechanization in chinese agriculture? role of farm size and fragmentation. *China Economic Review*.

WEITZMAN, M. L. (1974). Prices vs. quantities. *The review of economic studies*, **41** (4), 477–491.

WHO (2003). *Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)*. Tech. rep., World Health Organization.

WILEBORE, B., VOORS, M., BULTE, E. H., COOMES, D. and KONTOLEON, A. (2019). Unconditional transfers and tropical forest conservation: Evidence from a randomized control trial in sierra leone. *American Journal of Agricultural Economics*, **101** (3), 894–918.

WOESSMANN, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *The Journal of Economic Perspectives*, **30** (3), 3–31.

WOOLDRIDGE, J. M. (1997). Quasi-likelihood methods for count data. *Handbook of Applied Econometrics Volume 2: Microeconomics*, pp. 321–368.

— (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics*, **90** (1), 77–97.

WU, X., NETHERY, R. C., SABATH, B. M., BRAUN, D. and DOMINICI, F. (2020). Exposure to air pollution and covid-19 mortality in the united states. *medRxiv*.

WU, Z. and MCGOOGAN, J. M. (). Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention. *Jama*.

YADAV, R. (2007). Mechanization of sugar cane production in india. In *Proceedings of the XXVI ISSCT Congress held in Durban in July/August*, vol. 7.

YAMAUCHI, F. (2016). Rising real wages, mechanization and growing advantage of large farms: Evidence from indonesia. *Food Policy*, **58**, 62–69.

YANG, S., HE, H., LU, S., CHEN, D. and ZHU, J. (2008). Quantification of crop residue burning in the field and its influence on ambient air quality in suqian, china. *Atmospheric Environment*, **42** (9), 1961–1969.

YONGJIAN, Z., JINGU, X., FENGMING, H. and LIQING, C. (2020). Association between short-term exposure to air pollution and covid-19 infection: Evidence from china. *Science of The Total Environment*, p. 138704.

YOUNG, A. (2018). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics*, **134** (2), 557–598.

ZHANG, T. and WOOSTER, M. (2016). Small fire detection algorithm development using viirs 375m imagery: application to agricultural fires in eastern china. In *EGU General Assembly Conference Abstracts*, vol. 18.

ZHENG, S. and KAHN, M. E. (2017). A new era of pollution progress in urban china? *Journal of Economic Perspectives*, **31** (1), 71–92.

ZHOU, P., YANG, X.-L., WANG, X.-G., HU, B., ZHANG, L., ZHANG, W., SI, H.-R., ZHU, Y., LI, B., HUANG, C.-L. *et al.* (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, **579** (7798), 270–273.

ZHU, N., ZHANG, D., WANG, W., LI, X., YANG, B., SONG, J., ZHAO, X., HUANG, B., SHI, W., LU, R. *et al.* (2020). A novel coronavirus from patients with pneumonia in china, 2019. *New England Journal of Medicine*.

ZIMMERMANN, L. (2013). Why guarantee employment? evidence from a large indian public-works program. *Unpublished manuscript*.

Appendix A

Appendix to Chapter 1

A.1 Model proofs

Proposition 1:

By assumption:

1. $G(L)$ and $F(L)$ are continuous
2. $F(L) > G(L)$ for all L
3. $0 < F^{'}(L) < G^{'}(L)$
4. $F^{''}(L) < 0, G^{''}(L) < 0$
5. $F^{'-1}(x) < G^{'-1}(x)$

Farmers face the adoption condition $AF\Big(L\Big) - wL - \big[AG\Big(L\Big) - wL\big] = K$. Solving for A and substituting for farmers optimal labor choices in F and G yields the equation for $\hat{A}$ in the text:

$$\hat{A} = \frac{K - wL^{*} + wL^{**}}{F(L^{**}) - G(L^{*})}$$

The assumptions above imply $F(L^{**}) - G(L^{*}) > 0$ and that $wL^{**} - wL^{*} < 0$ so it is clear that $\hat{A}$ is increasing in K and decreasing in w.

To see that $\hat{A}$ is unique note that:

1. $\partial \pi^F / \partial A > \partial \pi^G / \partial A$ and $\partial^2 \pi^F / \partial A \partial A = \partial^2 \pi^G / \partial A \partial A$

2. K is weakly positive so that there exists some A where $\pi^F \leq \pi^G$.

Then by the intermediate value theorem there exists a $\hat{A}$ that defines the point at which farmers are indifferent between $F(L)$ and $G(L)$ and it is unique. That also implies, because e is monotonically increasing in the number of farmers who choose $F(L)$, that e is decreasing in $\hat{A}$.

Proposition 2:

By assumption:

1. $A = A_D + \mu$

2. μ has a constant single-peaked distribution $H(\mu)$ with mean zero

3. $\hat{A}$ is everywhere greater than A_D

Since the level of pollution is increasing in the number of farmers above $\hat{A}$ the level of pollution e in any given district D will be a function of the distribution of μ and $\hat{A}$ such that $e \equiv e(1 - H[\hat{A}])$. As A_D approaches $\hat{A}$ from below for a constant distribution of μ the area $1 - H(\hat{A})$ is strictly increasing. Further, I've shown that $\hat{A}$ is declining in w. Assume some increase in w that results in an ϵ shift down in $\hat{A}$. Then for a $A_D < A_{D'}$, for a single peaked distribution, $H(1 - [\hat{A} - \epsilon], A_D) - H(1 - \hat{A}, A_D) < H(1 - [\hat{A} - \epsilon], A_{D'}) - H(1 - \hat{A}, A_{D'})$ so the change in pollution moving from $\hat{A}$ to $\hat{A} - \epsilon$ is increasing in A_D.

A.2 Fires and agriculture

Fire has been used by humans to manage landscapes for at least 40,000 years (Pyne and Goldammer, 1997) and in the cultivation of corn for at least 5,000 years (Rue *et al.*, 2002). Sediment cores indicate that fire was used to manage agricultural fields in India at least 600 years ago (Morrison, 1994). The use of fires in agriculture is still widespread today in both developed and developing countries. Historically fires were used to clear land for planting in a swidden (“slash-and-burn") style of agriculture and its use evolved over time to include preparing fields for harvest and clearing residue to prepare fields for re-planting (Pyne, 2019).

Analysis of satellite imagery suggests that there are roughly 1.5 million fires annually with the largest number in Russia (Korontzi *et al.*, 2006). The widespread use of fires in Russia, Eastern Europe (namely the Ukraine) and North America mean that in absolute terms fire use in agriculture is more common in the developed world than the developing world (Cassou, 2018). On a per hectare basis, however, African countries are the most frequent users of fire. This is in part due to declines in recent years in the use of fires in North America and the European Union (Marlon *et al.*, 2008), driven in part by increasing regulation around the practice. It should be noted, however, that the use of fires to clear crop residue was widespread in California and Western Canada until the mid-1990s when increasing concerns over air pollution led to regulation to eliminate the practice (Cassou, 2018).

A.2.1 Agricultural mechanization in India and fires

Rice, wheat, and sugarcane are the three crops most associated with the use of fire on cropland in India. Of these, the harvest of rice and wheat can be mechanized using existing technology present in India (Yadav, 2007; Solomon, 2016). Mechanizing the rice and wheat harvest is done using combines, which data from the Agricultural Input Survey shows are present throughout India prior to the implementation of NREGA but tend to be present at higher levels in areas with higher use of fire.

The relationship between mechanization and the use of fires is driven by the fact that harvesting with a combine leaves more residue on the field than harvesting by hand (Yang *et al.*, 2008). Specifically, combine harvesting leaves stocks that tend to be around 30-40cm as opposed to the 10-15cm that harvesting by hand leaves. The higher stalks interfere with the ability of farmers to plant the following season’s crops and must be removed to facilitate planting (Jain *et al.*, 2014; Smil, 1999; Cassou, 2018; Bhuvaneshwari *et al.*, 2019).[1] Burning is the least expensive way of dealing with this residue (Cassou, 2018), as interviews with farmers indicate: “Ankit Choyal Jat...offers an answer. ’If I can clear my

[1] No-till agriculture is an alternative to clearing the residue after a combine harvest. In this practice the standing residue is plowed back into the field. However, it requires specialized planting equipment that is both expensive and not widely present in India (Jain *et al.*, 2014; Bhuvaneshwari *et al.*, 2019).

farm using a one-rupee matchbox, why will I spend thousands? (Jitendra *et al.*, 2017)'

While combines were used in harvesting prior to the implementation of NREGA, the shock that NREGA provided to agricultural labor markets may have led to a substantial increase in their use. The only existing study of the impact of NREGA on agricultural production processes that I am aware of Bhargava (2014) shows that the use of labor saving animal-based production processes increased after the implementation of NREGA. Data from the Agricultural Input Survey shows a substantial increase in the average number of combines from 2006 to 2011 (figure A.1) – an increase that is much larger both in absolute and percentage terms than the increase from 2001 to 2006. Such an increase is consistent with broad state level trends from the Cost of Cultivation survey that show a decline in the amount of labor used in agriculture and an increase in the amount spent on machine inputs to production over the period that NREGA was implemented.

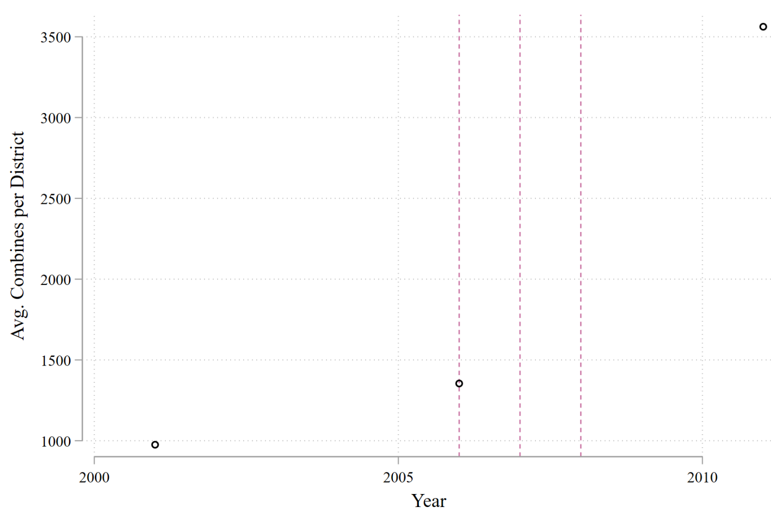
The notion that farmers responded to the impact of NREGA on agricultural labor markets by mechanizing is supported by their own statements as well. Media interviews with farmers from areas that have seen the largest increase in burning frequently include quotes like the following (Jitendra *et al.*, 2017):

> Hari Ram Karore, a 71-year-old farmer who owns more than 10 hectares (ha) in the same village, says, We started using combine harvester machines to tide over the labour scarcity. The machine finishes the task of reaping, threshing and winnowing in a few hours and is also economical, he adds.
>
> Residents of villages around Kota say that mechanisation has killed the practice of using wheat stalk and straw as fodder, and burning is the only way out. The cuttings left by the machines are too sharp. Not only do they injure us, even animals find it difficult to graze on, says Shital Devi.

These quotes are supported by more formal work interviewing farmers (FLA, 2012) that indicates farmers believe the supply of unskilled agricultural labor has declined as a result of NREGA. This is consistent with survey data from the 2011 India Human Development Survey that suggests a large share ($\approx$33%) of surveyed villages believe there to be fewer agricultural laborers than in the past and individual households report a 30% decline in the hours worked in agriculture. Fewer households in the IHDS report traveling for work after NREGA as well.

Broadly, the hypothesis that NREGA may have reduced the supply of unskilled labor to agriculture is consistent with the existing literature on its impacts Imbert and Papp (2014, 2015) that show small reductions in labor in the private sector and possible reductions in migration. The estimated effects in the literature are smaller than those suggested by interviews with farmers but Asher and Novosad (2018) suggest that survey-based estimates of the amount of labor in agriculture may underestimate the impacts of policies because of respondents misunderstanding the intention of the questions. Reductions in the supply of agricultural labor in response to NREGA are also consistent with the documented increase in local consumption and the effect that this has on local labor markets measured by Emerick (2018). It is also true that the research on NREGA has not reached a conclusion about its overall impact on the share of labor in agriculture, the productivity of that labor, or the production practices employed in agriculture (Sukhtankar, 2016). These remain important areas of research.

FIGURE A.1: PRESENCE OF COMBINES OVER TIME



NOTES*: The average number self-powered combines by district over time. Data scrapped from the Indian Agricultural Input survey conducted in 2001, 2006 and 2011. The dashed lines indicate the phases of NREGA.*

A.3 NREGA details

A.3.1 A brief history of Indian workfare

The NREGA program is the latest in a succession of work-based anti-poverty programs in India dating back to at least the British Raj (Imbert and Papp, 2015). The most notable program prior to NREGA was the Maharashtra Employment Guarantee Scheme (MEGS) introduced in 1977 by the Government of Maharashtra (Shah and Mehta, 2008). MEGS was not a national program but much of the design of MEGS was incorporated into NREGA. Like NREGA the aim of MEGS was to provide employment to rural residents, focused on labor-intensive work, and targeted the formation of public goods. The number of person-days of work that it generated reached an early peak 1980 and declined through the 1990s before climbing again through the early 2000s. Variation in the number of person-days supplied through MEGS appears to be tied to changes in the wage schedule and declines in the level of activist support for the program (Shah and Mehta, 2008). Of particular relevance to this study and NREGA, estimates suggest that MEGS increased agricultural labor wages by around 18%(Gaiha, 1997).

After independence, the national government experimented with a number of national rural workfare programs. A series of small-scale and pilot programs in the 1960s and 1970s were rolled into the national Food for Work Programme (FWP) in 1977. Despite receiving significant investment, there is little evidence that the FWP had a meaningful impact on reducing rural unemployment, due, at least in part, to poor implementation and exclusion of the poorest citizens (Deshingkar *et al.*, 2005). The FWP became the National Rural Employment Program in the 1980s. In 2001, the Sampoorn Grameen Rozgar Yojana (SGRY) program combined this with several existing poverty alleviation programs and rural infrastructure programs to consolidate effort and provide additional employment, food security, and infrastructure in rural areas (GOI, 2007). Wages were paid in a combination of cash and food supplies. By 2008 the SGRY program had been fully merged with NREGA.

A.3.2 Details of NREGA

The law creating NREGA was passed in September 2005 and the program was implemented in the first districts in February 2006. Figure A.2 shows the districts included in each phase. I map districts on 2001 geographies and apportion all data to the districts as they existed in 2001.

To participate in NREGA households obtain job cards from their local districts and then are able to apply for work whenever they would like. The district office is to provide work within 5 km of their house within 15 days of receiving their applications. The district must pay an unemployment allowance in cash if they fail to provide employment. Wages are to be paid at a statutorily set minimum wage that is not less than 60 Rs/day.

NREGA project lists are prepared at the district level and projects must be in one of the permitted categories. Those are: water conservation, drought proofing, flood protection, land development, minor irrigation, and rural connectivity (GOI, 2007).[2] All projects must have a ratio of labor expense to material expense of at least 60/40 and the use of contractors and machinery is not allowed.[3] The cost of NREGA is split between the central government and state governments but, crucially, the full cost of unskilled labor is borne by the central government. State governments bear none of the cost of unskilled labor and 25% of the cost of skilled labor and materials, giving states an incentive to prioritize projects that use a greater share of unskilled labor.

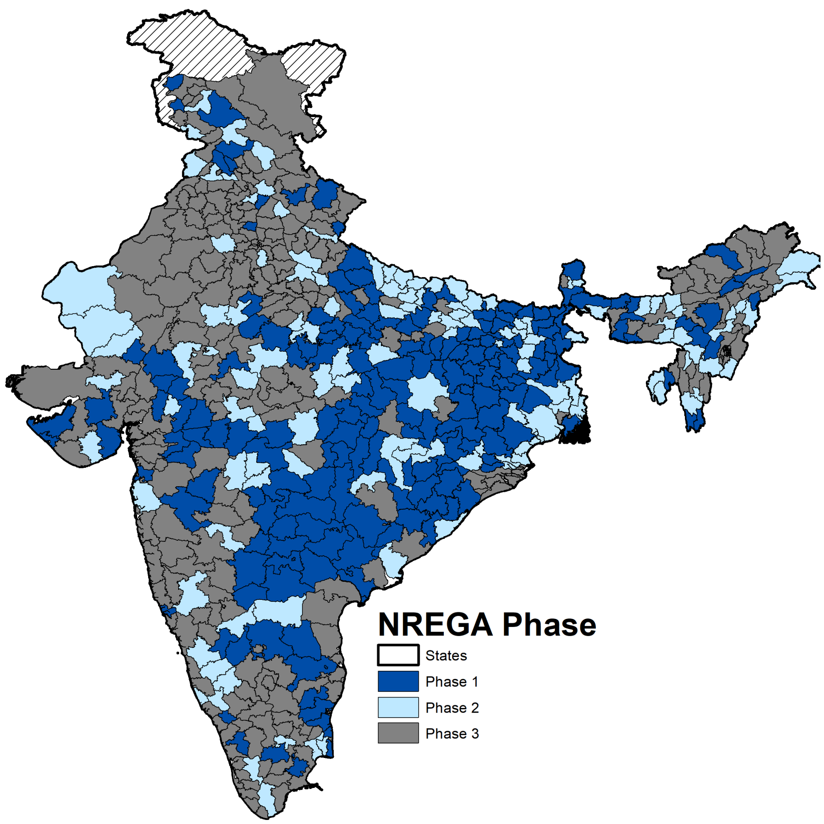
The scale of the program is remarkable. It is generally agreed that NREGA is the largest workfare/rural poverty reduction program in the world (Ambasta *et al.*, 2008). By 2014, 121 million job cards had been registered. In 2009-2010 there were 2.8 billion person-days of work conducted under the program (Sukhtankar, 2016). Participation appears to have grown steadily from implementation in 2006 to around 2013. Roughly 11% of the world's population is covered by the program(Niehaus and Sukhtankar, 2013). In principle, NREGA marked a shift from existing anti-poverty programs by being demand, as opposed to supply, driven (GOI, 2007). In practice, implementation challenges and

[2]This work can occur on private land if it is owned by a member of a scheduled caste or tribe. The operational guidelines of NREGA were modified in 2009 to allow for work to also be conducted on private land if the total holdings of the owner placed them in the "small" or "marginal" categories and the owner participated in the work GOI (2009).

[3]The use of contractors and machinery was understood to be an obstacle to the effectiveness of previous programs in providing pro-poor benefits (Ambasta *et al.*, 2008).

state capacity may have limited the extent to which it was able to fully meet demand (Sukhtankar, 2016; Niehaus and Sukhtankar, 2013). By March 2007, demand for work exceeded supply in at least 30% of the states with significant demand (GOI, 2007). Despite this, NREGA provided an average of three times the number of person-days of employment in its first years as SGRY provided in its final years.

FIGURE A.2: MAP OF NREGA DISTRICTS BY PHASE OF ROLL-OUT



NOTES: *Indian districts are mapped by their NREGA phase. Phase one districts received NREGA in February 2006. Phase two received it in April 2007. Phase 3 received it in April 2008.*

A.4 Measurement error and my estimates

There are at least three types of measurement error present in my measurement of cropland fires. The first is introduced by the presence of cloud cover, which makes it difficult for satellites to measure the presence of fires. The presence of clouds therefore leads me to systematically under count the true number of fires. This will attenuate my estimates towards zero. To see this consider the a district that has 5 “true" fires prior to treatment and 7 fires after treatment. In that case the “true" treatment effect is an increase of 2 fires. Now consider the case with a constant level of cloud cover that reduces my counts of fire by 50%. Now I measure 2.5 fires prior to treatment and 3.5 fires after treatment and estimate that treatment increased fires by only 1 fire. The same attenuation would exist in the case of a negative treatment effect.

The second case relates to the fact that widespread fire use can lead to the creation of clouds (Fromm *et al.*, 2010; Gatebe *et al.*, 2012; Jain *et al.*, 2014). This will exacerbate the attenuation described above because it suggests I under count fires by more when there are more fires. In the example above it suggests that clouds lead to counting 2.5 fires prior to treatment but, due to increased cloud cover driven by the increase in fires, I only count 3 fires after treatment and therefore estimate that treatment only increased fires by 0.5.

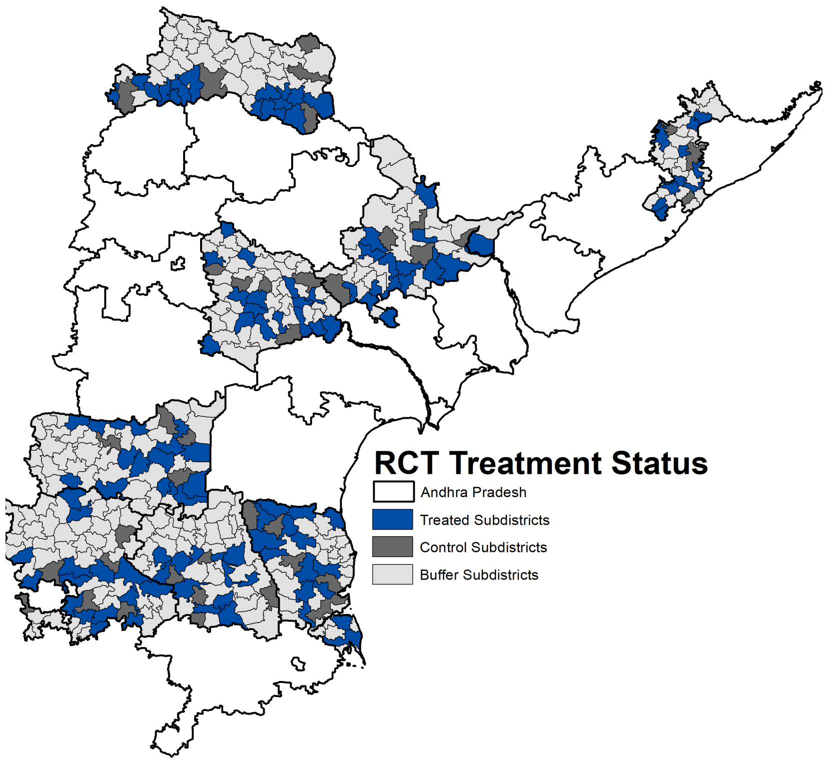
The final source of measurement error is due to the large resolution of MODIS. As I discuss below, MODIS under counts the number of fires because its resolution is 1 square kilometer and it cannot distinguish between pixels with 1 or 10 fires. The impact of this measurement error is more ambiguous than the previous two but recent work suggests it may also lead to attenuation (Abay *et al.*, 2019).

A.5 The Andhra Pradesh NREGA improvement RCT

Figure A.3 shows the spatial distribution of treated and control subdistricts from the RCT in Muralidharan *et al.* (2017) (MNS). Treated subdistricts received treatment following the completion of baseline surveys in June of 2010. Buffer subdistricts received treatment following treated subdistricts and meant that there was at least a two year gap between the treatment of the treated and control subdistricts.

In table A.1 I replicate the balance table from MNS showing balance across treated and control

FIGURE A.3: MAP OF MNS TREATED AND CONTROL SUBDISTRICTS



NOTES*: This map replicates the map of treatment shown in the appendix of Muralidharan et al. (2017). It shows the treated, control and buffer subdistricts of the RCT that improved the implementation of NREGA by providing biometric smart bank cards to participants as described in Muralidharan et al. (2017) and Muralidharan et al. (2016).*

subdistricts on socio-economic covariates. Column 3 reports the coefficient from a regression of treatment on the named covariate in a linear regression with state fixed effects.

In table A.2 I report summary data for the average number of pre-NREGA fires by subdistrict across all of India an the treatment and control subdistricts in the MNS RCT. I also report pre-NREGA levels of several covariates that are important in predicting fires. Except for monthly fires and average share of crop coverage all variables are reported at the district level for the All of India sample and at the subdistrict level for the MNS sample due to data limitations.

A.6 All of India compared to Andhra Pradesh

Figure A.4 shows the distribution of pre-NREGA fires in Andhra Pradesh and the MNS sample. I use the same scale as in figure 1.1 in the main text so the figure is exactly analogous to the figure showing the distribution of pre-NREGA fires across all of India. Subdistricts that are included in the MNS RCT are outlined, in black for treated subdisricts and blue for control subdistricts.

There are two main takeaways from figure A.4. The first is that the frequency of fires is lower in Andhra Pradesh than in the states of India with the most fires. This confirms the results in figure 1.2 that show Andhra Pradesh is near the median state in the number of pre-NREGA fires. The second is that the subdistricts included in the MNS sample are, broadly, not the same subdistricts in which fires most frequently occur in Andhra Pradesh but in the majority of the MNS sample districts fires are present. Further, the treated and control subdistricts appear roughly balanced on the frequency of fire.

A.7 Relationship between MODIS and VIIRs fire detection

VIIRs is similar to the MODIS platform in that it is a source for satellite based imagery. However, VIIRs is newer than MODIS, with the imagery available starting in 2012, and has higher resolution. The lack of data prior to 2012 means I cannot use VIIRs for the primary analysis.

VIIRs and MODIS are able to detect roughly the same size fires but VIIRs provides data at a much finer pixel resolution than MODIS. VIIRs resolution is roughly 375m compared to 1km for MODIS.

TABLE A.1: BALANCE TABLE

	Treatment (1)	Control (2)	Difference (3)	p-value (4)
(A) Numbers based on official records from GoAP in 2010				
% population working	.53	.52	.0072	.41
% male	.51	.51	.00043	.67
% literate	.45	.45	.0034	.72
% SC	.19	.18	.0019	.85
% ST	.1	.12	-.014	.48
Jobcards per capita	.54	.55	-.0074	.72
Pensions per capita	.12	.12	.0012	.76
% old age pensions	.48	.49	-.013	.09
% weaver pensions	.0088	.011	-.0018	.63
% disabled pensions	.1	.1	.0019	.59
% widow pensions	.21	.2	.013**	.04
(B) Numbers based on 2011 census rural totals				
Population	45697	45600	392	.85
% population under age 6	.11	.11	-.00074	.66
% agricultural laborers	.23	.24	-.0048	.61
% female agri. laborers	.12	.12	-.0031	.55
% marginal agri. laborers	.071	.063	.0078	.16
(C) Numbers based on 2001 census village directory				
# primary schools per village	2.9	3.2	-.3	.28
% village with medical facility	.67	.7	-.032	.41
% villages with tap water	.59	.6	-.0033	.94
% villages with banking facility	.12	.16	-.034**	.024
% villages with paved road access	.8	.82	-.01	.78
Avg. village size in acres	3394	3743	-316	.38

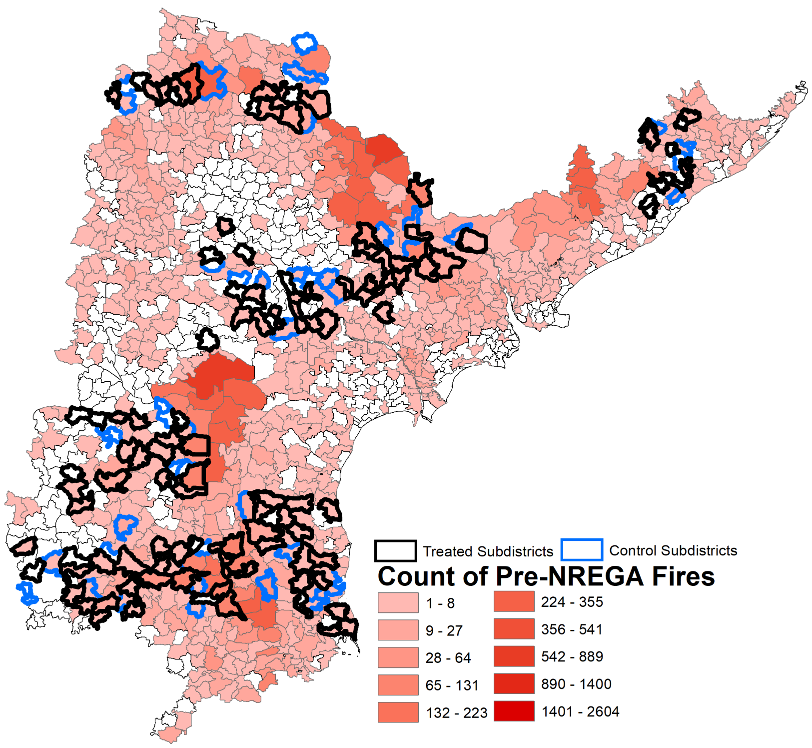
NOTES: I report here baseline characteristics across treated and control subdistricts. Column 3 reports the difference in treatment and control means. Column 4 reports the p-value on the treatment indicator from simple regressions of the outcome with district fixed effects as the only controls. The table exactly replicates that found in Muralidharan *et al.* (2016). They provide the following notes: "A "jobcard" is a household level official enrollment document for the NREGS program. "SC" ("ST") refers to Scheduled Castes (Tribes), historically discriminated-against sections of the population now accorded special status and affirmative action benefits under the Indian Constitution. "Old age", "weaver", "disabled" and "widow" are different eligibility groups within the SSP administration. "Working" is defined as the participation in any economically productive activity with or without compensation, wages or profit. "Main" workers are defined as those who engaged in any economically productive work for more than 183 days in a year. "Marginal" workers are those for whom the period they engaged in economically productive work does not exceed 182 days. The definitions are from the official census documentation. The last set of variables is taken from 2001 census village directory which records information about various facilities within a census village (the census level of observation). "# primary schools per village" and "Avg. village size in acres" are simple mandal averages - while the others are simple percentages - of the respective variable (sampling weights are not needed since all villages within a mandal are used). Note that we did not have this information available for the 2011 census and hence use the 2001 data. Statistical significance is denoted as: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$."

TABLE A.2: AVERAGE LEVEL OF COVARIATES PREDICTIVE OF FIRES NATIONALLY AND BY MNS TREATMENT STATUS

	All of India			MNS Treated			MNS Control		
	Mean	SD	Max	Mean	SD	Max	Mean	SD	Max
Avg. monthly fires	0.57	2.71	70.38	0.21	0.37	2.22	0.15	0.23	1.00
Avg. share planted in rice	0.32	0.27	0.92	0.32	0.26	1.00	0.30	0.29	0.92
Avg. share planted in wheat	0.17	0.18	0.62	0.00	0.00	0.00	0.00	0.00	0.00
Avg. share planted in sugarcane	0.03	0.07	0.57	0.01	0.03	0.21	0.01	0.02	0.13
Combines in 2005(000s)*	1.35	5.45	64.60	0.02	0.04	0.22	0.02	0.03	0.11
Share of holdings >4 HA	0.30	0.20	0.96	0.25	0.09	0.44	0.27	0.10	0.54
Avg. share of crop coverage	22.00	20.13	84.50	10.79	7.59	40.50	10.82	6.35	30.01

NOTES: Statistics for each covariate are calculated for years prior to NREGA implementation. The all of India sample includes all districts in India outside of Nicobar and Jammu & Kashmir. The MNS Treated and MNS Control refer to the subdistricts in the RCT conducted in Andhra Pradesh by Muralidharan *et al.* (2016) (MNS). Average monthly fires and average share of crop coverage are calculated at the subdistrict for both the All of India and MNS samples. All the remaining covariates are measured at the district level for the All of India sample and at the subdistrict level for the MNS sample because of data limitations. The average district in the MNS sample has 19 subdistricts. Data used to calculate the share planted in each crop in the All of India sample comes from the ICRISAT meso data (Rao *et al.*, 2012) and is the average over the years 2003-2005. Data used to calculate the share planted in the MNS sample comes from the Indian Agricultural Census in 2005 and is the level reported for 2005. Data on combines comes from the Indian Agricultural Input Survey in 2005. For the MNS sample I use the number of tractors reported to the Ministry of Statistics of Andhra Pradesh as a proxy for the number of combines. This likely overestimates the number of combines in Andhra Pradesh. Data on the share of holdings >4 HA comes from the Agricultural Census in 2005 for both samples. The share of crop coverage reports the average share of the subdistrict area that is classified as agricultural on October 31^{st} over the years 2003-2005 in the SEDAC Indian Winter Cropping dataset (Jain *et al.*, 2017; NASA, 2017b).

FIGURE A.4: PRE-MNS FIRES

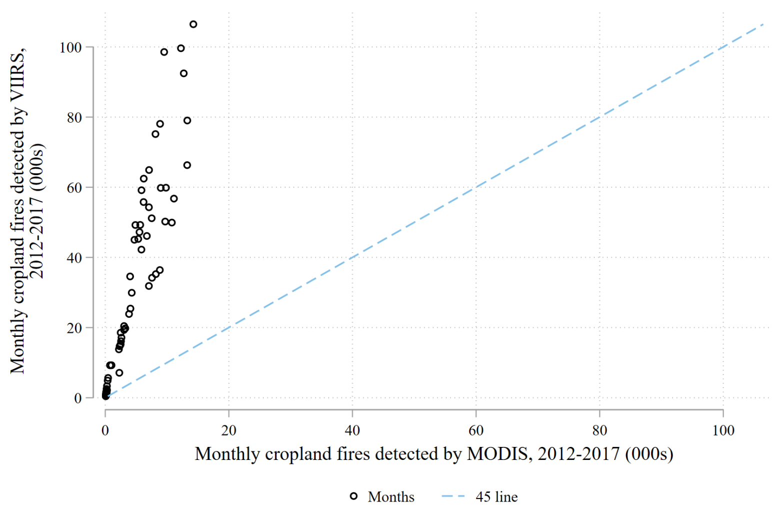


NOTES: *The count of fires by subdistrict in the years prior to implementation of the MNS RCT in Andhra Pradesh. The subdistricts that participated in the RCT are outlined. Darker areas had more fires. White areas have no fires. Data comes from the NASA FIRMS database.*

Both classify a pixel as having a fire if at least one fire is detected in that pixel. However, the finer resolution of VIIRs means it is able to count more pixels that contain fires. For example, two fires located 750m apart within a given square km would be counted as only one fire by MODIS but would likely be distinguished as two separate fires by VIIRs. Figure A.5 shows clearly that MODIS detects substantially fewer individual fires than VIIRs.

Korontzi *et al.* (2006) shows that MODIS and VIIRs are both highly accurate in counting a pixel that should contain a fire as containing a fire. So while MODIS underestimates fires it does not appear to mis-classify pixels that include low numbers of fires as non-fire.

FIGURE A.5: MONTHLY FIRES DETECTED BY MODIS AND VIIRS FROM 2012-2017



NOTES*: I count the number of fires detected by the MODIS and VIIRs platforms in each month from 2012-2017. Each black circle is a month plotted according to the fires detected by MODIS and VIIRs in that month. The dashed blue line is the 45° line. If MODIS and VIIRs detected the same number of fires each month would be on the 45° line. The observed distribution suggests that MODIS undercounts fires relative to VIIRs.*

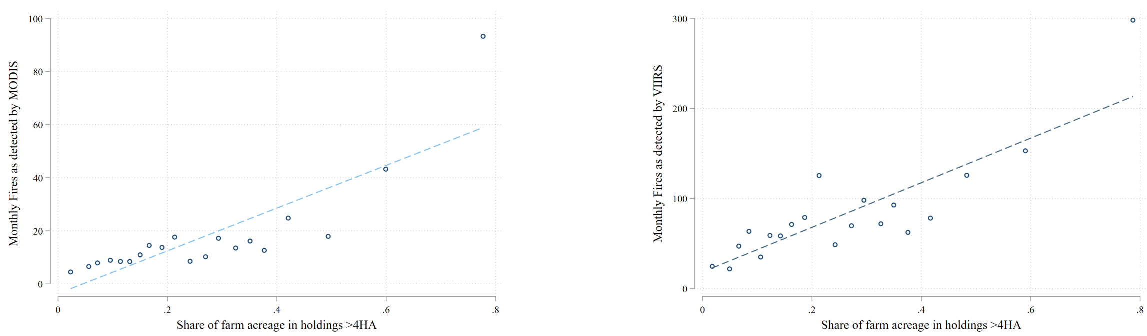
A.8 Plot size and fire relationship robustness

I show in the main text in figure 1.1 that districts in which a larger share of agricultural land is in large plots there are more frequent cropland fires. Given the large pixel size of MODIS one might think this is simply due to the increased likelihood that MODIS detects a fire on a large plot because they are likely to be larger as a result of the plot size. I can test this prediction by comparing the relationship between share of agricultural land in large plots and cropland fires detected by MODIS and cropland fires detected by VIIRs. If the relationship is driven only by the lower resolution of MODIS then the relationship between size and fires should be weaker when I use fires detected by VIIRs.

Figures A.6a and A.6b show that the relationship appears to be stronger when I use the fires detected by VIIRs. This suggests that it is not being driven by areas with larger plots having larger fires that are easier for MODIS to detect.

FIGURE A.6: COMPARISON OF PLOT SIZE AND FIRES RELATIONSHIP BY SATELLITE

(A) *MODIS* (B) *VIIRS*



NOTES*: Panel **a** shows the relationship between the number of monthly fires in a district and the share of farmland in that district in plots greater than 4 hectares when the fires were detected with the MODIS platform. Panel **b** shows the same but when the fires were detected by the VIIRS platform. In both figures the sample period is 2012-2017 and the sample covers all districts in India. The VIIRS platform can detect fires up to 10x smaller than the MODIS platform (Zhang and Wooster, 2016).*

A.9 Changes in pollutant concentrations after NREGA implementation

Because I cannot measure the dispersion of pollutants from biomass burning I do not estimate the causal effect of NREGA on pollution. Instead, I show that emissions rates are correlated with concentrations at the district month level. The correlation I measure suggests that an emissions rate ten percent higher than average has $PM_{2.5}$ concentrations that are between 0.5% and 0.8% higher than average.

TABLE A.3: CORRELATION BETWEEN EMISSIONS RATES AND CONCENTRATION OF $PM_{2.5}$

	Black Carbon	Organic Carbon	SO_2
Log Emissions Rate	0.08***	0.05***	0.08***
	(0.005)	(0.003)	(0.005)
Districts	556	556	556
N	80,064	80,064	73,392
District FE	X	X	X
Weather Controls	X	X	X
Year × Month FE	X	X	X

NOTES: The outcome is the log of the monthly concentration $PM_{2.5}$ in $\mu g/m^3$.The coefficient can be interpreted as the approximate percentage change in concentration for a percentage change in emissions rates. The specification is a linear fixed effects regression of the form $\log(PM_{imy}) = \beta log(ER_{imy}) + \gamma_i + \delta_{my}$ where PM_{imy} is the concentration of $PM_{2.5}$ in district i in month m in year y. ER_{imy} is a the rate of emissions of the named pollutant in district i. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. N refers to the number of district × months included in each regression. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel of districts in India from 2003-2014. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

A.10 Cropping patterns across India

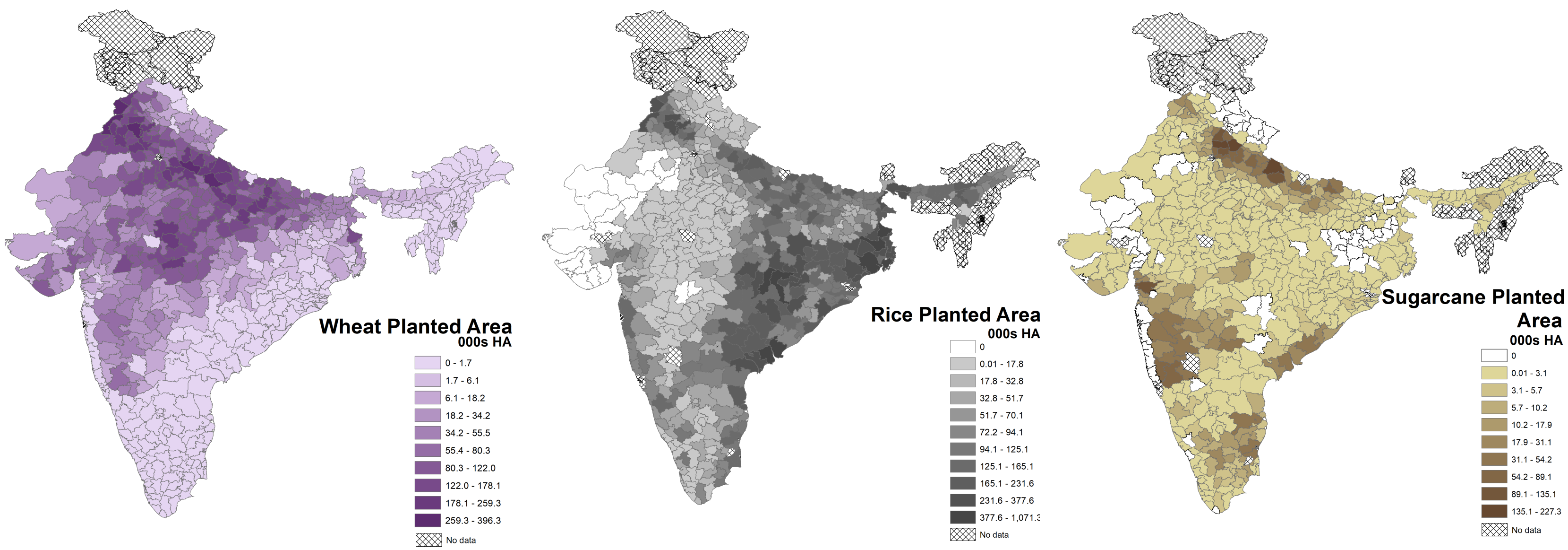
Figures A.7a-A.7c show the distribution of planted area in wheat, rice, and sugarcane across India. I show the annual average in 000s of hectares in each crop over the years 2003-2005. Data comes from the ICRISAT meso database (Rao *et al.*, 2012).

Wheat production is clearly concentrated in Punjab and along the Indo-Gangetic plain. This broadly aligns with the areas that have the highest frequency of fires as shown in figure 1.1 in the main text. The relationship between fires and coupled rice-wheat production is highlighted by comparing the map of rice production with the map of wheat production and fires. The districts with the highest frequency of fires are the districts where high production of rice and wheat appear to overlap. Notably, there are districts of high wheat production with low rice production and vice-versa, these districts do not appear to have as high a frequency of fire. Districts with a high area in sugarcane production also appear to have more frequent fires although the visual correlation is not as strong.

The importance of the coupled rice-wheat production for fire use is brought out more clearly in figure A.8. Here I show the average share of a subdistrict's area that is covered in crops on October 31^{st} where the average is calculated across the years 2003-2005. The area covered by crops is calculated by NASA from remotely sensed imagery (NASA, 2017b; Jain *et al.*, 2017) and measures not the area of cropland but the share of a pixel on which crops are actively growing on October 31^{st} each year. This is not a perfect proxy for areas that engage in coupled rice-wheat production but it captures areas that are growing rice crops during the post-monsoon season and the visual correlation between these areas and those that grow wheat in figure A.7a is high. Figure A.8 highlights that areas that appear to most intensively engage in coupled rice-wheat production are also the areas that have the highest frequency of fires. Figure A.9 shows that much of AP, and many of the sample subdistricts in MNS do not appear to engage in coupled rice-wheat production.

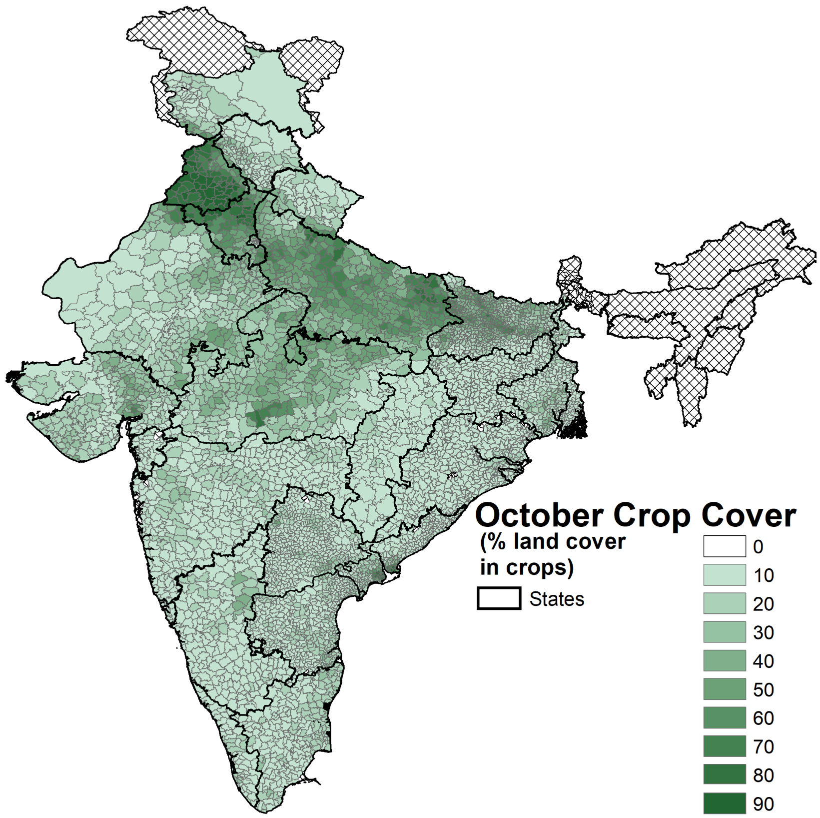
FIGURE A.7: CROPPING PATTERNS IN WHEAT, SUGARCANE, AND RICE ACROSS INDIA

(A) *Wheat* (B) *Rice* (C) *Sugarcane*



NOTES*: The average area planted by district annually in 000s of hectares in wheat, rice and sugarcane in the pre-NREGA period from 2003-2005. Data comes from the ICRISAT meso dataset (Rao et al., 2012).*

FIGURE A.8: AVERAGE CROP COVERAGE ON OCTOBER 31^{st}

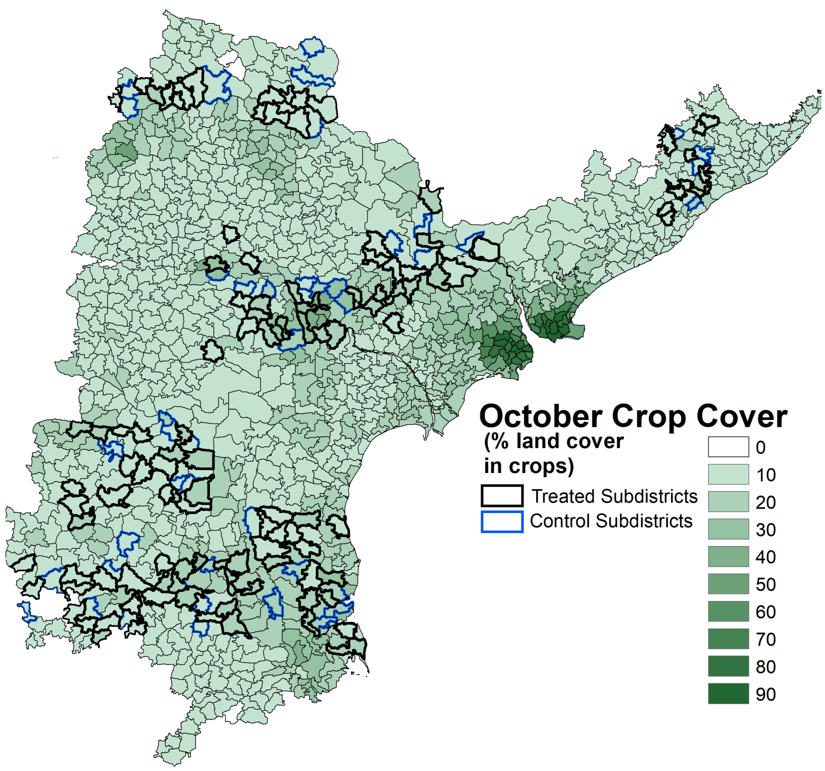


NOTES*: This figure shows the average share of pixels in a subdistrict across the areas of India for which data is available that have crop coverage on October 31^{st} over the years 2003-2005. Crop coverage is measured by reflectivity detected by satellite as described in Jain et al. (2017). Data comes from the Center for International Earth Science at Columbia University NASA (2017b).*

A.11 Event studies for non-fire outcomes

Figures A.10-A.13 show the event studies of NREGA's impact on the area planted in rice, wheat, sugarcane and all other crops in 000s of HA. Data comes from the ICRISAT meso dataset (Rao *et al.*, 2012). In all cases I confirm that the assumption of no differential pre-trends appears to hold. Further, each shows little to no evidence of an increase in area planted after the implementation of NREGA. Consistent with the results in table 1.8 rice shows a small increase after NREGA's implementation

FIGURE A.9: AVERAGE CROP COVERAGE ON OCTOBER 31st IN ANDHRA PRADESH



NOTES*: This figure shows the average share of pixels in a subdistrict across Andhra Pradesh is available that have crop coverage on October 31st over the years 2003-2005. Crop coverage is measured by reflectivity detected by satellite as described in Jain et al. (2017). Data comes from the Center for International Earth Science at Columbia University NASA (2017b).*

but this decays quickly and is never statistically different from zero. It is distinctly different from the impact of NREGA on fires that shows initial increases that persist over time.

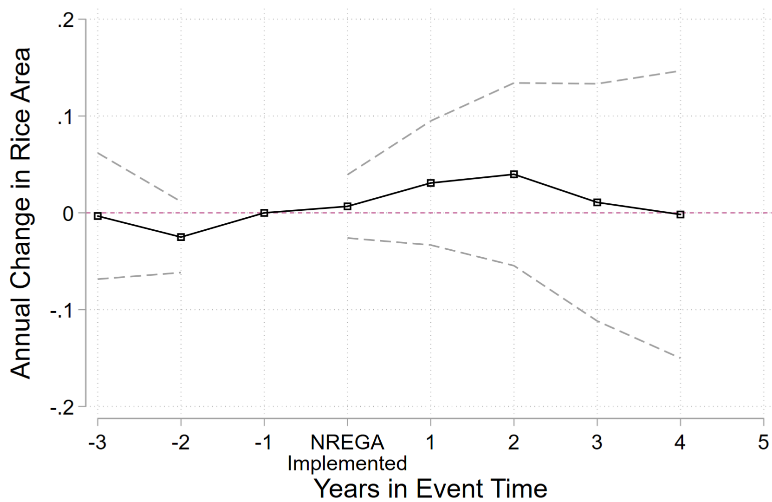
Figure A.11 suggests NREGA had no impact on area planted in wheat while figure A.12 shows weak evidence that sugarcane production declines after the implementation of NREGA. A decline in sugarcane production is consistent with both anecdotal evidence (FLA, 2012) and the model presented in the main text. That model suggests that facing higher labor costs farmers can either reduce production or mechanize. However, the technology to mechanize the harvest of sugarcane in India is not widespread (Yadav, 2007; Solomon, 2016). That suggests that mechanization is not as feasible in the short-term, which leads farmers to reduce production in the face of higher labor costs (Clemens *et al.*, 2018).

The area planted in other crops (figure A.13) follows a similar pattern to rice production, showing a small initial increase that decays. This is consistent with farmers shifting into other, higher value crops (e.g. cotton) after the implementation of NREGA (Rabotyagov *et al.*, 2014; Gehrke, 2013).

In figure A.14 I show the event study of NREGA's impact on monthly cropland fires in the first and fourth quartiles of the mechanization index. This corresponds to the regression results in panel A and panel D of table 1.7 in the main text. The event study confirms that there is no evidence of pre-trends within the two quartiles. Further, it shows the same increase in fires in the districts with the highest score in the mechanization index that are reported in the table 1.7.

In figure A.15 I show the average level of night lights appears to decline initially and then recover after the implementation of NREGA. This is consistent with the findings in Cook and Shah (2019) where they find that night lights decline after NREGA implementation in phase 1 and 2 districts but to increase in phase 3 districts. If one is concerned that may measure of fires is simply identifying areas with more night lights the decline in night lights after the implementation of NREGA suggests that measurement error of this type cannot explain my results. I also show in figure A.16 that there does not appear to be a strong relationship between districts with high night lights and high fires.

FIGURE A.10: AREA PLANTED IN RICE EVENT STUDY



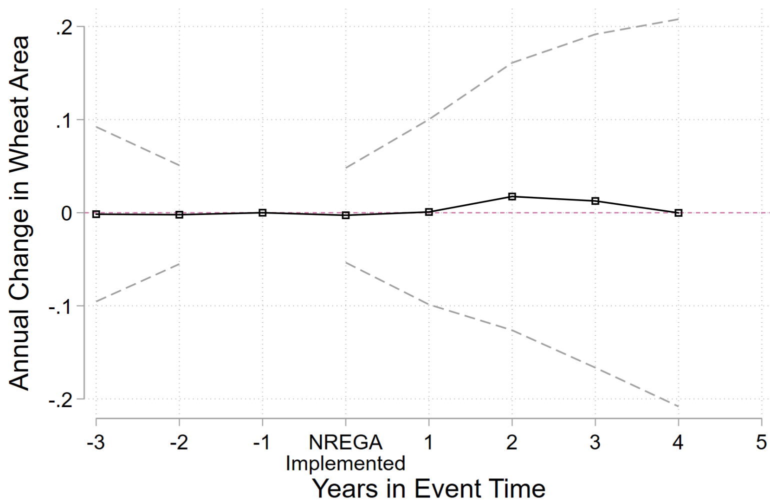
NOTES: *Each point is the estimated ω_τ coefficient from the regression* $log\left(\mathbb{E}\left[C_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}$, *where Y_τ is an indicator for event-time year τ in the set $T = \{-3,-2,-1,0,1,2,3,4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\boldsymbol{W}_{imy}$ are weather controls. C_{imy} is the area planted in rice in 000s of hectares in month m in year y in district i. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year before NREGA is implemented.*

A.12 Predictive power of mechanization index

The goal of the mechanization index is to identify areas in which mechanization may be a more viable option for farmers. While I cannot observe the direct impact of NREGA on mechanization because of data limitations I do observe combine counts in 2011, five years after the first implementation of NREGA. As a result, I can test the ability of the mechanization index to predict the number of combines five years after NREGA's implementation.

In figure A.17 I show that the mechanization index appears to predict combine levels in 2011 reasonably well. I show the binscatter of the district level count of combines in 2011 and the mechanization index score calculated based on data from 2003 to 2005. Data on combines comes

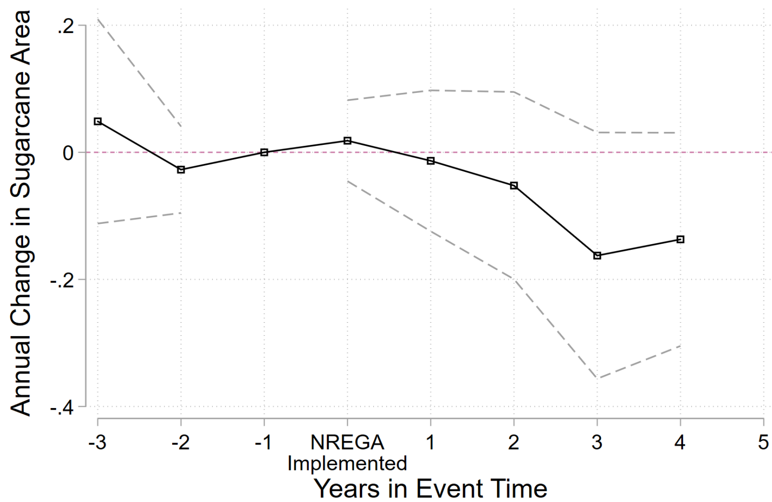
FIGURE A.11: AREA PLANTED IN WHEAT EVENT STUDY



NOTES: *Each point is the estimated ω_τ coefficient from the regression* $log\left(\mathbb{E}\left[C_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}$, *where* Y_τ *is an indicator for event-time year* τ *in the set* $T = \{-3,-2,-1,0,1,2,3,4\}$, ψ_i *is a district fixed effect ,* δ_{my} *is a month* $\times$ *year fixed effect and* $\boldsymbol{W}_{imy}$ *are weather controls.* C_{imy} *is the area planted in wheat in 000s of hectares in month* m *in year* y *in district* i*. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year before NREGA is implemented.*

from the agricultural input survey. The mechanization index is calculated as described in the main text. I plot the linear best fit line using all the data in the dashed light blue line. Because there is a clear outlier in the binscatter I also show the linear best fit line excluding that data in the darker dotted line. Note that the excluded data is not a single district but rather the binned data for the approximately 30 districts with the highest mechanization index scores and the highest level of combines in 2011.

FIGURE A.12: AREA PLANTED IN SUGARCANE EVENT STUDY

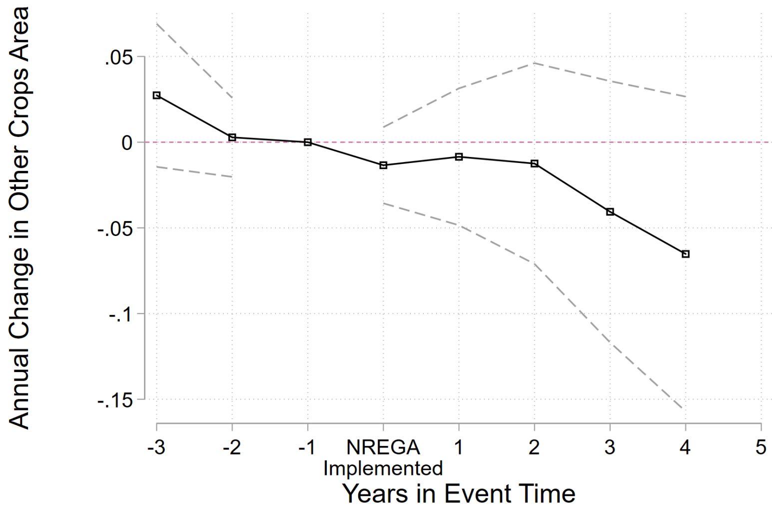


NOTES*: Each point is the estimated ω_τ coefficient from the regression $\mathbb{E}\left[C_{imy}|\boldsymbol{X}_{imy}\right] = exp\bigg(\sum_{\tau \in T} \omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}\bigg)$, where Y_τ is an indicator for event-time year τ in the set $T = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\boldsymbol{W}_{imy}$ are weather controls. C_{imy} is the area planted in sugarcane in 000s of hectares in month m in year y in district i. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year before NREGA is implemented.*

A.13 Effect of NREGA on crop production by mechanization index and pre-NREGA fires

To verify that the lack of a mean effect on cropping levels is not masking heterogeneity across districts in cropping responses that is correlated with the mechanization index, and so allowing for changes in cropping levels to drive the results shown in table 1.7 in the main text, I show that the impact of NREGA on cropping levels does not vary by the mechanization index in table A.4. Across all crops and all levels of the mechanization index the results indicate that NREGA has little impact on cropping

FIGURE A.13: AREA PLANTED IN OTHER CROPS EVENT STUDY



NOTES: *Each point is the estimated ω_τ coefficient from the regression* $log\left(\mathbb{E}\left[C_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}$, *where Y_τ is an indicator for event-time year τ in the set $T = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\boldsymbol{W}_{imy}$ are weather controls. C_{imy} is the area planted in all other crops in 000s of hectares in month m in year y in district i. 95% CIs are show in dashed grey lines. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year before is the year NREGA is implemented.*

levels, consistent with the results in table 1.8 in the main text.

There appear to be slight increases in production and area planted of rice in the second quartile of the mechanization index and slight increases in area planted in the fourth quartile. Sugarcane shows a slight increase in production in the second quartile as well. However, none of these changes can explain the pattern of results in the main text where the increase in fires appears to be strongly concentrated in the districts with the highest mechanization index score.

TABLE A.4: EFFECT OF NREGA ON CROP PRODUCTION BY MECHANIZATION INDEX

	Wheat		Rice		Sugarcane		All Crops	
	Area	Tons	Area	Tons	Area	Tons	Area	Tons
(A) Mech. Index, Q1								
Post-NREGA	0.001	0.007	0.017	0.055	-0.009	-0.246*	0.003	-0.032
	(0.012)	(0.017)	(0.019)	(0.033)	(0.033)	(0.129)	(0.011)	(0.038)
Districts								
Months	120	120	120	120	120	120	120	120
N	13,224	12,384	16,860	16,860	15,972	15,972	16,860	16,860
Mean								
(B) Mech. Index, Q2								
Post-NREGA	0.011	0.054	0.070*	0.097*	0.093	0.204**	0.005	0.122
	(0.056)	(0.077)	(0.037)	(0.053)	(0.066)	(0.096)	(0.021)	(0.081)
Districts								
Months	111	111	111	111	111	111	111	111
N	8,880	8,760	8,976	8,976	8,112	8,112	9,240	9,240
Mean								
(C) Mech. Index, Q3								
Post-NREGA	-0.031	-0.012	0.006	0.058	0.077	-0.283	0.019	0.103**
	(0.042)	(0.065)	(0.020)	(0.043)	(0.088)	(0.182)	(0.015)	(0.044)
Districts								
Months	108	108	108	108	108	108	108	108
N	12,456	12,096	13,620	13,620	11,988	11,976	13,980	13,980
Mean								
(D) Mech. Index, Q4								
Post-NREGA	-0.044	-0.058	0.023*	-0.004	0.056	0.024	0.008	-0.033
	(0.032)	(0.041)	(0.012)	(0.023)	(0.061)	(0.100)	(0.010)	(0.027)
Districts								
Months	116	116	116	116	116	116	116	116
N	15,828	15,708	14,724	14,724	15,192	15,072	16,296	16,296
Mean								
District FE	X	X	X	X	X	X	X	X
Year × Month FE	X	X	X	X	X	X	X	X
Weather Controls	X	X	X	X	X	X	X	X

NOTES: Each column represents separate regressions. Outcomes for each column are listed in the column headings. Area is measured as area planted in 000s of HA and Tons measures annual production in Tons. In all columns the base regression is a poisson fixed effects of the form $log\Big(\mathbb{E}\left[C_{imy}|\mathbf{X}_{imy}\right]\Big) = \beta\sum_{z=1}^{4}\left[Post_{imy} \times Mech_{iz}\right] + \gamma_i + \delta_{my}$ where C_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Mech$_{iz}$ is an indicator for where district i falls in the distribution of the ease of mechanization index. The ease of mechanization index is the sum of a district's Z score across measures of land concentration, combine presence and crop types. The mechanization index is calculated based on levels of each component variable in the district prior to 2006. γ_i are district fixed effects while δ_t is a year by month fixed effect. The Ease of Mechanization Index is an index that considers, in NREGA districts, the type of crops planted, the average area of holdings and the number of combines prior to treatment in a given district. Areas that have larger farms, plant more wheat and/or rice, and have more pre-treatment combines are given higher scores. In the Andhra Pradesh subdistricts the index omits combines for lack of data. N refers to the number of district × months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2014. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

A.14 Heterogeneity by mechanization index components

My index of mechanization is the sum of a district's Z score across several different measures of how easy it may be to mechanize the harvest in a given district. These are all measured over the pre-NREGA period from 2003-2005 and are:

1. The share of agricultural land in holdings larger than 4 hectares. I consider this because the efficiency of harvesting by combines increases as the area of land harvested increases. As Clemens *et al.* (2018) shows, mechanization is more efficient at lower levels of labor per unit land. Larger farms have more available land to spread the labor of operating the combine over.

2. The share of land in marginal holdings. This is not mechanically determined by the share of land in large holdings because there are several size classes in between marginal and large. Increasing the share of land in marginal holdings, holding the share in large holdings constant reduces mechanization. This happens for two reasons. The first is simply the inverse of the reason mechanization is more frequently used on larger holdings; combines are less efficient on smaller plots with higher levels of labor per unit land. There is an additional reason why mechanization occurs less used less on marginal land however. NREGA allows marginal farmers to use labor from NREGA on projects on their private land (GOI, 2009). While marginal farmers cannot use NREGA labor on the harvest, to the extent that money is fungible, using NREGA labor on projects that farmers would have otherwise paid for themselves frees money to pay for harvest labor.[4] Because more land in marginal plots should reduce mechanization I invert the Z score by multiplying by -1 when I calculate the mechanization index.

3. The number of combines in the district in 2005. Combines are the unit of capital most directly related to mechanization of the harvest and their use is a primary reason for the increase in the use of fires (Yang *et al.*, 2008; Bhuvaneshwari *et al.*, 2019; Shyamsundar *et al.*, 2019). Using combines to harvest leaves more residue on the field than harvesting by hand and it is this increased residue that interferes with the next season's planting. Farmers often do not own their

[4] Fungibility may not be an innocuous assumption in this context if farmers are engaging in any form of mental accounting.

own combines but rather rent a combine and operator's time for a specific harvest (Shyamsundar *et al.*, 2019). Having more combines at implementation of NREGA facilitates mechanization, all else equal, by reducing congestion in this rental market.

4. The area planted in rice and wheat. As discussed above, the use of fires is most intense in areas of coupled rice-wheat production. To account for this I measure the average annual area planted in rice and wheat.

5. The area planted in sugarcane. Farmers who grow primarily sugarcane do not have the ability to easily mechanize in India (Yadav, 2007; Solomon, 2016). Areas that have more land planted in sugarcane should, as a result, see less mechanization than other areas. To account for this I invert the Z score for sugarcane area by multiplying by -1 before I calculate the mechanization index.

In table A.6 I show that across each individual component of the mechanization index the use of fires increases in areas that the component would predict are easier to mechanize. For several of the components the effect seems to begin after the median, as opposed to only in the top quartile as with the overall mechanization index, but others follow the same pattern as the overall mechanization index. The only exception is the area planted in sugarcane. However, this is consistent with the distribution of districts by the area planted in sugarcane. All districts except the first quartile, the areas with the most sugarcane production, have essentially no area planted in sugarcane. As a result, the large increases in fires outside in the second through fourth quartile are consistent with farmers in sugarcane producing areas not being able to mechanize easily while those in non-sugarcane producing areas are more able to mechanize the harvest.

A.15 Heterogeneity by mechanization index in the MNS sample

Data limitations restrict my construction of the mechanization index for Andhra Pradesh so my ability to examine heterogeneity in the response of fires across the mechanization index in the RCT is constrained. Specifically, I do not have access to data that measures the number of combines in Andhra Pradesh prior to the implementation of NREGA. This is an important part of the mechanization index because the ability to rent time from a combine operator, a primary method of mechanizing in this context, is facilitated when there are more combines. Despite this, I can construct a measure of mechanization index that uses the number of tractors to proxy for the number of combines. This is likely an overestimate of the number of combines in a given district.

Examining heterogeneity across my Andhra Pradesh specific mechanization index using the RCT sample confirms the results in the main text. While the results are very noisily estimated I see the same pattern where the only quartile in which fires increase is that in which mechanization is predicted to be easiest. Despite the noise in the estimates, this is also the only quartile in which the estimate is statistically different from zero at conventional (10%) levels.

A.16 Changes in agricultural wages after NREGA

Using data from the Indian Agriculture Ministry on wages in specific agricultural occupations across some Indian states I attempt to replicate the results in Imbert and Papp (2015). I do not find a general increase in agricultural wages in the sample of states for which I have data (table A.8). When I attempt to replicate their specification focusing on star states and restricting to districts in phase 1 relative to untreated districts in phase 3 of NREGA (table A.9) I find increases in the wage of field labor – perhaps the best approximately of general unskilled agricultural labor – that is consistent with their results.

There are many potential reasons why I fail to replicate their results. The most likely is a combination of a small sample - my data on wages is far less comprehensive than theirs with data on wages in fewer than half the districts in India - and inaccuracies in the collection of wage data. While my data reports wages by occupation it is collected by the Agriculture Ministry and is likely less complete and less accurate than the NSS data that Imbert and Papp (2015) uses.

A.17 Robustness

A.17.1 RGGVY comparison

In 2005 the Indian government rolled out a national program ("Rajiv Gandhi Grameen Vidyutikaran Yonana" (RGGVY)) intended to electrify those villages that remained un-electrified or were "under-electrified" (see Burlig and Preonas (2016) for more detail). The program had a similar financing structure as NREGA – funding came from the Federal government but projects were implemented at a local level. Crucially, funding was dispersed under two different five year plans, the 10^{th} and the 11^{th} and not all districts receive funding under both.

In order to receive funding a State had to submit a district specific proposal to the Rural Electrification Corporation (REC), overseen by the Ministry of Power. Proposals were reviewed by the REC and funds were disbursed by them on approval.

Submitting a proposal was a costly act by the state, requiring surveys, and a detailed village-by-village implementation plan indicating which households and public places were to be electrified (Burlig and Preonas, 2016). Performing this costly action earlier or faster may be an indicator of a government's ability to effectively implement programs. Under that assumption, districts that received funding in the 10^{th} Five Year plan may be more effective at implementing government programs than those that did not. However, this assumption may not be valid. It may have been the case that lower capacity or less developed districts were specifically targeted by the REC for assistance in putting together their applications in order to facilitate participation in the 10^{th} five year plan. In that case, participation in the earlier round of funding may indicate lower government capacity. There are a number of reasons district government capacity might impact the fire response to NREGA. One is simply that if higher capacity districts can better implement NREGA the labor market shock may be larger (Imbert and Papp, 2015).

I divide districts into those that receive funding under RGGVY in the 10^{th} Five Year Plan and the 11^{th} Five Year Plan (districts that receive funding in both are included in the 10^{th}). I show a map of these districts in A.18. I then run the primary specification described in equation 1 of the main text on each sub-sample.

The results in table A.10 indicate no difference in the impact of NREGA on fires between the early implementing and late RGGVY implementing districts. This is consistent with what Burlig and Preonas (2016) find in measuring the direct impacts of the program.[5] I also find no evidence of a direct effect of the RGGVY program on fires. This is not evidence that government capacity did not impact the implementation of NREGA - in part due to the lack of certainty around how district participation in each phase of the RGGVY program was determined in practice - but it is reassuring that I do not find substantial differences in the impact of NREGA on fires based on the timing of an unrelated government program.

A.17.2 Placebo tests

To ensure that the results I report in the main text are not due to underlying differences in districts that are correlated with the implementation timing of NREGA by chance I run a placebo test where I maintain the order of treatment but move it forward in time by two years for all districts. As a result, phase 1 districts are treated in 2004, phase 2 in 2005 and phase 3 in 2006. As I report in table A.11 I find no treatment effect in this placebo treatment. That suggests that my results are not driven by differences in districts across phases that are correlated with the treatment timing.

A.17.3 Randomization test

As a further test of the robustness of the main results I conduct something similar to a randomization inference test on the full, national implementation of NREGA. I keep the timing of implementation the same (i.e. phase 1 districts receive NREGA in February 2006) but I randomly assign districts to phases. This imposes a null hypothesis of no effect of program implementation. I then run the primary specification, with weather controls, 1,000 times and plot the distribution of the estimated impact of NREGA implementation in figure A.19.

Comparing the distribution of estimated effects under the null of no effect to the effect I estimate in table 1.3 in the main text (shown in the dashed red line in figure A.19) shows that it is highly unlikely

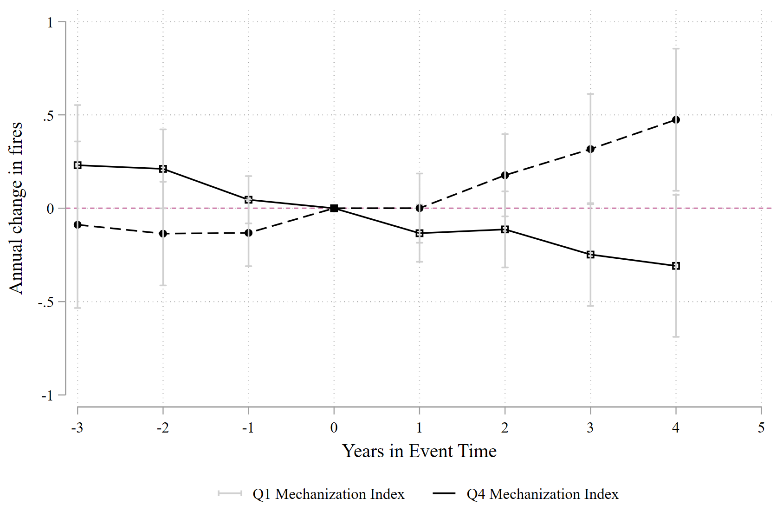
[5] They find no differential effects by early or late. However, while they find very strong increases in the rate of electrification and use of electricity they find negligible effects on several measures of income and economic activity.

the effect I estimate is due to chance assignment of districts to phases. The implied p-value on my estimated effect from this randomization exercise is < 0.001.

A.17.4 Changing timing of harvest

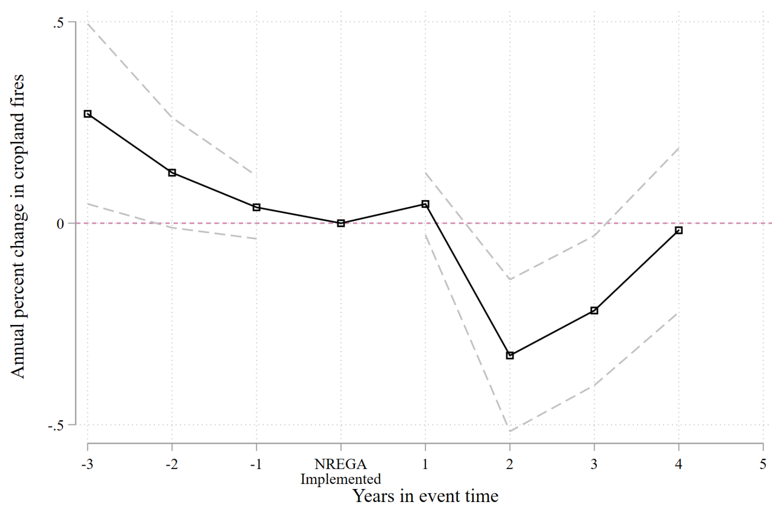
I find some evidence that the timing of the kharif harvest shifts over time from October to November (figures A.20 and A.21). Some have argued that shifting the kharif harvest to later in the year is the primary cause of the increased pollution from crop fires in Delhi in December in January. This occurs because of a shift in wind patterns in early December that cause more pollution to be blown from Punjab to Delhi. I cannot rule this out as a potential cause of the increase in pollution in Delhi. However, the major increase in fires occurs in the post-rabi harvest. This can clearly be seen in the monthly patterns. The timing of fires shifts during the kharif harvest but there is no evidence that the overall number of fires declines. Further, as table 1.3 shows, NREGA substantially increases pollution *within* districts. Thus, while the increase in Delhi's pollution is likely due to many causes there is still clear evidence that NREGA increases the frequency of cropland fires.

FIGURE A.14: MECHANIZATION INDEX EVENT STUDY



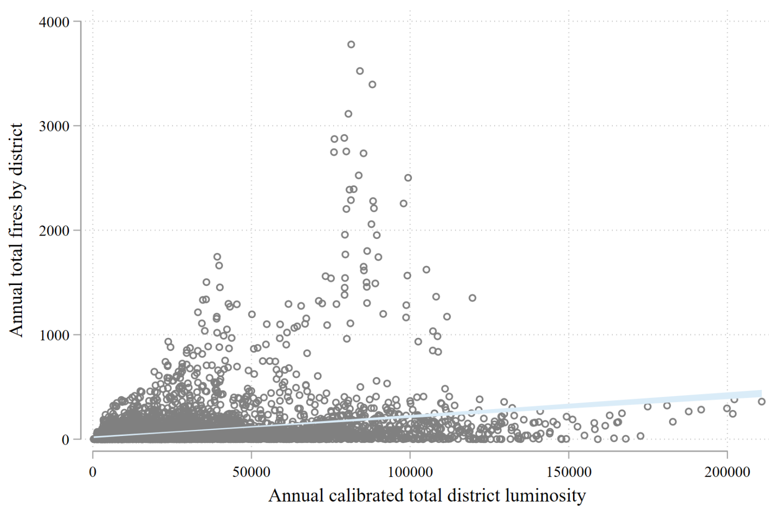
NOTES*: Each point is the estimated ω_τ coefficient from the regression $log\left(\mathbb{E}\left[F_{imy}|\boldsymbol{X}_{imy}\right]\right) = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \boldsymbol{W}_{imy} + \psi_i + \delta_{my}$, where Y_τ is an indicator for event-time year τ in the set $T = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, ψ_i is a district fixed effect , δ_{my} is a month $\times$ year fixed effect and $\boldsymbol{W}_{imy}$ are weather controls. F_{imy} is number of cropland fires in month m in year y in district i. 95% CIs are show in dashed grey lines. I run the regression separately on districts in the first quartile of the mechanization index, areas where mechanization is predicted to be more difficult, and the fourth quartile, areas where mechanization is predicted to be easier, to generate each line. The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year before NREGA is implemented.*

FIGURE A.15: NREGA'S IMPACT ON AVERAGE NIGHT LIGHTS EVENT STUDY



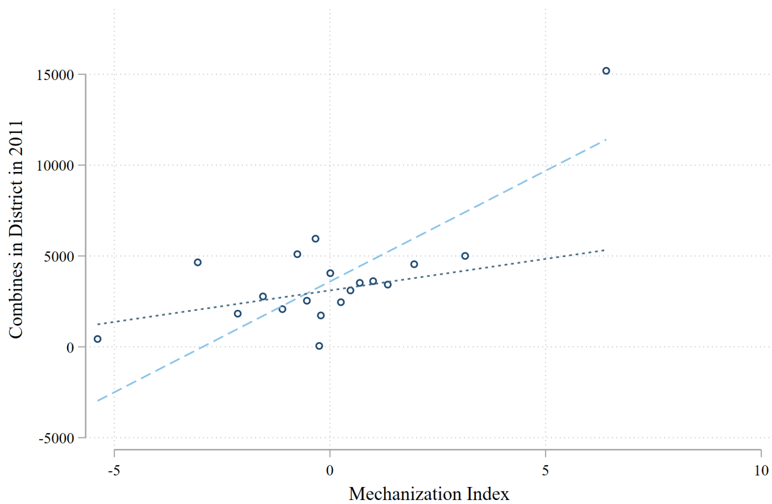
NOTES*: Each point is the estimated ω_τ coefficient from the regression $L_{imy} = \sum_{\tau \in T} \omega_\tau Y_{\tau i} + \psi_i + \delta_y$, where Y_τ is an indicator for event-time year τ in the set $T = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, ψ_i is a district fixed effect , δ_y is a year fixed effect. L_{iy} is average of the night lights in year y in district i. 95% CIs are show in dashed grey lines. Data on night lights comes from (Asher et al., 2019). The figure uses the full sample. I pool event years less than -3 and greater than 4 into those boundary values. The base year is the year before NREGA is implemented.*

FIGURE A.16: SCATTER OF DISTRICTS BY FIRES AND TOTAL NIGHT LIGHT LUMINOSITY



NOTES*: The scatter of monthly fires and the total annual luminosity of a district as reported in Almås et al. (2019).*

FIGURE A.17: ABILITY OF MECHANIZATION INDEX TO PREDICT COMBINE LEVELS IN 2011



NOTES: *The binscatter of mechanization index against the average number of combines by district in 2011. Higher values of the mechanization index indicate that mechanization was expected to be easier in that district. The mechanization index is calculated based on data from 2003-2005. The average number of combines by district in 2011 is collected by scraping the Agricultural Input Survey data for 2011. The lighter dashed line is the OLS best fit line including the districts with the highest mechanization score and the greatest number of combines. The darker dotted line is the same OLS best fit line excluding the 30 districts with the highest mechanization index and highest number of combines in 2011.*

TABLE A.5: EFFECT OF NREGA ON CROP PRODUCTION BY PRE-NREGA FIRES

	Wheat		Rice		Sugarcane		All Crops	
	Area	Tons	Area	Tons	Area	Tons	Area	Tons
(A) Quartile 1 of pre-treatment fires								
Post-NREGA	-0.013	0.086**	0.029	0.064	0.045	0.124	0.031**	0.009
	(0.028)	(0.039)	(0.021)	(0.045)	(0.082)	(0.195)	(0.013)	(0.045)
Districts	99	99	107	107	99	98	118	118
Months	144	144	144	144	144	144	144	144
N	10,608	10,608	11,640	11,640	10,932	10,812	12,864	12,864
(B) Quartile 2 of pre-treatment fires								
Post-NREGA	-0.019	-0.024	0.002	0.043	0.013	0.213	-0.011	-0.004
	(0.026)	(0.041)	(0.013)	(0.040)	(0.069)	(0.203)	(0.012)	(0.036)
Districts	112	103	123	123	111	111	126	126
Months	144	144	144	144	144	144	144	144
N	12,660	11,580	13,872	13,872	12,816	12,804	14,244	14,244
(C) Quartile 3 of pre-treatment fires								
Post-NREGA	-0.000	0.046	0.028	0.084*	0.017	-0.156	0.011	0.039
	(0.050)	(0.069)	(0.025)	(0.043)	(0.061)	(0.199)	(0.016)	(0.050)
Districts	108	106	111	111	104	104	114	114
Months	144	144	144	144	144	144	144	144
N	12,468	12,228	12,804	12,804	12,180	12,180	13,164	13,164
(D) Quartile 4 of pre-treatment fires								
Post-NREGA	0.022	-0.004	0.015	-0.002	0.042	0.031	0.004	0.060*
	(0.041)	(0.041)	(0.018)	(0.028)	(0.079)	(0.086)	(0.010)	(0.033)
Districts	122	121	132	132	128	128	134	134
Months	144	144	144	144	144	144	144	144
N	14,652	14,532	15,864	15,864	15,336	15,336	16,104	16,104
District FE	X	X	X	X	X	X	X	X
Year × Month FE	X	X	X	X	X	X	X	X
Weather Controls	X	X	X	X	X	X	X	X

NOTES: The outcome is the total area planted and tons produced of each crop. The coefficient can be interpreted as the approximate percentage change after NREGA was statutorily implemented in a district. The sample is all districts in India that were part of the NREGA program. The specification is a fixed effects Poisson of the form $log\Big(\mathbb{E}\left[C_{imy}|\mathbf{X}_{imy}\right]\Big) = \beta\sum_{z=1}^{4}\left[Post_{imy} \times Pre-Fires_{if}\right] + \gamma_i + \delta_{my}$ where C_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Pre-Fires$_{if}$ is an indicator for where district falls in the distribution of total pre-NREGA fires. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. Each panel is a different quartile of pre-NREGA fires with Q4 corresponding to the largest number of pre-NREGA fires. N refers to the number of district × months included in each regression. Districts reports districts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel of districts in India from 2003 to 2014. All columns include controls for weather in the month the outcome number of fires is measured. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

Table A.6: Heterogeneity of treatment impact by components of mechanization index

	Share of land >4HA	Inverted share of marginal land	Combines	Area in Wheat & Rice	Inverse area sugarcane
(A)Quartile 1					
Post-NREGA	0.166	-0.074	0.111	0.126	0.008
	(0.150)	(0.178)	(0.091)	(0.084)	(0.070)
Districts	100	100	243	122	123
Months	144	144	144	144	144
N	14,400	14,400	34,992	17,568	17,712
Avg. monthly fires pre-NREGA	1.73	1.81	2.46	1.93	6.42
(B)Quartile 2					
Post-NREGA	-0.044	-0.014	0.111	0.119	0.389***
	(0.111)	(0.107)	(0.091)	(0.080)	(0.120)
Districts	100	100	243	123	123
Months	144	144	144	144	144
N	14,400	14,400	34,992	17,712	17,712
Avg. monthly fires pre-NREGA	4.21	3.58	2.46	3.57	10.21
(C)Quartile 3					
Post-NREGA	0.212	0.351***	0.028	0.209**	0.333***
	(0.142)	(0.132)	(0.120)	(0.106)	(0.103)
Districts	100	100	103	123	123
Months	144	144	144	144	144
N	14,400	14,400	14,832	17,712	17,712
Avg. monthly fires pre-NREGA	3.57	3.98	4.15	3.86	4.74
(D)Quartile 4					
Post-NREGA	0.289***	0.298***	0.264***	0.176**	0.266*
	(0.101)	(0.102)	(0.090)	(0.078)	(0.154)
Districts	100	100	116	122	121
Months	144	144	144	144	144
N	14,400	14,400	16,704	17,568	17,424
Avg. monthly fires pre-NREGA	17.08	17.22	14.73	14.68	2.57
District FE	X	X	X	X	X
Month × Year FE	X	X	X	X	X
Weather Controls	X	X	X	X	X

Notes: The outcome is monthly cropland fires. The coefficient can be interpreted as the approximate percentage change in fires after NREGA was statutorily implemented in a district. The sample is all districts in India that were part of the NREGA program. The specification is a fixed effects Poisson of the form $\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] = exp(\beta \sum_{z=1}^{4}\left[Post_{imy} \times Mech_{iz}\right] + \gamma_i + \delta_{my})$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Mech$_{iz}$ is an indicator for where district i falls in the distribution of the ease of mechanization index. The ease of mechanization index is the sum of a district's Z score across measures of land concentration, combine presence and crop types. The mechanization index is calculated based on levels of each component variable in the district prior to 2006. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. Each panel is a different quartile of the mechanization index with quartile 4 corresponding to the places mechanization is predicted to be easiest. N refers to the number of district × months included in each regression. Districts reports districts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel of districts in India from 2003 to 2014. All columns include controls for weather in the month the outcome number of fires is measured. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

TABLE A.7: HETEROGENEITY OF TREATMENT IMPACT BY EASE OF MECHANIZATION INDEX

	Andhra Pradesh
(A)Quartile 1 of Ease of Mechanization Index	
Post-NREGA	-0.305
	(0.231)
Subdistricts	T:17 C:8
Months	79
N	2,844
Avg. monthly fires pre-NREGA	.35
(B)Quartile 2 of Ease of Mechanization Index	
Post-NREGA	-0.050
	(0.217)
Subdistricts	T:20 C:9
Months	88
N	3,432
Avg. monthly fires pre-NREGA	.17
(C)Quartile 3 of Ease of Mechanization Index	
Post-NREGA	-0.637
	(0.476)
Subdistricts	T:22 C:8
Months	86
N	3,010
Avg. monthly fires pre-NREGA	.16
(D)Quartile 4 of Ease of Mechanization Index	
Post-NREGA	0.452*
	(0.264)
Subdistricts	T:10 C:8
Months	64
N	2,240
Avg. monthly fires pre-NREGA	.11
District FE	X
Month × Year FE	X
Weather Controls	X

NOTES: The outcome is monthly cropland fires. The coefficient can be interpreted as the approximate percentage change in fires after NREGA was statutorily implemented in a district. The sample is the subdistricts in Andhra Pradesh included in the MNS RCT. The specification is a fixed effects Poisson of the form $\mathbb{E}\left[F_{imy}|\mathbf{X}_{imy}\right] = exp(\beta \sum_{z=1}^{4} \left[Post_{imy} \times Treated_i \times Mech_{iz}\right] + \gamma_i + \delta_{my})$ where F_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after MNS treatment and Treated$_i$ is a dummy indicating that subdistrict i was among the treated subdistricts. Mech$_{iz}$ is an indicator for where subdistrict i falls in the distribution of the ease of mechanization index within Andhra Pradesh. The ease of mechanization index is the sum of a district's Z score across measures of land concentration, tractor presence and crop types. Note this is not directly comparable to the Z score in the main text because it lacks a measure of combine presence. The mechanization index is calculated based on levels of each component variable in the subdistrict prior to 2006. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. Each panel is a different quartile of the mechanization index with quartile 4 corresponding to the places mechanization is predicted to be easiest. N refers to the number of subdistrict × months included in each regression. Subdistricts reports the treated and control subdistricts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. T The samples are a balanced, monthly panel of subdistricts in Andhra Pradesh from 2003 to 2012. All columns include controls for weather in the month the outcome number of fires is measured. Heteroskedasticity robust standard errors clustered at the subdistrict level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

TABLE A.8: EFFECT OF NREGA ON AGRICULTURAL WAGES

	Harvesters	Field Labor	Sowers	Weeders	Ploughman	Carpenters	All Ag
Dry-season	-0.013	0.018	0.001	0.001	0.001	-0.006	0.004
	(0.020)	(0.022)	(0.020)	(0.020)	(0.027)	(0.027)	(0.017)
Wet-season	-0.024	0.024	0.010	0.006	0.007	-0.011	0.010
	(0.024)	(0.024)	(0.021)	(0.023)	(0.032)	(0.028)	(0.019)
Districts	240	250	231	213	231	258	283
N	5,692	6,367	5,773	5,257	5,706	6,454	7,327
District FE	X	X	X	X	X	X	X
Year × Month FE	X	X	X	X	X	X	X

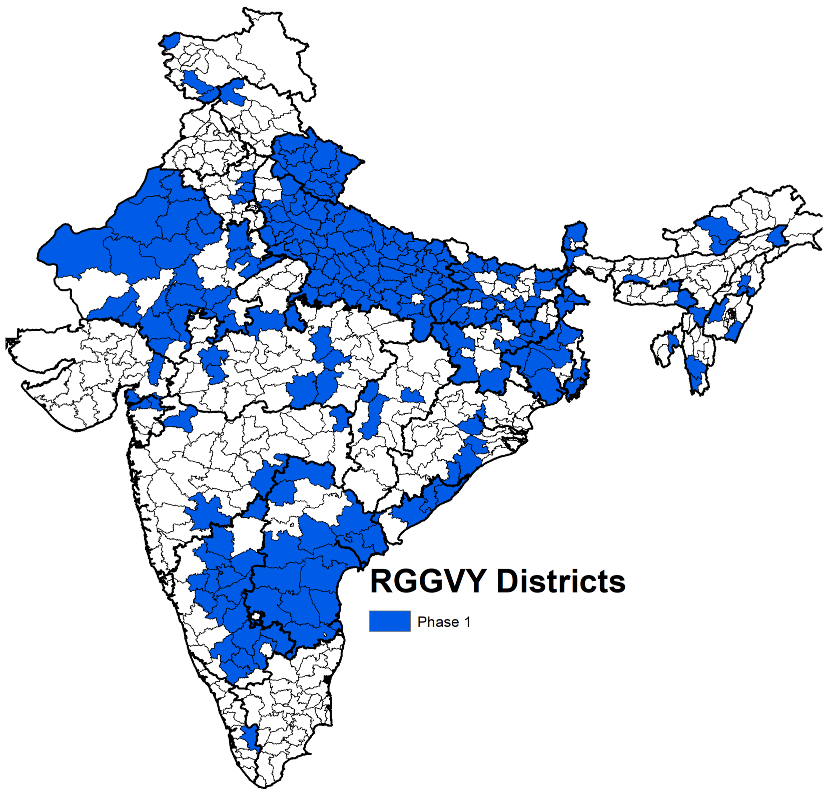
NOTES: The outcome is the log of the monthly wage in the occcupation identified in the column title. The coefficient can be interpreted as the approximate percentage change in wages in each season after NREGA was statutorily implemented in a district. The sample is phase 1 and phase 3 districts in the months prior to April 2008, replicating the specification in citeimbert2015. The specification is a linear fixed effects regression of the form $W_{imy} = Post_{imy} \times Season_{my} + \gamma_i + \delta_{my}$ where W_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Season$_{my}$ is an indicator for whether the month occurs in the dry (January to June) or wet (July to December) season. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. Each panel is a different quartile of the mechanization index with quartile 4 corresponding to the places mechanization is predicted to be easiest. N refers to the number of district × months included in each regression. Districts reports districts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

TABLE A.9: EFFECT OF NREGA ON AGRICULTURAL WAGES (IP STAR REPLICATION)

	Harvesters	Field Labor	Sowers	Weeders	Ploughman	Carpenters	All Ag
Star $\times$ Post $\times$ Dry-season	-0.047	0.033	-0.010	-0.037	-0.059	0.009	0.009
	(0.031)	(0.022)	(0.029)	(0.027)	(0.060)	(0.032)	(0.032)
Star $\times$ Post $\times$ Wet-season	-0.086**	0.048**	-0.021	-0.035	-0.058	0.010	0.010
	(0.037)	(0.023)	(0.029)	(0.032)	(0.071)	(0.036)	(0.036)
Non-star $\times$ Post $\times$ Dry-season	-0.002	0.006	0.005	0.016	0.023	-0.015	-0.015
	(0.022)	(0.031)	(0.023)	(0.023)	(0.027)	(0.027)	(0.027)
Non-star $\times$ Post $\times$ Wet-season	-0.002	0.007	0.022	0.022	0.030	-0.025	-0.025
	(0.026)	(0.033)	(0.024)	(0.025)	(0.031)	(0.029)	(0.029)
Districts	240	250	231	213	231	258	258
N	5,692	6,367	5,773	5,257	5,706	6,454	6,454
District FE	X	X	X	X	X	X	X
Year $\times$ Month FE	X	X	X	X	X	X	X

NOTES: The outcome is the log of the monthly wage in the occcupation identified in the column title. The coefficient can be interpreted as the approximate percentage change in wages after NREGA was statutorily implemented in a district. The sample is phase 1 and phase 3 districts in the months prior to April 2008, replicating the specification in citeimbert2015. The specification is a linear fixed effects regression of the form $W_{imy} = Post_{imy} \times Season_{my} \times Star_i + \gamma_i + \delta_{my}$ where W_{imy} is the outcome in district i in month m in year y. Post$_{imy}$ is a dummy variable equal to one after NREGA treatment takes effect in district i. Season$_{my}$ is an indicator for whether the month occurs in the dry (January to June) or wet (July to December) season. Star$_i$ is an indicate for whether Imbert and Papp (2015) classify the state of district i as a star state. γ_i are district fixed effects while δ_{my} is a year by month fixed effect. γ_i are district fixed effects while δ_t is a year by month fixed effect. Each panel is a different quartile of the mechanization index with quartile 4 corresponding to the places mechanization is predicted to be easiest. N refers to the number of district $\times$ months included in each regression. Districts reports districts in each sample. The average number of monthly fires (the outcome) in the pre-treatment period in each quartile are reported. The sample is a balanced, monthly panel. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

FIGURE A.18: RGGVY PHASE 1 DISTRICTS



NOTES*: Districts that were included in the first phase of the RGGVY electrification program are shown here. The first phase districts are those that received funding from the program in the 10th Five Year plan. Districts that receive funding in both phases are included in the first phase in this figure. The figure is based on data from Burlig and Preonas (2016).*

TABLE A.10: EFFECT OF NREGA ON BY RGGVY PHASE

	Cropland Fires
Post-NREGA, RGGVY Phase 1	0.214***
	(0.070)
Post-NREGA, RGGVY Non-phase 1	0.213***
	(0.053)
Districts	558
Months	144
N	80,352
District FE	X
Year × Month FE	X
Weather Controls	X

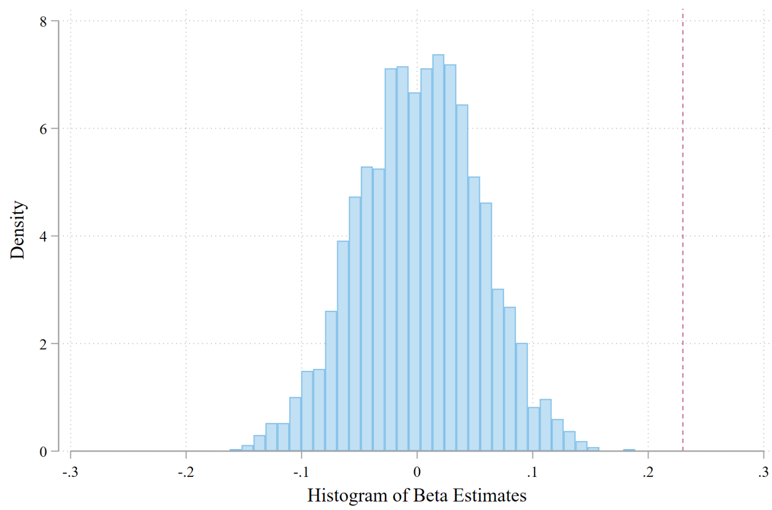
Notes: Each column represents separate regressions. In all columns the outcome is monthly cropland fires. In all columns the base regression is a fixed effects Poisson of the form $y_{it} = \beta Post + \psi[Post \times NREGA \times RGGVY_i] + \gamma_i + \delta_t$ where y_{it} is the outcome in district i in month t. Post is a dummy variable equal to one after NREGA treatment takes effect in a given phase and NREGA is a dummy indicating the NREGA phase of district *i*. RGGVY is an indicator for whether district *i* was in the first phase of the RGGVY program. γ_i are district fixed effects while δ_t is a year by month fixed effect. In all cases N refers to the number of district × months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2012. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

TABLE A.11: EFFECT OF NREGA ON FIRES (PLACEBO 1)

	Cropland Fires	
post	0.032	0.018
	(0.048)	(0.052)
Districts	558	558
Months	144	144
N	80,352	80,352
District FE	X	X
Year × Month FE	X	X
Weather Controls		X

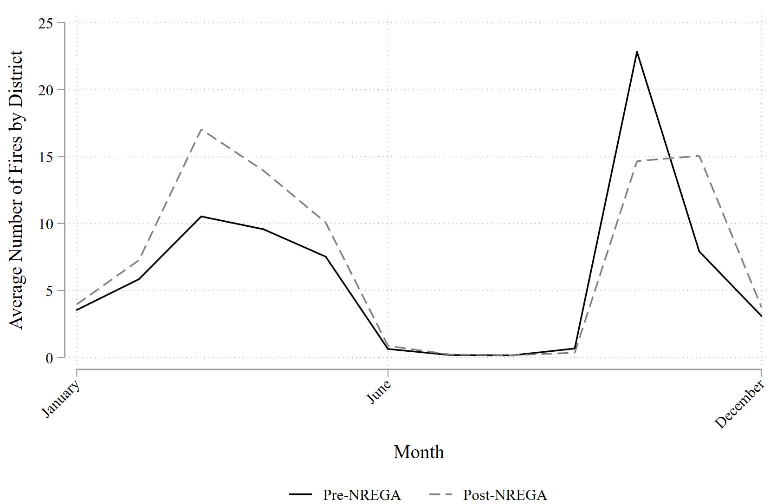
Notes: Each column represents seperate regressions. In all columns the outcome is monthly cropland fires. In all columns the base regression is a fixed effects poisson of the form $y_{it} = \beta Post + \psi[Post \times NREGA] + \gamma_i + \delta_t$ where y_{it} is the outcome in district i in month t. Post is a dummy variable equal to one after NREGA treatment takes effect in a given phase and NREGA is a dummy indicating the NREGA phase of district *i*. γ_i are district fixed effects while δ_t is a year by month fixed effect. In columns 2 I include controls for the monthly average cloud cover, precipitation and minimum and maximum temperature in district *i* in month *t*. N refers to the number of district × months included in each regression. The sample is a balanced, monthly panel of districts in India from 2003 to 2012. Heteroskedasticity robust standard errors clustered at the district level are in parentheses. (* p<.10 ** p<.05 *** p<.01).

FIGURE A.19: DISTRIBUTION OF ESTIMATED IMPACT COEFFICIENT WITH RANDOM NREGA ASSIGNMENT



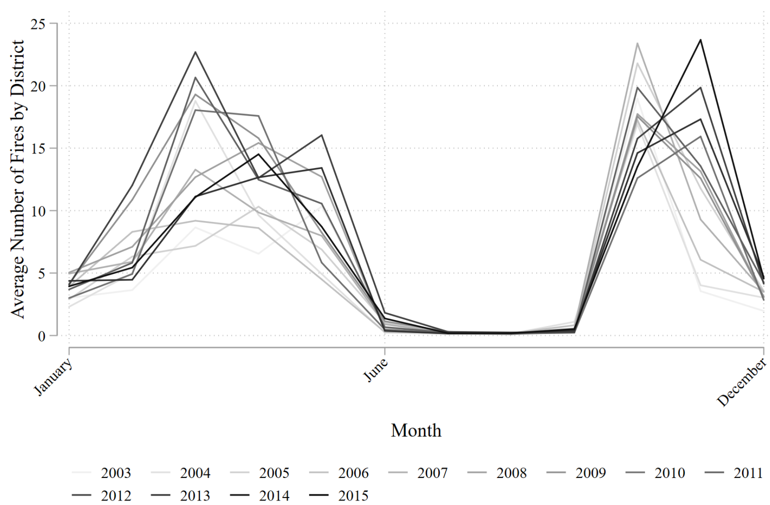
NOTES*: The distribution of the β coefficient from the primary specification in the paper under the random assignment of NREGA phase to districts. The dashed vertical line is the actual estimated coefficient in the paper. I randomly assign districts to be in phase 1, 2, or 3 of NREGA and re-estimate the primary specification, with weather controls, 1,000 times to get the distribution under random assignment. This is in the spirit of a randomization inference exercise. I can reject the null of no effect with a* $p < 0.001$.

FIGURE A.20: MONTHLY PATTERN OF FIRE USE PRE- AND POST-NREGA



NOTES*: The monthly pattern of fires averaged over all districts pre- and post-NREGA. Implementation of NREGA is measured at the district level. The count of fires by month by district is averaged over the pre- and post-NREGA period for that district. Data comes from the NASA FIRMS database.*

FIGURE A.21: MONTHLY PATTERN OF FIRE USE OVER TIME



NOTES*: The monthly pattern of fires averaged over all districts by each year in the sample. The count of fires by month by district is averaged over within each year. Lighter lines indicate earlier years. Data comes from the NASA FIRMS database.*

Appendix B

Appendix to Chapter 2

TABLE B.1: OTHER POLLUTION EXPOSURE AND COVID19 MORTALITY

	NH_4	SO_3	NO_3
Avg. NH_4, 2008-2017	0.017		
	(0.015)		
Avg. SO_3, 2008-2017		-0.018	
		(0.014)	
Avg. NO_3, 2008-2017			0.013***
			(0.005)
N	3,096	3,096	3,096
State FE	X	X	X
LASSO Controls	X	X	X

NOTES: The outcome in each column is the change in the level of the pollutant specified in the column heading.LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* p<.10 ** p<.05 *** p<.01).

TABLE B.2: INSTRUMENTAL VARIABLE RESULTS - LIML

	LN1+	LN1+	IHS	IHS
Avg. $PM_{2.5}$, 2008-2017	0.121**	0.111*	0.148**	0.138**
	(0.056)	(0.058)	(0.067)	(0.070)
N	3,097	3,097	3,097	3,097
State FE	X	X	X	X
LASSO Controls		X	X	X
NAAQS Controls		X		X

NOTES: The outcome in columns 1-2 is the log of deaths+1 in a county as of April 19^{th}. In column 3-5 it is the inverse hyperbolic sine. In all columns we estimate a LIML model rather than 2SLS. Coefficients should be interpreted as the percent change in deaths for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

TABLE B.3: INSTRUMENTAL VARIABLE RESULTS - GMM

	LN1+	LN1+	IHS	IHS
Avg. $PM_{2.5}$, 2008-2017	0.118***	0.111**	0.147***	0.137**
	(0.046)	(0.047)	(0.056)	(0.058)
N	3,097	3,097	3,097	3,097
State FE	X	X	X	X
LASSO Controls	X	X	X	X
NAAQS Controls		X		X

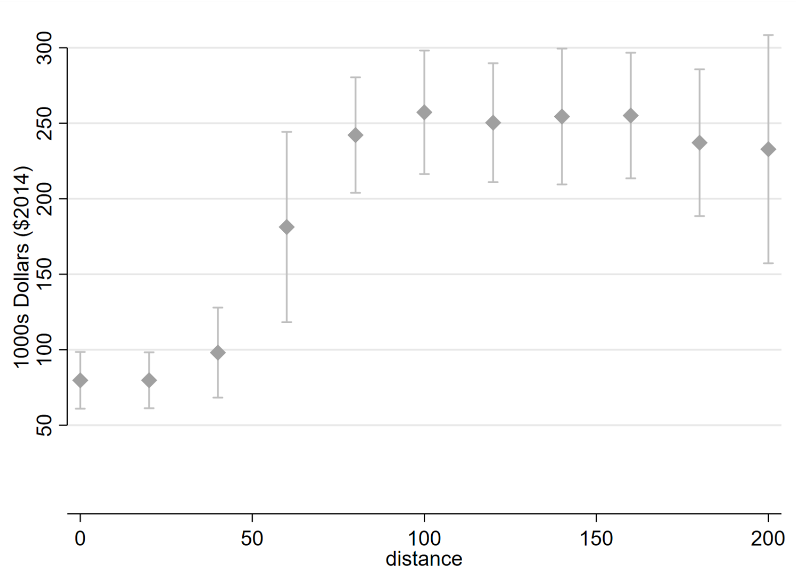
NOTES: The outcome in columns 1-2 is the log of deaths+1 in a county as of April 19^{th}. In column 3-5 it is the inverse hyperbolic sine. In all columns we estimate a two stage GMM model rather than 2SLS. Coefficients should be interpreted as the percent change in deaths for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

Table B.4: Instrumental variable results without NYC or KC

	LN1+	LN1+	IHS	IHS
Avg. $PM_{2.5}$, 2008-2017	0.103^{**}	0.092^{*}	0.128^{**}	0.116^{*}
	(0.048)	(0.049)	(0.058)	(0.060)
N	3,095	3,095	3,095	3,095
State FE	X	X	X	X
LASSO Controls	X	X	X	X
NAAQS Controls		X		X

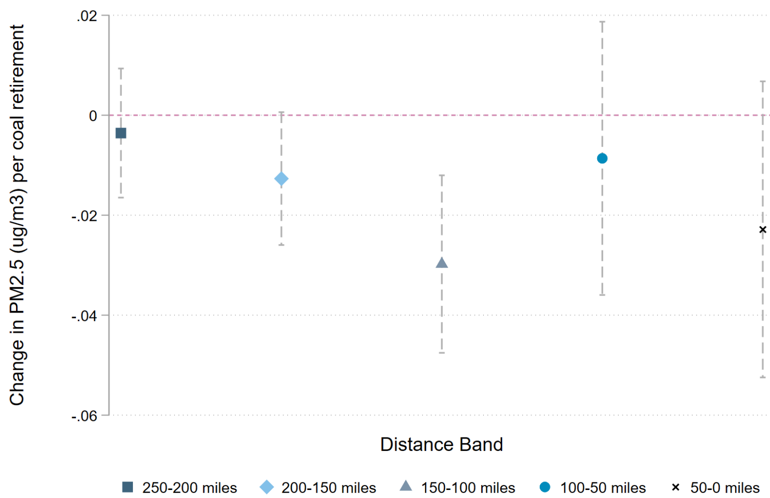
Notes: The outcome in columns 1-2 is the log of deaths+1 in a county as of April 19^{th}. In column 3-5 it is the inverse hyperbolic sine. In all columns we exclude NYC and Kansas City due to the way the NYT aggregates data in these cities. Coefficients should be interpreted as the percent change in deaths for a one unit change in annual average $PM_{2.5}$ from 2008 to 2017. At the mean a one unit change in $PM_{2.5}$ represents a 12% change in $PM_{2.5}$. LASSO controls include controls for the number of days since the first reported case, mortality rates from diabetes and obesity, population density, levels of health insurance, the racial and age makeup of a county, and days since the county experience a lockdown as well as post-lockdown activity levels. Heteroskedasticity robust standard errors are in parentheses (* $p<.10$ ** $p<.05$ *** $p<.01$).

FIGURE B.1: RANGE OF HYDRAULIC FRACTURING'S ECONOMIC SHOCK



NOTES*: We replicate the wage results in Figure 3 of Feyrer et al. (2017). They examine how the economic shock of hydraulic fracturing propagates through space. Diamonds indicate the coefficients from a regression of the one year change in wages on the total value of new production aggregated within circles with radii indicated by the mileage counts on the x-axis. 95% CIs are show in light grey. The key result is that the economic impact of new wells appears to decline to zero at 100 miles from the well location.*

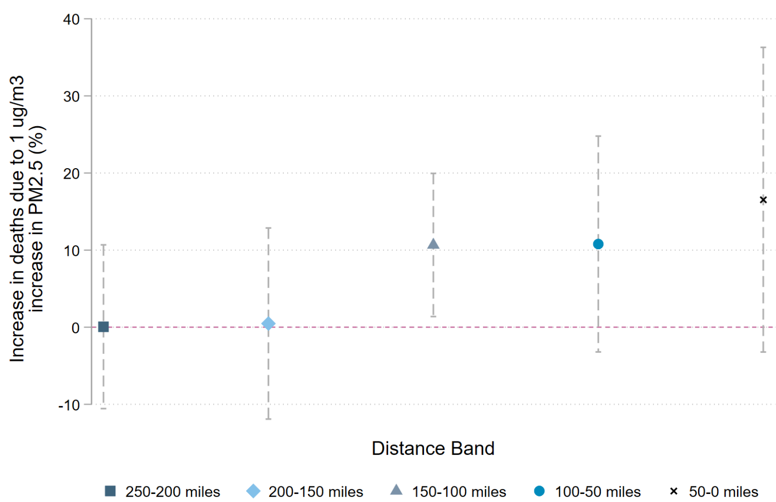
FIGURE B.2: FIRST STAGE BY BANDWIDTH



NOTES*: Each point represents the average impact of closing a coal fired power plant in any of the four quadrants estimated from our first stage regression using a range of different distance bandwidths. 95% CIs are show in light grey. Consistent with the pattern of emissions transport in figure 2.3 closure of plants in each distance band less than 150 miles has a meaningful negative impact on air quality in a given county. The large confidence bounds on the estimates in the 0 to 50 mile and 50 to 100 mile bands are likely due to relatively low number of closing plants in each band (see figure 2.4)*

.

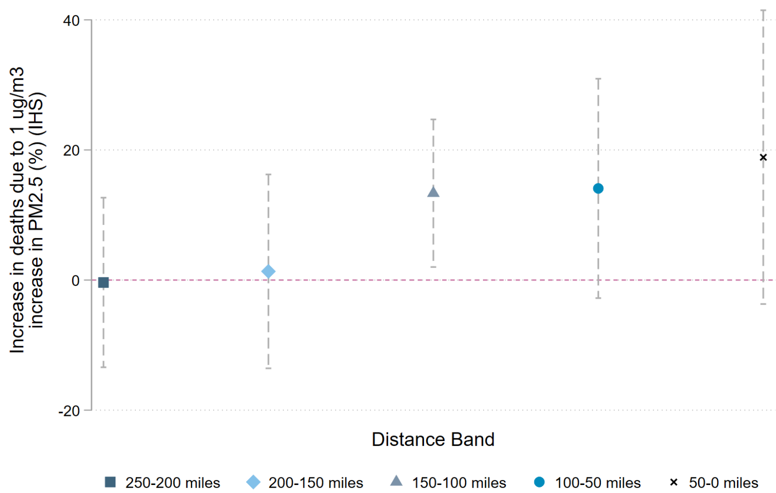
FIGURE B.3: SECOND STAGE BY BANDWIDTH



NOTES*: Each point represents the percentage change in deaths in a county from a change of a 1 $\mu g/m^3$ in the ten year annual average $PM_{2.5}$ from our main specification using various distance bands*

.

FIGURE B.4: SECOND STAGE BY BANDWIDTH - IHS



NOTES*: Each point represents the percentage change in deaths in a county from a change of a 1 $\mu g/m^3$ in the ten year annual average $PM_{2.5}$ from our main specification using the inverse hyperbolic sine using various distance bands*

.

Appendix C

Appendix to Chapter 3

TABLE C.1: SUMMER AND WEEKEND TEMPERATURES - PISA

	PISA Scores	PISA Scores	PISA Scores	PISA Scores
Total hot days	-0.126			
	(0.045)			
	[0.007]			
	{-0.22,-0.04}			
Hot school days		-0.198	-0.290	-0.309
		(0.084)	(0.072)	(0.074)
		[0.022]	[0.000]	[0.000]
		{-0.37,-0.03}	{-0.43,-0.15}	{-0.46,-0.16}
Hot summer days		0.051		0.083
		(0.135)		(0.133)
		[0.708]		[0.534]
		{-0.22,0.32}		{-0.18,0.35}
Hot school weekend days			0.341	0.353
			(0.163)	(0.164)
			[0.040]	[0.035]
			{0.02,0.67}	{0.03,0.68}
N	282	282	282	282

Notes: Heteroskedasticity robust standard errors clustered by country are in parentheses. *p*-values reported in brackets and 95% confidence intervals are reported in curly brackets. All regressions include additional controls so that heat effects are measured relative to days between 60 and 70°F. All columns include country and year fixed effects, a continent-specific linear trend, and controls for temperature in the year of the exam and precipitation in both the year of the exam and the three years proceeding the exam. All columns also control for measures of development from the World Bank.

Table C.2: Impact of crop seasons - PISA

	PISA Scores	PISA Scores
(A) All Days		
Total hot days	-0.206	-0.278
	(0.050)	(0.154)
	[0.000]	[0.083]
	{-0.31,-0.10}	{-0.60,0.04}
Hot rice season days		0.102
		(0.260)
		[0.699]
		{-0.43,0.64}
N	120	120
(B) School Days		
Hot school days	-0.410	-0.312
	(0.088)	(0.116)
	[0.000]	[0.013]
	{-0.59,-0.23}	{-0.55,-0.07}
Hot non-school days	0.039	0.305
	(0.089)	(0.259)
	[0.665]	[0.251]
	{-0.14,0.22}	{-0.23,0.84}
Hot rice season days		-0.297
		(0.294)
		[0.322]
		{-0.90,0.31}
N	120	120

Notes: Heteroskedasticity robust standard errors clustered by country are in parentheses. *p*-values reported in brackets and 95% confidence intervals are reported in curly brackets. All regressions include additional controls so that heat effects are measured relative to days between 60 and 70°F. All columns include country and year fixed effects, a continent-specific linear trend, and controls for temperature in the year of the exam and precipitation in both the year of the exam and the three years proceeding the exam. All columns also control for measures of development from the World Bank. We define rice growing seasons based on the data in Sacks *et al.* (2010). The reduced sample is due to the fact that some countries in the full sample do not grow rice.

TABLE C.3: LIST OF RICH AND POOR COUNTRY SAMPLES - PISA

Rich Countries	Poor Countries
Australia	Argentina
Austria	Azerbaijan
Belgium	Brazil
Canada	Bulgaria
Denmark	Chile
Finland	Costa Rica
France	Croatia
Germany	Czech Republic
Iceland	Estonia
Ireland	Greece
Israel	Hungary
Italy	Indonesia
Japan	Jordan
Luxembourg	Kazakhstan
Netherlands	Korea, Rep.
New Zealand	Kyrgyz Republic
Norway	Latvia
Qatar	Lithuania
Singapore	Mexico
Spain	Montenegro
Sweden	Peru
Switzerland	Poland
United Kingdom	Portugal
United States	Romania
	Russian Federation
	Serbia
	Slovak Republic
	Slovenia
	Thailand
	Trinidad and Tobago
	Tunisia
	Turkey
	Uruguay
	Vietnam