Handle With Care: The Local Environmental Costs of Coal Piles at U.S. Power Plants

Akshaya Jha and Nicholas Z. Muller[*†]

December 27, 2017

Abstract

Current and proposed policies subsidize the on-site storage of coal at U.S. power plants; many criticize these policies because mining, transporting, and burning coal are all known to have environmental costs. Our paper demonstrates that the storage and handling of coal also emits significant quantities of fine particulates ($PM_{2.5}$) which adversely impact the health of local populations. We find that a 10% increase in the coal stockpiles held at U.S. power plants results in a 0.09% increase in average $PM_{2.5}$ concentration levels within 25 miles of these plants. Using coal stockpiles as an instrument, we show that a 10% increase in $PM_{2.5}$ leads to a 1.1% (3.2%) increase in average adult (infant) mortality rates. Using a value of statistical life approach, our estimates suggest that a one ton increase in coal stockpiles results in local air pollution costs of \$197. These costs disproportionately impact the economically disadvantaged communities living near power plants.

[*]Akshaya Jha: H. John Heinz III College, Carnegie Mellon University, 4800 Forbes Avenue, Pittsburgh, PA 15213. Email: akshayaj@andrew.cmu.edu. Nicholas Z. Muller: Tepper School of Business and Department of Engineering and Public Policy, Carnegie Mellon University, and NBER, 5000 Forbes Avenue, Pittsburgh, PA, 15213. Email: nicholas.muller74@gmail.com.

[†]The authors would like to thank Max Auffhammer, Prashant Bharadwaj, Jim Bushnell, Steve Cicala, Karen Clay, Lucas Davis, Meredith Fowlie, Katrina Jessoe, David Rapson, Jim Sallee, and Wally Thurman as well as seminar and conference participants (listed chronologically) at Penn State University, the Midwest Energy Fest, the Property and Environmental Research Center (PERC) Seminar, UC Berkeley, UC Davis, the Heartland Workshop, and the Workshop on Environmental and Energy Policy hosted by Utah State University. Akshaya would also like to gratefully acknowledge that work on this project was supported by the Lone Mountain Fellowship provided by PERC. Any remaining errors are our own.

1 Introduction

Roughly 74% of U.S. coal-fired electricity generation came from sources under output price regulation in 2016; coal-fired power plants under output price regulation typically receive a working capital allowance from regulators based on the level of coal stockpiles held on-site. Moreover, the Grid Resiliency Pricing Rule (GRPR) proposed by the United States Department of Energy on September 9^{th}, 2017 would require wholesale electricity market operators to subsidize power plants that store at least 90 days of fuel on-site. The stated goal of these subsidies is to increase the reliability of electricity supply. However, subsidizing the storage of coal has been intensely criticized because the reliance on coal as an energy source has a multitude of well-known environmental consequences.[1] For example, burning coal emits global pollutants which contribute to climate change as well as criteria air pollutants which affect the health status of local populations; mining coal can cause acid mine drainage and the ecological impacts of mountaintop removal are significant.

However, proponents of coal use in the United States argue that coal is already subject to a plethora of regulations that mitigate the environmental consequences from the mining, transportation, and burning of coal. For example, the effects of current mining operations as well as abandoned mine sites are managed by the Surface Mining Control and Reclamation Act of 1977. Transportation of coal by trains, trucks, or barges is governed by fuel and emission standards set by the United States Environmental Protection Agency. Finally, the combustion of coal for power generation and manufacturing is regulated by the Clean Air Act. Seemingly every stage of the coal supply chain from coal production at mines to coal burned by power plants is subject to a policy constraint.

This paper provides evidence that the process of coal storage and handling itself emits significant quantities of fine particulates which adversely impact the health of local populations. We argue for the necessity of environmental policy targeted specifically at mitigating the emissions from coal storage and handling, especially in light of current and proposed policies that subsidize coal storage at U.S. power plants. To substantiate this call for the management of the emissions from coal stockpiles, this paper quantifies the local air pollution costs associated with the storage and handling of coal at U.S.

[1] See NRC and NAS (2010) for a comprehensive report on the environmental externalities associated with energy use.

power plants. We do so in three steps. The first part of our analysis demonstrates that increases in the level of coal stockpiles held at a power plant results in significant increases in the concentrations of fine particulates ($PM_{2.5}$) within 25 miles of this plant.[2] We next estimate how $PM_{2.5}$ concentration levels affect adult and infant mortality rates, using variation in local air pollution due to changes in coal stockpiles (an instrumental variables analysis). Finally, we combine our estimates relating coal storage to $PM_{2.5}$ and $PM_{2.5}$ to mortality rates, monetizing the increased mortality due to coal-storage-induced increases in local air pollution using a Value of Statistical Life (VSL) approach.

Our paper combines several sources of data for the sample period 2002-2012. First, we use monthly, plant-level data on coal purchases and stockpiles provided by the Energy Information Administration. These data are linked to monthly average $PM_{2.5}$ concentration levels at roughly 1,000 monitored sites across the United States, which are collected from the Air Quality System database maintained by the United States Environmental Protection Agency. Weather conditions play a crucial role in dictating $PM_{2.5}$ levels; moreover, wind speed, wind direction, and precipitation are important determinants of the extent to which coal stockpiles generate $PM_{2.5}$. Due to this, our analysis incorporates monthly meteorological data collected from roughly 1,600 monitors across the United States from the National Climatic Data Center. Finally, we assess the extent to which $PM_{2.5}$ affects mortality using county-by-month specific mortality rates provided by the Centers for Disease Control and Prevention (CDC Wonder, 2016).

To estimate the impact of coal stockpiles on local $PM_{2.5}$ concentration levels, we specify a spatial econometric framework that matches coal-fired power plants to all air quality (AQ) monitoring stations within a given distance radius. This framework allows us to test for the distance at which plants' coal storage behavior ceases to affect pollution levels. We find that a 10% increase in coal stockpiles results in a 0.09% increase in average $PM_{2.5}$ concentration levels for power plants within 25 miles of their matched air quality monitor. In contrast, there is no statistical link between coal stockpiles and $PM_{2.5}$ levels measured more than 25 miles away from the plant; the effect of coal storage on $PM_{2.5}$ concentration levels is highly localized. This is intuitive, given that we posit that the emissions from coal storage and handling stem primarily from wind blowing over

[2] Jaffe et al. (2015) studies the emissions of diesel particulate matter (DPM) and coal dust from trains in the Columbia River Gorge in Washington State; they find that the passage of a diesel powered open-top coal train results in nearly twice as much respirable $PM_{2.5}$ compared to the passage of a diesel-powered freight train not carrying coal in an open hopper.

uncovered coal piles, gaseous emissions from these coal piles, as well as the displacement of coal when it's moved; these emissions occur at ground level and are therefore not likely to be entrained in upper-level winds. In contrast, the environmental impacts of coal combustion on $PM_{2.5}$ are primarily regional because the emissions from coal combustion are released from tall smokestacks.[3]

Importantly, conditional on our rich set of covariates, we find no statistical link between the coal stockpiles held at power plants and locally monitored carbon monoxide (CO), sulfur dioxide (SO_2), or nitrogen dioxide (NO_2) levels. This provides strong evidence that our estimated relationship between coal storage and $PM_{2.5}$ is not capturing any combustion-based source of $PM_{2.5}$ emissions such as the coal burned by power plants or the fuel burned by the trains, trucks, and barges carrying coal to power plants. Finally, our estimated effect of coal stockpiles on $PM_{2.5}$ is higher: (1) for areas downwind from a plant, and (2) for lower levels of precipitation.[4] These empirical results using spatiotemporal variation in weather, our placebo tests considering other pollutants (CO, SO_2, and NO_2) as well as our battery of controls give us confidence that we have causally identified how changes in the level of unburned coal piles affect $PM_{2.5}$.

We next estimate the impact of $PM_{2.5}$ concentration levels on mortality rates using variation in $PM_{2.5}$ arising from plants' coal storage behavior. This instrumental variables (IV) strategy allows us to overcome two important challenges in identifying the effect of $PM_{2.5}$ on mortality. First, mortality rates and local air pollution are both undoubtedly tied to local economic conditions; we argue that power plants purchase and store coal on the basis of regional factors, such as electricity demand, coal prices, and natural gas prices, that do not vary with these local economic conditions. Also, as discussed in Greenstone and Hanna (2014) and Benmarhnia, Bharadwaj and Romero (2017), it is nearly impossible to find variation in polluting behavior or environmental regulations that affects only one pollutant; for example, burning fossil fuels emits CO, SO_2 and NO_2

[3]Clay, Lewis and Severnini (2015) examines the economic benefits versus environmental costs of coal-fired power plant openings in the United States from 1938-1962, a period in which smokestacks were significantly shorter than they currently are. Similar to us, Clay, Lewis and Severnini (2015) finds that the local air pollution costs from coal-fired power plants are concentrated primarily within 30 miles of these plants.

[4]There is a growing literature that utilizes wind direction to identify how local air pollution affects economic and environmental outcomes; among others, this identification strategy is used by Anderson (2015) to examine how long-term air pollution exposure affects mortality, Herrnstadt and Muehlegger (2015) to study the effect of air pollution on crime, and Deryugina et al. (2016) to investigate how air pollution affects health outcomes and health utilization for the elderly.

in addition to $PM_{2.5}$. We demonstrate that the storage and handling of coal emits $PM_{2.5}$ but not these other pollutants, allowing us the unique opportunity to isolate the impact of a single pollutant ($PM_{2.5}$) on mortality rates. We construct these mortality rates using data from the Centers for Disease Control and Prevention (CDC) on annual, county-level total number of deaths associated with: (1) the cardiovascular system, (2) respiratory system, (3) any cause for people over 30 years old ("adults"), (4) any cause for children aged 0 to 4 ("infants"), and (5) external causes such as accidents as a placebo test.[5] Our specifications include a bevy of controls as well as county-year fixed effects. We find an economically small and negative association between $PM_{2.5}$ and mortality rates if we simply estimate this relationship using ordinary least squares. However, our estimated effects of $PM_{2.5}$ on all of the aforementioned mortality rates, excepting the external-cause mortality rate, are positive, statistically significant, and economically significant when we instrument using the monthly level of coal stockpiles held at nearby power plants. Of particular note, we find that a 10% increase in $PM_{2.5}$ leads to a 1.1% (3.2%) increase in average overall adult (infant) mortality rates.

Our estimated semi-elasticities indicate that a one microgram per cubic meter increase in $PM_{2.5}$ causes a 1.1% increase in adult mortality rates; the two epidemiological studies most commonly used by the United States Environmental Protection Agency (USEPA) estimate that a one microgram per cubic meter increase in $PM_{2.5}$ increases average mortality rates by 0.6% (Krewski et al., 2009) and 1.3% (Lepeule et al., 2012). However, recent work in economics has challenged the causal basis for the link between exposure to PM and increases in mortality risk reported in epidemiological studies.[6] Thus, we contribute both to the economics literature testing for a causal relationship between $PM_{2.5}$ and mortality rates as well as the epidemiological literature used by policymakers such as the USEPA to quantify the impacts of $PM_{2.5}$ emissions. This estimated relationship is arguably the most important component of any cost-benefit analysis of any environmental policy designed to improve air quality.[7]

[5] Exposure to fine particulates has been linked specifically to cardiovascular and respiratory conditions; Nel (2005) describes the biological mechanisms underpinning this link.

[6] This work includes, among others, Chay and Greenstone (2003), Currie, Neidell et al. (2005), Currie, Neidell and Schmieder (2009), Chen et al. (2013), and Dominici, Greenstone and Sunstein (2014).

[7] Previous empirical research finds that the damages from air pollution are significant in magnitude and that these damages are mostly due to the increased mortality risk from exposure to $PM_{2.5}$; for examples, see NRC and NAS (2010), EPA (2010), Muller, Mendelsohn and Nordhaus (2011), and Muller (2014).

Finally, we combine our estimates of the effect of coal stockpiles on $PM_{2.5}$ with our IV estimates of the effect of $PM_{2.5}$ on adult and infant mortality rates in order to calculate the increased number of deaths due to the $PM_{2.5}$ emitted from coal storage and handling. These health risks are monetized using the Value of Statistical Life (VSL) approach ((EPA, 1999); (EPA, 2010)). We calculate that the local air pollution cost per ton of coal stockpiled is $197. Roughly 86% of this external cost is due to increases in adult mortality while 14% is from infant mortality since far more adults are exposed to $PM_{2.5}$ relative to infants. Our air pollution cost estimates are sizable given that the average U.S. coal-fired power plant pays roughly $48 per ton for coal and stockpiles 212,782 tons of coal. However, translating tons of coal into MWh of electricity, our local air pollution costs are $106 per MWh-equivalent of coal stored. Levy, Baxter and Schwartz (2009) finds that the external cost of increased $PM_{2.5}$ exposure from burning coal at U.S. power plants is roughly $230 per MWh. Thus, our empirical estimates of the local environmental damages from storing coal are sizable, but not unreasonably so when compared to estimates of the local air pollution costs of burning coal. Finally, unlike the environmental impacts from burning coal, the air pollution costs of storing coal are borne primarily by people living within 25 miles of coal-fired power plants. People living in census tracts with power plants have lower per-capita incomes and educational attainment on average relative to people living in census tracts without power plants; thus, the highly localized environmental costs of coal storage and handling disproportionately affect economically disadvantaged communities.[8]

Coal is also transported by trains in uncovered cars as well as handled and stored at mines and coal export terminals; people living near mines, rail tracks, or export terminals are likely also exposed to higher levels of $PM_{2.5}$ due to the storage and handling of unburned coal. Coal export terminals in particular are located in densely populated cities such as Los Angeles, Houston, and Baltimore; our empirical connection between unburned coal and ambient $PM_{2.5}$ informs the ongoing debate regarding the potential adverse health impacts of these terminals on local communities.[9] In particular, the final

[8]Davis (2011) similarly finds that neighborhoods near fossil-fuel fired power plants have lower average household income and educational attainment using restricted-access census microdata from 1990 and 2000.

[9]There is staunch local opposition to proposed coal export terminals in Washington (near Bellingham and Longview) and California (in Oakland) due in part to concerns about the negative health consequences from coal dust; these proposed terminals have not been built as of March 2017. Moreover, the city of Oakland voted to ban the transportation of coal within their jurisdiction; a recent New York Times article (Fuller, 2016) provides more details on this ban.

section of our paper argues for the exploration of policies aimed at mitigating the $PM_{2.5}$ from coal storage; potential low-cost policy interventions range from frequently spraying down coal piles with a chemical solvent (known as a "topper") to simply covering coal piles. Given that the effect of coal piles on $PM_{2.5}$ is highly local, policymakers at the municipal, county, or state-level can intervene if federal regulators are unwilling to act. The need for policy aimed at mitigating the local air pollution from coal storage and handling is increasingly urgent given current and proposed regulations that subsidize power plants holding higher levels of coal stockpiles.

The remainder of the paper proceeds as follows. In Section 2, we provide background on the coal and electricity generation sectors in the United States. This section also elaborates on the physical process by which plants' coal storage and handling behavior results in higher $PM_{2.5}$ concentration levels. Section 3 describes the data sources and methodology used to estimate the relationship between plants' coal stockpiling behavior and $PM_{2.5}$ concentration levels, while Section 4 presents our empirical findings demonstrating that increases in coal stockpiles result in increases in $PM_{2.5}$ concentration levels. Our instrumental variables approach for identifying the effect of $PM_{2.5}$ concentration levels on mortality rates is discussed in Section 5. In Section 6, we quantify the local environmental health costs of increases in average mortality rates due to the $PM_{2.5}$ emissions from coal piles. Finally, we conclude in Section 7 by illustrating how our findings are applicable to current policy debates, such as the environmental costs of coal dust at export terminals and the environmental costs of economic policy that subsidizes coal storage at U.S. power plants.

2 Economic and Environmental Background

2.1 Input Coal Purchase and Storage Behavior of U.S. Power Plants

Coal is not homogeneous; coal mined in different regions of the United States differs across many different dimensions such as heat content, sulfur content, and ash content. The vast majority of the coal mined in the United States is burned by power plants in order to generate electricity.[10] Thus, power plants primarily value the heat content of coal; burning a ton of coal with higher heat content generates a larger amount of output electricity. Coal with a lower sulfur content or ash content produces less emissions when burned; plants are willing to pay a premium for coal with a lower sulfur or ash content only to the extent they need to reduce emissions in order to comply with environment regulations.

Power plants in the United States purchase the majority of their input coal from long-term contracts with coal suppliers, purchasing the remainder from spot markets. These coal supply contracts typically feature coal prices that are indexed to measures outside of either party's control, such as productivity indices and/or cost indices associated with coal mining.[11] Power plants also enter into separate short or long term agreements with transportation companies; roughly 67 percent of coal is transported from mine to plant by rail. The remaining coal is shipped by barge (12%), trucks (10%) and various other modes of transportation used primarily for short distances such as conveyors and pipelines.[12] Finally, the electricity produced by power plants is injected into the transmission grid and distributed regionally based on where it is needed.

Summarizing, plants purchase and store coal based primarily on four factors: regional/grid-level electricity demand, contract and spot coal prices, as well as natural gas prices[13] (Jha, 2017*a*). Importantly, these factors are not affected by the local economic activity around the power plant. This is vital because both local air pollution and mortality rates are undoubtedly correlated with this nearby economic activity. Thus, in order to identify the effect of local air pollution on mortality rates, we need an instrument that impacts mortality rates only through its effect on local air pollution; we argue that plant-level coal storage behavior satisfies this exclusion restriction.

[10] Roughly 93% of the total amount of coal consumed in the United States in 2012 was burned in order to produce electricity. This statistic is from the Annual Energy Review released by the Energy Information Administration (EIA).

[11] See Joskow (1988), Joskow (1990), or Jha (2017*b*) for more details on the structure of coal supply contracts.

[12] These statistics are for the year 2013 and come from EIA's "Today In Energy": http://www.eia.gov/todayinenergy/detail.cfm?id=16651.

[13] Natural gas fired electricity generation is the primary substitute to coal-fired electricity generation during our 2002-2012 sample period. Roughly 30% (37%) of overall U.S. electricity generation came from natural gas fired (coal-fired) sources in 2012.

2.2 Current and Proposed Economic Policies Affecting Coal Storage

74% of coal-fired electricity generation in the United States came from sources under output price regulation in 2016. Under output price regulation, the regulator is obligated to set the electricity price charged by the firm in order to allow that firm an opportunity to recover its costs, including a return on capital. Regulators in practice typically provide the plant with a working capital allowance based on how much coal they store on-site. Consistent with the intuition that this working capital allowance incentivizes plants under output price regulation to store more coal, Jha (2017*a*) demonstrates within a matched difference-in-differences framework that regulated plants hold 13% more coal stockpiles on average relative to similar plants facing electricity market mechanisms. Though plants facing electricity market mechanisms do not currently receive a subsidy for storing coal, the Federal Energy Regulatory Commission, at the behest of the U.S. Department of Energy, is currently exploring a rule that would subsidize power plants that keep 90 days of fuel on site (the so-called "90-day rule"). This paper argues that there are significant local air pollution costs associated with storing and handling unburned coal; these environmental costs are exacerbated by economic policies that subsidize the storage of coal at power plants.

2.3 Environmental Impacts of U.S Coal-fired Generation

It is well-known that NO_x, SO_2, and $PM_{2.5}$ emissions result in elevated mortality risk among exposed populations. Earlier work has shown that U.S. coal-fired power plants emit significant levels of NO_x, SO_2, and $PM_{2.5}$ when burning coal. These emissions caused approximately 27,000 deaths in the United States in 1999. However, emissions-induced deaths resulting from power plants burning coal fell to roughly 9,500 in 2011 due to differences in the type of coal burned by plants as well as the increased prevalence of scrubbing technology (Muller, 2014).

While the environmental costs of burning coal are well-documented (NRC and NAS, 2010), this paper explores the local environmental costs of the emissions from coal stockpiles. We posit two mechanisms for emissions from these uncovered piles of coal. The first is wind erosion. Namely, wind blowing over uncovered coal stockpiles entrains fine particulates; these passive, or fugitive, dust emissions become a constituent of ambient $PM_{2.5}$. Second, coal stockpiles emit volatile gases. Specifically, coal in open stockpiles undergoes oxidation, which releases a set of pollutants including hydrocarbons and sulphuric gases (Zhang, 2013). The gases result in the formation of secondary organic $PM_{2.5}$ that is a constituent of the ambient $PM_{2.5}$ collected at monitoring stations. Finally, the coal handling process at power plants can also entrain fine particulates; this process includes offloading coal from the trains, trucks, and barges delivering coal to the plant, moving coal to different places within the plant site, separating "light dust" from coal, and crushing coal in order to make it suitable for burning.[14]

Our empirical strategy does not allow us to dis-entangle what portion of our estimated link between $PM_{2.5}$ and coal stockpiles comes from the wind erosion of coal piles, volatiles gases from these piles, or the coal handling process; local populations exposed to the $PM_{2.5}$ emissions from any of these three sources face adverse health consequences and ultimately increased risk of mortality. However, we demonstrate that our estimated impact of coal stockpiles on $PM_{2.5}$ is *not* capturing the combustion of any fuel, whether it be the power plant burning coal in order to generate electricity or the diesel burned by the trains, trucks, and barges delivering coal to the plant.[15] In particular, we show that there's no statistical impact of coal stockpiles on locally monitored sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and carbon monoxide (CO) emissions, noting that the combustion of any fuel emits SO_2, NO_2, and CO in addition to $PM_{2.5}$. Importantly, the fact that coal storage and handling emits $PM_{2.5}$ but not these other pollutants gives us the unique opportunity to examine how a single pollutant ($PM_{2.5}$) affects mortality rates using the level of coal stockpiles held at U.S. power plants as an instrument for $PM_{2.5}$.

[14]An article on Electrical Engineering Portal (Raman, 2012) provides a brief description of the coal handling process at power plants.

[15]Using data on crude oil shipments originating in North Dakota in 2014, Clay et al. (2017) finds that the air pollution cost associated with freight rail transportation is sizable; Clay et al. (2017) also shows that the majority of this air pollution cost associated with trains burning diesel comes from NO_x emissions.

3 Local Environmental Impacts of Coal Storage: Data and Methodology

This section describes the data used to estimate the link between the monthly level of coal stockpiles held at each plant and $PM_{2.5}$ concentration levels at nearby air quality monitors. We also present the empirical framework we use to measure this relationship, which includes specifying our set of controls for alternative sources of $PM_{2.5}$, such as the burning of coal, as well as other factors that increase or decrease the $PM_{2.5}$ emitted from a given set of sources, such as wind speed, wind direction, and precipitation.

3.1 Data Sources

We use monthly, plant-level data from 2002-2012 on end-of-month fuel inventories and fuel purchases from the Energy Information Administration (EIA).[16] Regarding plants' coal purchases, we have order-level data on the month of purchase, quantity purchased, delivered price, heat content, sulfur content, ash content, and the coal's county of origination. We only consider electricity generation plants whose primary business purpose is the sale of electricity to the public; this excludes plants that also sell significant quantities of heat ("combined heat and power plants") as well as commercial and industrial plants that generate electricity for their own use.

We also use the Air Quality System (AQS) data provided by the United States Environmental Protection Agency (USEPA). This publicly available database includes hourly readings of ambient $PM_{2.5}$ concentrations at roughly 1,000 monitored sites across the contiguous United States. We aggregate these data to obtain monthly average $PM_{2.5}$ levels for each air quality monitor for the sample period 2002-2012.

Our meteorological controls come from the quality controlled local climatological data (QCLCD) collected by the National Climatic Data Center (NCDC). These data include hourly wind speed and direction, dry bulb temperature, wet bulb temperature, dew-point temperature, relative humidity, station pressure, and precipitation at approximately 1,600 locations. We aggregate these data to the meteorological monitor/month-

[16]Monthly, plant-level data on fuel inventories for 2002-2012 are confidential; we obtained a research contract from the EIA in order to use these restricted-access data.

of-sample level by taking time-weighted averages over hours of dry bulb temperature, wet bulb temperature, dew-point temperature, relative humidity, and station pressure; we use the meteorological monitor/month-of-sample level sum of hourly precipitation. Wind speed is an important factor in determining both how much $PM_{2.5}$ is generated from various sources as well as how this $PM_{2.5}$ is dispersed. Thus, we also control for the (5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95) hourly percentiles of wind speed, calculated over all hours-of-sample for each meteorological monitor/month-of-sample. Finally, we use the wind-speed weighted monthly average wind direction measured at each meteorological monitor.

We additionally control for a variety of other factors that can affect monthly $PM_{2.5}$ concentration levels. First, the USEPA's Continuous Monitoring Emissions System (CEMS) collects hourly data from each plant's smokestack on SO_2, CO_2, and NO_x emissions (in tons) resulting from coal burned; we sum these hourly data to the monthly level and control for the total SO_2, CO_2, and NO_x emissions from each plant in each month-of-sample. We also control for the monthly total quantity of coal purchased by each plant as well as the monthly total electricity generation produced by each plant; both of these variables come from EIA data.

The AQS database provides the latitude and longitude for each air quality monitor, the NCDC database provides the latitude and longitude for each meteorological monitor, and the EPA eGrid database provides the latitude and longitude for each coal-fired power plant. We use these geocoded locations to merge air quality monitors to coal-fired power plants and meteorological monitors.

3.2 Data Merge

We merge each air quality monitor i to meteorological monitors and coal-fired power plants as follows:

1. For each month-of-sample, we find all meteorological monitors within M miles of air quality monitor i. We take a weighted average of the meteorological data (for example, wind speed and wind direction) across these meteorological monitors for each air quality monitor i, where we weight by the inverse of the distance between the air quality monitor and the meteorological monitor.

2. If $M = 25$ miles, we consider all coal-fired power plants less than 25 miles away from air quality monitor i. If $M = 50$ miles, we consider all coal-fired power plants between 25 miles and 50 miles away from air quality monitor i.

Thus, our unit of observation is an air quality monitor/power plant pair for each month-of-sample, emphasizing that each air quality monitor can be linked to multiple power plants for a given month-of-sample. We examine how the effect of coal stockpiles on $PM_{2.5}$ concentration levels decays with distance by separately estimating these effects for plants within 25 miles of their air quality monitor versus plants between 25 miles and 50 miles away from their air quality monitor. We also show how the effect of coal storage on $PM_{2.5}$ varies with the relative wind direction between air quality monitor and power plant (read: upwind versus downwind) as well as locally-monitored precipitation levels. We provide further details on our data sources and data construction in Appendix Section B.

3.3 Empirical Framework

This subsection describes the regression specifications used to estimate the effect of coal stockpiles held at power plants on locally monitored $PM_{2.5}$ concentration levels. We consider both "Log-Log" and "Levels-Levels" specifications linking coal storage to $PM_{2.5}$. Our Log-Log specification is:

$$log(PM2.5_{i,t}) = \alpha_{c,y} + \theta_{i,p} + \delta_t + \gamma log(CS_{p,t}) + \psi log(1 + \sum_{k=1,k\neq p}^{P_{i,t}} CS_{k,t}) + X_{i,p,t}\beta + \epsilon_{i,p,t} \quad (1)$$

where a unit of observation is an air quality monitor i and a coal-fired power plant p in month-of-sample t. $P_{i,t}$ is the number of coal-fired power plants merged with each air quality monitor i in each month t.[17,18]

We control for whether plant p received any coal deliveries in month-of-sample t.

[17]This Log-Log specification excludes the small number of observations for which monthly average $PM_{2.5}$ levels are zero; our empirical results are similar if we instead consider $log(PM_{2.5} + 1)$ as the dependent variable.

[18]Coal-fired power plants only stock-out in roughly 0-5% of the months-of-sample for which we observe them; thus, we drop observations with zero stockpiles for the Log-Log specification. The empirical results are similar if we instead keep these zero stockpile observations in the sample; for these specifications, we include $log(CS_{p,t} + 1)$ and $1(CS_{p,t} > 0)$ (an indicator variable for coal stockpile levels greater than zero) rather than $log(CS_{p,t})$.

This term accounts for the $PM_{2.5}$ emissions both from the fuel burned by the trains, trucks or barges delivering coal as well as the $PM_{2.5}$ emissions associated with handling newly delivered coal at the plant site. We also control for the total amount of coal stockpiles held by all plants $k \neq p$ ($\sum_{k=1,k\neq p}^{P_{i,t}} CS_{k,t}$) so that the $log(CS_{p,t})$ term doesn't capture positive correlations in stockpile increases across plants. For example, if a one ton increase in coal stockpiles at plant p is typically associated with a 0.5 ton increase in coal stockpiles at plant q, we do not want to include the effect of the 0.5 ton increase at plant q on locally monitored $PM_{2.5}$ in our estimate of γ.

Our regression specification controls for a wide variety of alternative factors that affect local $PM_{2.5}$ levels, such as the annual averages of the sulfur content and ash content of coal purchased by the plant, the plant's monthly total electricity generation, the monthly total quantity of coal purchased by the plant, the plant's monthly total SO_2, CO_2, and NO_x emissions from coal burned (in tons), wind speed, dry bulb temperature, wet bulb temperature, dew-point temperature, relative humidity, station pressure, and precipitation. We also include fixed effects for each county-of-air-quality-monitor/year-of-sample ($\alpha_{c,y}$), air quality monitor/power plant ($\theta_{i,p}$), and month-of-sample (δ_t). We report a specification considering one overall coefficient for coal stockpiles as well as specifications allowing γ to vary by wind direction and precipitation.

Our goal is to estimate how a *one ton* increase in coal stockpiles at a plant affects the *level* of $PM_{2.5}$ measured by nearby air quality monitors, controlling for all of the other factors listed above. The Log-Log specification in Equation 1 implies that this partial effect in levels for observation (i, p, t) is:

$$\frac{dP\hat{M}2.5_{i,t}}{dCS_{p,t}} = \hat{\gamma}(\frac{PM2.5_{i,t}}{CS_{p,t}})$$

We also directly estimate this partial effect in levels. Our "Levels-Levels" specification, presented below, additionally serves as a sensitivity check of the functional form relating coal stockpiles to $PM_{2.5}$:

$$PM2.5_{i,t} = \alpha_{c,y} + \theta_{i,p} + \delta_t + \gamma CS_{p,t} + \psi \sum_{k=1,k\neq p}^{P_{i,t}} CS_{k,t} + X_{i,p,t}\beta + \epsilon_{i,p,t}$$

This specification includes the same set of controls as listed above; the partial effect relating coal stockpiles to $PM_{2.5}$ is simply $\hat{\gamma}$ in the Levels-Levels case.

For all of our specifications, we cluster standard errors at the air quality monitor level and weight by the inverse distance between air quality monitor and power plant.[19]

4 Local Environmental Impacts of Coal Storage: Empirical Findings

This section presents our empirical results regarding the link between coal stockpiling behavior and $PM_{2.5}$ concentration levels. In particular, we find that a 10% increase in coal stockpiles results in a 0.09% increase in average $PM_{2.5}$ concentration levels for populations within 25 miles of power plants. In contrast, our estimates indicate that there is no average effect of coal stockpiles on $PM_{2.5}$ when examining plants farther than 25 miles away from their corresponding air quality (AQ) monitor. This is intuitive, given that the $PM_{2.5}$ emissions associated with the wind erosion of coal piles, the volatile gases emitted from these piles, as well as the coal handling process at plants, occur at or near ground level and can thus only travel so far.

We should also expect more severe increases in $PM_{2.5}$ concentrations for local populations downwind from coal piles. Consistent with this intuition, we estimate that the average $PM_{2.5}$ increase from coal storage is larger for AQ monitors that are downwind of nearby coal-fired power plants. Also, precipitation is known to reduce ambient $PM_{2.5}$ concentration levels; we show that the average $PM_{2.5}$ increase from coal stockpiles is lower for higher levels of precipitation. Finally, we consider a variety of robustness checks in order to demonstrate that our empirical results are not due to the emissions from the coal burned by power plants or the fuel burned by the trains, barges, and trucks delivering coal to these plants. For example, carbon monoxide (CO), sulfur dioxide (SO_2) and nitrogen dioxide (NO_2) are emitted when any fuel is burned; we demonstrate that there is no statistical effect of coal stockpiles on the average concentrations of CO, SO_2 and NO_2 measured within 25 miles of the power plant. Our estimated effect of coal storage on $PM_{2.5}$ is also very similar for specifications where we include flexible controls for each

[19] We re-ran our regressions: (1) weighting each observation by the inverse of the number of power plants matched to each air quality monitor, and (2) not weighting at all. Our empirical findings are very similar to those reported below for both of these alternative weighting schemes.

plant's monthly total electricity generation. Summarizing, we are confident that we've identified a causal relationship between coal stockpiles and $PM_{2.5}$ for three reasons. First, we find no effect of coal stockpiles on CO, SO_2 and NO_2, which are pollutants emitted when burning coal but not storing or handling it. Moreover, our regression specifications control for an extensive array of other factors that affect $PM_{2.5}$. Finally, consistent with predictions from the environmental science literature, we find that increasing the size of coal piles results in higher levels of $PM_{2.5}$ for air quality monitors downwind of power plants and locations/months-of-sample with lower levels of precipitation.

4.1 Overall Effect of Coal Stockpiles on $PM_{2.5}$ Concentrations

The top (bottom) panel of Table 1 displays the results from the Levels-Levels (Log-Log) specification measuring the impact of the coal stockpiles held at U.S. power plants on locally monitored $PM_{2.5}$ concentration levels. The summary statistics corresponding to all of the regressions discussed in this section are in Appendix Tables A.1 and A.2. The unit of observation for these regressions is an air quality monitor/power plant pair in a given month-of-sample.

The top panel of Table 1 reports the results from the Levels-Levels specification where the size of the coal stockpiles held at each power plant in each month-of-sample (in tons) is the covariate of interest. We see from this table that a 100,000 ton increase in coal stockpiles increases average ambient $PM_{2.5}$ levels by 0.1 micrograms per cubic meter when considering only plants within 25 miles of their corresponding air quality (AQ) monitor (Column 1). However, there is no statistical link between coal stockpiles and $PM_{2.5}$ when focusing on plants between 25-50 miles of their AQ monitor (Column 2). This result is intuitive because the $PM_{2.5}$ emissions associated with the wind erosion of stationary coal stockpiles, the volatile gases emitted from these coal piles, and the coal handling process occur at or near ground-level; these ground-level $PM_{2.5}$ emissions are less likely to travel long distances relative to the emissions from the power plant's smokestack.

We draw the same qualitative conclusions when examining the bottom panel of Table 1, which shows the estimated elasticities between coal stockpiles and $PM_{2.5}$ using the Log-Log specification. In particular, a 10% increase in coal stockpiles results in a 0.09%

increase in average $PM_{2.5}$ concentration levels for plants within 25 miles of their AQ monitor. As before, there is no average effect of coal stockpiles on $PM_{2.5}$ for plants between 25-50 miles of their AQ monitors, which makes sense given that we do not typically expect the $PM_{2.5}$ emissions from ground-level coal piles to travel very far.

[Table 1 about here.]

4.2 The Effects of Coal Stockpiles on $PM_{2.5}$ Concentrations By Wind Direction

In this subsection, we interact coal stockpiles with the relative bearing from each coal-fired power plant to its corresponding air quality (AQ) monitor. A relative bearing of 0° means that the wind is blowing directly from the power plant to the AQ monitor, while a relative bearing of 180° means that the wind is blowing directly from the AQ monitor to the power plant. For each plant/air quality (AQ) monitor pair in each month-of-sample, we code the AQ monitor as "downwind" from the plant if their relative bearing is less than 90° and code the AQ monitor as "upwind" from the plant if their relative bearing is greater than 90°.

[Table 2 about here.]

The top panel of Table 2 displays the results for the directional models in which the size of coal stockpiles (in tons) is the independent variable of interest; the bottom panel of this table presents our findings instead considering the log of coal stockpiles. Focusing on the specifications considering plants within 25 miles of their AQ monitor (Column 1), we see that the coefficient estimate on the interaction between coal stockpiles and an indicator for downwind monitors is larger than the coefficient estimate on the interaction between coal stockpiles and an indicator for upwind monitors for both the Levels-Levels (top panel) and Log-Log (bottom panel) specifications; as expected, the positive effect of coal stockpiles on $PM_{2.5}$ concentration levels is greater downwind from plants relative to upwind from plants. Our estimated effect of coal stockpiles on downwind $PM_{2.5}$ concentration levels is statistically significant for both the Levels-Levels and Log-Log specifications, while the positive effect of coal stockpiles on upwind $PM_{2.5}$ is statistically significant only for the Levels-Levels specification. As described in Section 2, $PM_{2.5}$ particulates are generated both from wind blowing over coal piles as well as from the volatile gases emitted by these piles. Our empirical findings are consistent with the intuition that these $PM_{2.5}$ particulates are transported from power plant to AQ monitor via downwind currents.

However, for both the Levels-Levels specification (top panel of Table 2) and the Log-Log specification (bottom panel of Table 2), neither the upwind nor downwind coefficient estimates are statistically significant for plants between 25-50 miles from their AQ monitor (Column 2). This is to be expected; as discussed above, the $PM_{2.5}$ emissions from coal storage and handling, whether it be wind blowing over the coal pile, gaseous emissions from this pile, or the coal handling process on-site, occur at ground-level. Therefore, we should not typically expect the $PM_{2.5}$ emissions from coal storage and handling to travel very far.

4.3 The Effect of Coal Stockpiles on $PM_{2.5}$ Concentrations Interacted with Precipitation

In this subsection, we explore how the effect of coal stockpiles on $PM_{2.5}$ varies with monthly total precipitation as measured by the set of meteorological monitors within M miles of the plant's corresponding air quality (AQ) monitor. For example, if there are four meteorological monitors within M miles of a given AQ monitor, we take the monthly sum over the hourly precipitation data for each meteorological monitor and then take the inverse-distance weighted average over all four of these meteorological monitors for each month-of-sample.[20] We consider $M = 25$ for the within 25 mile bandwidth specifications and $M = 50$ for the 25-50 mile bandwidth specifications.

[Table 3 about here.]

Table 3 displays the results when interacting monthly, plant-level coal stockpiles with an indicator for whether there was any precipitation in that month; the top panel of this table presents our findings using the Levels-Levels specification (both coal stockpiles

[20] The empirical results are very similar if we instead take an unweighted average across meteorological monitors for each air quality monitor/month-of-sample.

and $PM_{2.5}$ are in levels) and the bottom panel presents our findings using the Log-Log specification (both coal stockpiles and $PM_{2.5}$ are in logs). We see from Table 3 that the interaction term between the indicator for positive monthly precipitation and coal stockpiles is statistically significant and negative when considering plants less than 25 miles away from their air quality monitor (Column 1); the average effect of coal stockpiles on locally monitored $PM_{2.5}$ levels is substantially larger for areas and months-of-sample without any precipitation. We similarly conclude that higher levels of precipitation mitigate the impact of coal stockpiles on $PM_{2.5}$ if we instead interact coal stockpiles with the log of total monthly precipitation (see Appendix Table A.3).

This finding is intuitive because local pollution levels are known to decrease with rainfall due to "wet deposition": $PM_{2.5}$ particulates are brought from the atmosphere to the ground by rain. Thus, if storing coal generates a given amount of $PM_{2.5}$, less of this $PM_{2.5}$ remains in the air for higher levels of monthly total precipitation. This highlights the importance of rainfall in reducing the level of local ambient $PM_{2.5}$ concentration levels. Our findings also provide suggestive evidence that spraying down coal piles with water may be a low-cost mitigation strategy for plants to reduce the amount of $PM_{2.5}$ emitted by these piles.

4.4 Robustness Checks and Sensitivity Analyses

One potential concern is that our estimates capture the $PM_{2.5}$ increases from burning coal rather than storing it. We consider two robustness checks in order to rule out this concern. First, we include more flexible controls for the monthly total amount of electricity generated at each coal-fired power plant; in particular, we control for linear, quadratic, and cubic terms for monthly, plant-level thermal generation as well as the log of monthly, plant-level thermal generation. Our empirical results including these controls (presented in Appendix Table A.4) are very similar to those presented above, indicating that variation in electricity generation is not driving our findings regarding the relationship between $PM_{2.5}$ and coal storage.

Second, we re-run our regressions considering carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2) as dependent variables rather than $PM_{2.5}$. Combustion of any fuel, be it coal, diesel, or gasoline, emits CO, SO_2, and NO_2 as well as

$PM_{2.5}$. Thus, we would expect to see a positive statistical effect of coal stockpiles on CO, SO_2, and NO_2 if our empirical findings are due to combustion-based sources of $PM_{2.5}$ such as the coal burned by the power plant or the fuel burned by trains, trucks, and barges delivering coal to the plant. Instead, we demonstrate in Table 4 that there is **no** statistical link between changes in coal stockpiles and locally monitored CO, SO_2, or NO_2 concentration levels for plants less than 25 miles away from their corresponding air quality (AQ) monitor; this is true whether we consider specifications where both the pollutant and coal stockpiles are in levels (the top panel of Table 4) or specifications where both the pollutant and coal stockpiles are in logs (the bottom panel of Table 4). Summarizing, Table 4 provides strong evidence that our effect of coal stockpiles on $PM_{2.5}$ is not capturing any alternative source of $PM_{2.5}$ based on combustion; combustion-based sources of $PM_{2.5}$ include the burning of coal by power plants in order to generate electricity or the burning of fuel by the trains, barges, and trucks delivering coal to these plants.

[Table 4 about here.]

One may also be concerned that our empirical results are driven by weighting our regressions by the inverse distance between air quality monitor and power plant. To alleviate this concern, we re-run our analyses: (1) weighting each observation by the inverse of the number of power plants matched to each air quality monitor, and (2) not weighting at all. We find similar results to those reported in the paper when using either of these two alternative weighting schemes.

5 The Effect of $PM_{2.5}$ on Mortality Rates

The previous section estimated the impact of coal stockpiles on $PM_{2.5}$. This section quantifies the impact of $PM_{2.5}$ concentration levels on mortality rates using coal stockpiles as an instrument for $PM_{2.5}$. The first subsection describes our data on mortality rates. We next specify the ordinary least squares (OLS) and instrumental variables (IV) regression frameworks used to measure the impact of $PM_{2.5}$ on mortality rates. Finally, we present our empirical results for both OLS and IV specifications. Summarizing our findings, we see an economically small (and sometimes negative) association between

$PM_{2.5}$ and mortality rates when examining the OLS results. However, the effect of $PM_{2.5}$ on mortality rates is positive, statistically significant, and economically significant when we instrument $PM_{2.5}$ with monthly, plant-level coal stockpiles. Among other results, we find that a 10% increase in $PM_{2.5}$ leads to a 1.1% (3.2%) increase in average overall adult (infant) mortality rates.

Both $PM_{2.5}$ and mortality rates are undoubtedly correlated with local economic activity; this potentially explains why OLS does not identify an effect of $PM_{2.5}$ on mortality rates. Our instrumental variables framework relies on the assumption that plants' coal storage behavior affects mortality rates only through its effect on local $PM_{2.5}$ concentration levels; for example, we assume that plants do not store more or less coal based on nearby economic conditions. We argue in Section 2 that this identifying assumption is likely to hold because power plants purchase and store coal based primarily on electricity demand, spot and contract coal prices, as well as natural gas prices; these factors vary at the regional rather than local level. For example, the electricity produced by plants is injected into the transmission grid; this electricity is distributed throughout the grid as needed. Due to this near-costless transmission of electricity across long distances, plants likely purchase and store coal based on regional/grid-level electricity demand rather than local electricity demand.

Also, almost any source of variation in $PM_{2.5}$ concentration levels affects the concentration levels of other pollutants as well.[21] For example, burning fossil fuels emits SO_2, NO_2 and CO in addition to $PM_{2.5}$. Environmental regulations targeting $PM_{2.5}$ are likely to have "co-benefits": reductions in other pollutants when firms take action to comply with the environmental regulations associated with $PM_{2.5}$. However, we show in Section 4 that coal storage and handling affects $PM_{2.5}$ concentration levels but not the concentration levels of other pollutants (SO_2, NO_2 and CO). By using changes in $PM_{2.5}$ due to coal storage and handling, we have the unique opportunity to isolate how a single pollutant ($PM_{2.5}$) affects mortality rates.

[21] See Greenstone and Hanna (2014) or Benmarhnia, Bharadwaj and Romero (2017) for further discussion of the difficulties associated with isolating changes in a single pollutant in order to assess the impact of that single pollutant on mortality rates.

5.1 Data Sources: Constructing Mortality Rates

We collect monthly, county-level total number of deaths by age group and cause of death from the Centers for Disease Control and Prevention (CDC); mortality rates are calculated from these data using annual, county-level population data by age group from the Survey of Epidemiology and End Results (SEER). Our analysis considers how $PM_{2.5}$ affects four different types of mortality rates associated with: (1) deaths related to the cardiovascular system, (2) deaths related to the respiratory system, (3) deaths due to external causes such as accidents, and (4) deaths due to any cause.

The United States Environmental Protection Agency (USEPA) uses the results from two epidemiological studies, Krewski et al. (2009) and Lepeule et al. (2012), in order assess the impact of $PM_{2.5}$ on mortality rates. In order to remain consistent with these two studies, our primary specifications focus on people who are at least 30 years old ("adults"). The empirical results considering people of all ages, provided in Appendix A, are very similar to those presented below for the adult sub-population. Finally, we also present specifications considering overall infant (children ages 0-4) mortality rates. Our findings indicate that the increase in mortality rates due to a given level of $PM_{2.5}$ exposure is higher for infants relative to adults; infants are especially vulnerable to the adverse health consequences associated with $PM_{2.5}$ exposure.

5.2 Empirical Framework

We estimate the following specification relating mortality rates to local $PM_{2.5}$ concentration levels:

$$log(\frac{Deaths_{c,t}}{Pop_{c,y}}) = \alpha_{c,y} + log(PM2.5_{i,t})\gamma + X_{i,p,t}\beta + \epsilon_{i,p,t} \tag{2}$$

where c indexes the county where air quality (AQ) monitor i is located, p indexes a coal-fired power plant linked to AQ monitor i, t indexes the month-of-sample, and y indexes the year-of-sample. This specification includes the same set of controls $X_{i,p,t}$ as described in Section 3 as well as county-of-AQ-monitor/year fixed effects ($\alpha_{c,y}$). We estimate this equation separately for each type of mortality rate: cardiovascular, respiratory, external, all-cause adult, and all-cause infant.

We first present the results from estimating Equation 2 using ordinary least squares (OLS). However, we also consider specifications where we instrument for the log of

$PM_{2.5}$ concentration levels ($log(PM_{2.5})$) in Equation 2 with the log of coal stockpiles ($log(CS_{p,t})$). For both OLS and IV specifications, our standard errors are clustered by air quality monitor and we weight by annual, county-level population.[22] Finally, we focus on plants within 25 miles of their corresponding air quality monitor because we find that there's a statistical link between coal stockpiles and $PM_{2.5}$ for the less than 25 miles distance bandwidth but not the 25-50 mile distance bandwidth in Section 4.

5.3 Empirical Results

The top panel of Table 5 lists the empirical results when regressing $PM_{2.5}$ concentration levels on mortality rates using the OLS framework described in the previous subsection. We run separate regressions for mortality rates associated with the cardiovascular system, the respiratory system, all-cause adult deaths, and external causes such as accidents. The summary statistics for the regressions in this subsection are relegated to Appendix Section A. The OLS coefficient estimates in Column 1 (Column 2) of Table 5 relating $PM_{2.5}$ and adult cardiovascular (respiratory) mortality rates are negative and imprecisely estimated; moreover, the coefficient estimate on $log(PM_{2.5})$ is negative and statistically significant when considering overall, all-cause mortality rates (Column 3). Taken at face value, the OLS estimates suggest that increases in locally monitored $PM_{2.5}$ levels *decrease* overall adult mortality rates. Finally, deaths due to external causes such as accidents should not be related to $PM_{2.5}$ concentration levels; however, the OLS results in Column 4 of Table 5 indicate a negative and statistically significant association between the external mortality rate and $PM_{2.5}$. Summarizing, it is clear that the statistical association between $PM_{2.5}$ and mortality rates estimated using ordinary least squares is being confounded by other factors, such as economic growth or urbanization, that are correlated with both $PM_{2.5}$ and mortality rates.

[Table 5 about here.]

In order to estimate the causal impact of $PM_{2.5}$ on mortality rates, we instrument for $PM_{2.5}$ with the level of coal stockpiles held by power plants in each month-of-sample. The empirical results relating $PM_{2.5}$ concentration levels to mortality rates using this

[22] Our empirical results are similar if we do not weight by population.

instrumental variables (IV) framework are presented in the bottom panel of Table 5. In contrast to the OLS results, we find a positive and statistically significant link between $PM_{2.5}$ concentration levels and both cardiovascular and respiratory mortality rates within our IV framework; our IV findings indicate that a 10% increase in $PM_{2.5}$ concentration levels results in a 1.6% (2.8%) increase in average cardiovascular (respiratory) mortality rates. In addition, we find a positive and statistically significant link between $PM_{2.5}$ exposure and overall, all-cause mortality rates in Column 3. Finally, when using our IV framework, there is **not** a statistically significant relationship between the external mortality rate and $PM_{2.5}$ exposure (Column 4); this finding is consistent with the intuition that the probability of death due to external causes such as accidents should not change substantially with $PM_{2.5}$ concentration levels. Thus, the lack of a relationship between $PM_{2.5}$ and the external mortality rate provides evidence that we are not capturing some other source of variation that simultaneously increases both $PM_{2.5}$ concentration levels and mortality rates.

[Table 6 about here.]

Finally, we focus on the effect of $PM_{2.5}$ on the mortality rates of infants (children ages 0 to 4) in Table 6. Consistent with the intuition that infants are especially vulnerable to the negative health effects of $PM_{2.5}$, we see from the instrumental variables (IV) results in the bottom panel of Table 6 that the average increase in infant mortality rates from a 1% increase in $PM_{2.5}$ (0.32%) is much larger than the corresponding average increase in adult mortality rates (0.11%). Finally, we see from the top panel of Table 6 that our ordinary least squares (OLS) estimate suggests that there is no statistical link between $PM_{2.5}$ and infant mortality rates. This highlights both the necessity of our instrumental variables identification strategy based on coal storage as well as the strength of the statistical link between coal stockpiles and $PM_{2.5}$ concentration levels.

5.4 Robustness Checks/Sensitivity Analyses

Our empirical results are quantitatively similar to those presented in Table 5 and Table 6 if we instead: (1) consider mortality rates based on people of all ages rather than just people over 30 years old (see Appendix Table A.7) or (2) instrument for $PM_{2.5}$ concentration levels using both an indicator for whether the plant purchased any coal in that month as well as coal stockpiles (see Appendix Table A.8). Also, we draw the same qualitative conclusions as discussed in the previous subsection if we instead estimate semi-elasticities relating $PM_{2.5}$ in levels (read: microgram per cubic meters) to the log of mortality rates; see Appendix Table A.6) for these estimated semi-elasticities. Given that the environmental science literature does not give clear guidance on the "dose-response" function (read: functional form) relating $PM_{2.5}$ and mortality, it is comforting that we find a similar magnitude effect of $PM_{2.5}$ on mortality rates when using the Log-Log specification (read: both $PM_{2.5}$ and mortality rates in logs) versus the Log-Levels specification (read: $PM_{2.5}$ in levels but the log of mortality rates).

5.5 Policy Implications of our Estimates

The estimated relationship between $PM_{2.5}$ and mortality rates is arguably the most important component of any cost-benefit analysis of any environmental policy pertaining to local air pollution. For example, the United States Environmental Protection Agency (USEPA) estimates that roughly 85% of the benefits from the 1990 Clean Air Act Amendments in the year 2020 will come from "reductions in premature mortality associated with reductions in ambient particulate matter".[23] The USEPA typically uses the association between $PM_{2.5}$ levels and overall adult mortality rates reported by Krewski et al. (2009) or Lepeule et al. (2012) in its regulatory impact analyses for $PM_{2.5}$; Krewski et al. (2009) (Lepeule et al. (2012)) reports that a 1 microgram per cubic meter increase in $PM_{2.5}$ is associated with a 0.6% (1.4%) increase in overall adult mortality rates on average. Our estimate of the *semi-elasticity* between $PM_{2.5}$ and adult mortality rates falls squarely in this 0.6%-1.4% range; in particular, our estimate from Appendix Table A.6 indicates that a 1 microgram per cubic meter increase in $PM_{2.5}$ results in a 1.1% increase in overall, adult mortality rates on average. It is comforting that our causal estimate of the semi-elasticity relating $PM_{2.5}$ and mortality rates is similar in magnitude to the epidemiological estimates used by the USEPA in their regulatory impact analyses (RIAs); our findings lend further credence to these RIAs.

[23]This quotation is from the USEPA website: https://www.epa.gov/clean-air-act-overview/benefits-and-costs-clean-air-act-1990-2020-second-prospective-study.

6 Local Environmental Health Costs of Coal Stockpiles

This section combines our estimated impact of coal stockpiles on $PM_{2.5}$ from Section 4 with our estimated effect of $PM_{2.5}$ on mortality rates from Section 5. In particular, we calculate the increased number of deaths due to the $PM_{2.5}$ emissions from coal storage and handling, monetizing the cost of these deaths using the Value of Statistical Life (VSL) approach. We also present two alternative methods for calculating the local air pollution cost of coal storage and handling as sensitivity analyses. First, we use a concentration-response relationship from the epidemiological literature in order to calculate the total number of adult deaths due to our estimated $PM_{2.5}$ increase from coal stockpiles; this alternative approach does not use our estimated effect of $PM_{2.5}$ on mortality rates. We also present our local air pollution cost in life-years lost per ton of coal stockpiled as an alternative to the VSL approach. Summarizing our primary findings, we calculate that a one ton increase in coal stockpiles has local environmental costs of \$196.98. Roughly 86% (14%) of these external costs come from adult (infant) mortality simply because a far larger number of adults are exposed to the $PM_{2.5}$ emissions from coal storage and handling relative to infants.

6.1 Translating Partial Effects to Local Environmental Costs: Methodology

This subsection describes the method we use to quantify the local environmental health cost of the $PM_{2.5}$ emissions from coal storage and handling. We focus exclusively on how this $PM_{2.5}$ exposure affects mortality risk because prior research has shown that the majority of the damage from $PM_{2.5}$ exposure can be attributed to elevated mortality risk (EPA (1999); Muller, Mendelsohn and Nordhaus (2011)). Our quantification of local air pollution costs relies on plant-level average partial effects linking coal stockpiles (in tons) to $PM_{2.5}$ increases (in microgram per cubic meter). For each power plant p, we calculate this partial effect ($\Delta PM2.5_p^{CS}$) by averaging the air quality monitor/plant/month-of-sample level partial effects estimated in Section 4 over air quality monitors and months-of-sample. For the local air pollution costs presented in the next subsection, we compute the partial effects implied by the estimates from our Log-Log specification allowing for different impacts of coal stockpiles on $PM_{2.5}$ for plants upwind versus downwind from their corresponding air quality monitor; see the bottom panel of Table 2 in Section 4 for these estimates.[24] We focus on the local air pollution costs of coal storage and handling faced by populations within 25 miles of power plants because we find no statistical impact of coal stockpiles on the $PM_{2.5}$ concentration levels measured by air quality monitors farther than 25 miles away from the plant. Summarizing, our plant-level average partial effects ($\Delta PM2.5_p^{CS}$) translate a ton of coal stored into $PM_{2.5}$ increases in micrograms per cubic meter.

We also use our estimated effects of $PM_{2.5}$ on average adult and infant mortality rates from Section 5; we denote these impacts ΔMR^{adult} and ΔMR^{infant} respectively. Importantly, though we present elasticities in Section 5, ΔMR^{adult} and ΔMR^{infant} are *semi-elasticities* that convert $PM_{2.5}$ in micrograms per cubic meter into percentage changes in average adult and infant mortality rates; these semi-elasticity estimates are presented in Appendix Table A.6.

The following equation formally describes how we calculate the annual total number of adult deaths associated with increases in $PM_{2.5}$ concentration levels due to coal storage and handling:

$$M_p = \sum_{a=1}^{13} MR_{c,a}(\Delta MR^{adult})(\Delta PM2.5_p^{CS})POP_{c,a} \tag{3}$$

where $MR_{c,a}$ is the mortality rate of age group a in the county c where power plant p is located. To calculate these mortality rates, we first collect data on overall, county-level total number of deaths in the year 2011 from the Centers for Disease Control and Prevention (CDC); the CDC provides these mortality data separately for 13 age groups. We divide county-level total number of deaths for each age group by the county's total population for this same age group in 2011 in order to construct mortality rates $MR_{c,a}$; these population data are provided by the Survey of Epidemiology and End Results (SEER). However, for $POP_{c,a}$, we instead use census block level population data from the 2010 U.S. Census in order to calculate the total number of people in age group a

[24] Our local air pollution cost estimates are very similar to the ones presented in the paper if we instead calculate plant-level partial effects ($\Delta PM2.5_p^{CS}$) using the Levels-Levels specification (read: regressing $PM_{2.5}$ in micrograms per cubic meter on coal stockpiles in tons) that accounts for wind direction; see the top panel of Table 2 for our estimated effect of coal stockpiles on $PM_{2.5}$ using this specification and Appendix Table A.9 for the external costs based on this specification.

living in census blocks located in county c whose centroid is within 25 miles of power plant p.[25]

Multiplying mortality rate by population ($MR_{c,a} \times POP_{c,a}$) gives us the total number of deaths for each age group; we use the partial effects estimated in Section 4 ($\Delta PM2.5_p^{CS}$) as well as the semi-elasticities estimated in Appendix Table A.6 (ΔMR^{adult}) in order to calculate the annual total number of adult deaths (M_p) attributable to the increased $PM_{2.5}$ emissions from coal storage and handling at power plant p.

We similarly calculate the annual total number of infant deaths due to the $PM_{2.5}$ from coal stockpiles:

$$M_p = MR_c(\Delta MR^{infant})(\Delta PM2.5_p^{CS})POP_{c,infant}$$

where $POP_{c,infant}$ is the population of children ages 0 to 4 ("infants") living in census blocks both within 25 miles of power plant p and in the same county c as power plant p. MR_c is the county-level infant mortality rate, calculated using a combination of the CDC data on county-level infant deaths and the SEER data on county-level infant population.

As a sensitivity analysis, we also use the concentration-response relationship estimated in Krewski et al. (2009) in order to translate the effect of coal stockpiles on $PM_{2.5}$ into total number of adult deaths. This approach, specified below, does not use our estimated effect of $PM_{2.5}$ on mortality rates from Section 5:

$$M_p = \sum_{a=1}^{13} POP_{c,a} MR_{c,a} (1 - \frac{1}{exp(\rho \Delta PM2.5_p^{CS})})$$

where ρ is a statistically estimated parameter reported by Krewski et al. (2009). Both population ($POP_{c,a}$) and mortality rates ($MR_{c,a}$) are defined above in the discussion of Equation 3.

The next step is quantifying the cost in dollars of these deaths. We monetize the increased mortality from the $PM_{2.5}$ due to coal stockpiles using the Value of a Statistical Life (VSL) methodology(Viscusi and Aldy, 2003); we consider a VSL of \$9.85 million (expressed in 2013 U.S. Dollars) based on the regulatory impact analyses of air pollution

[25] This census block based population is typically only slightly smaller than county-level population because the median county land area in the United States is 640 square miles (this statistic is from the 2010 Census).

conducted by the USEPA. The monetary marginal damage due to the $PM_{2.5}$ emissions from an additional ton of coal stockpiled at plant p is given by:

$$MD_p = M_p VSL$$

As an alternative approach to the VSL methodology, we also calculate the life-years lost due to the increases in $PM_{2.5}$ concentration levels associated with coal storage and handling. To do this, we employ data on the life expectancy of each age group conditional on having reached that age ("conditional life expectancy") from the National Vital Statistics Reports (U.S. 2011 Life Tables). In particular, life-years lost is calculated as the change in mortality risk for a given age group due to $PM_{2.5}$ increases times the conditional life expectancy of that age group. As an example, if pollution exposure increases mortality risk for a given age group by 5%, then the years of life lost due to this pollution exposure is 0.05 times the conditional life expectancy for that age group. In contrast, the VSL-based approach implicitly values changes in the risk of death for each age group due to $PM_{2.5}$ exposure equally across age groups.

6.2 Translating Partial Effects to Local Environmental Costs: Empirical Results

This subsection presents our estimates of the local environmental health costs associated with the $PM_{2.5}$ emissions from coal storage and handling. In particular, Table 7 shows the local air pollution cost per ton of coal stockpiled at the *median* plant; we present the median external cost rather than the average external cost because the plant-level distribution of these external costs is right-skewed. Row 1 of Table 7 uses the link between $PM_{2.5}$ and adult mortality rates estimated in Krewski et al. (2009), Row 2 of this table uses our estimated effect of $PM_{2.5}$ on adult mortality rates from Section 5, and Row 3 uses our estimated effect of $PM_{2.5}$ on infant (ages 0-4) mortality rates from Section 5.

We see from Row 2 of Table 7 that the external cost of the increased adult mortality due to the $PM_{2.5}$ from coal storage and handling is \$168.48 per ton of coal stockpiled. The additional external cost of coal stockpiles due to increased infant mortality is \$28.50 per ton (see Row 3 of Table 7). Despite the fact that a 1% increase in $PM_{2.5}$ results in a larger *percentage* increase in infant mortality rates relative to adult mortality rates, the local environmental cost of coal stockpiles is smaller when considering infant mortality rather than adult mortality simply because the population of adults is substantially larger than the population of infants.

The United States Environmental Protection Agency (USEPA) uses the semi-elasticity estimates between $PM_{2.5}$ (in micrograms per cubic meter) and adult mortality rates reported by either Krewski et al. (2009) or Lepeule et al. (2012) when assessing the costs versus benefits of any policy or action related to local air pollution. Comparing Row 1 versus Row 2 of Table 7, our external cost estimates for adults are substantially smaller when we use the $PM_{2.5}$/adult mortality rate semi-elasticity estimated by Krewski et al. (2009) rather than our own estimate. This makes sense given that we estimate a larger effect of $PM_{2.5}$ on adult mortality rates than Krewski et al. (2009). However, as discussed in Section 5, our $PM_{2.5}$/adult mortality rate semi-elasticity estimate is smaller than the one found in Lepeule et al. (2012), indicating that our estimate is broadly consistent with the magnitudes reported epidemiological literature. Summarizing, our results highlight the importance of the relationship between $PM_{2.5}$ and adult mortality rates in quantifying the local environmental costs of air pollution; even small differences in the $PM_{2.5}$/mortality rate semi-elasticity used to calculate local air pollution costs result in large differences in these external costs.

The local environmental damages per ton of coal stockpiled are quite large given that the average plant pays roughly $48 per ton for coal and stockpiles 212,781.6 tons of coal on-site. However, our estimated local environmental cost of coal storage and handling are not unreasonably large when compared to the local air pollution cost of burning coal. In particular, Column 2 of Table 7 presents our external costs on a per-MWh basis rather than a per-ton basis; we convert tons of coal burned to MWh of electricity generated using monthly, plant-level data on coal consumption (in tons) and electricity generation (in MWh). Our estimates of the external cost of coal stockpiles summing across adults and infants is $106.46 per MWh-equivalent. As a basis for comparison, Levy, Baxter and Schwartz (2009) reports a median external cost (across roughly 400 plants) of $140 per MWh associated with the $PM_{2.5}$ emissions from burning coal; this $140 per MWh translates to roughly $230 per MWh when we adjust for the difference in the value of statistical life (VSL) used in our paper versus Levy, Baxter and Schwartz (2009). Burning coal emits significantly more $PM_{2.5}$ than storing and handling coal; it is thus comforting that our estimates of the local air pollution costs associated with coal storage and handling (roughly \$106 per MWh-equivalent) are substantially smaller than the external cost of the $PM_{2.5}$ from burning coal (roughly \$230 per MWh).

[Table 7 about here.]

The last column of Table 7 presents our local environmental cost estimates in life-years per 100,000 tons. Focusing on Row 2 of this table, we estimate that a 100,000 ton increase in coal stockpiles results in 26.3 life-years lost for adults.[26] As a point of comparison, our VSL-based external cost estimate indicates that a 100,000 ton increase in coal stockpiles results in 16.85 million dollars in environmental damages (see Column 1 of Table 7); this 16.85 million dollars in damages is calculated by monetizing roughly 1.71 adult deaths using a value of statistical life of 9.85 million dollars.[27] Combined, our estimates imply that an adult death due to the $PM_{2.5}$ from coal storage corresponds to roughly 15.38 life-years lost on average.[28] Thus, the reader can pick either a different value of statistical life (VSL) or a different conversion rate between life-years lost and adult deaths in order to scale up or scale down our local environmental cost estimates.

6.3 Heterogeneity in the Local Environmental Costs of Coal Storage

Figure 1 displays the geographic dispersion across the United States of our plant-specific estimates of the local environmental damages from the increased $PM_{2.5}$ emissions associated with a one ton increase in coal stockpiles. This figure highlights that the environmental costs of coal storage and handling are highly local; most counties do not have a power plant located nearby, and thus do not incur any of the local air pollution costs associated with the coal storage and handling at these plants. We also see from Figure

[26] We present our external cost estimates in life-years per GWh in Appendix Table A.10. This allows readers to compare our local air pollution costs of coal storage and handling to estimates from previous literature of the external costs associated with burning coal.

[27] We calculate that a 100,000 ton increase in coal stockpiles results in 1.71 adult deaths by taking our local air pollution cost in dollars per ton of coal stockpiled (\$168.48), multiplying by 100,000 tons, and finally dividing by our value of statistical life (9.85 million dollars per statistical life lost); this gives us 1.71 ($= \frac{168.48 \times 100{,}000}{9{,}850{,}000}$) adult deaths.

[28] We simply divide our estimate of the life-years lost from an additional 100,000 tons of coal stockpiled (26.3) by our estimate of the total number of adult deaths due to this 100,000 ton increase in coal stockpiles (1.71 adult deaths); this gives us 15.38 ($= \frac{26.3}{1.71}$) life-years lost per adult death.

1 that the local air pollution costs per ton of coal stockpiled are larger in more densely populated areas; for example, the external costs per ton of coal stockpiled are particularly high in the Northeastern region of the United States. Figure 2 demonstrates directly that local air pollution cost per ton of coal stockpiled increases with total county-level adult population. This figure also tells us that the distribution of local environmental health costs are right-skewed; for example, we estimate that air pollution costs per ton of coal stockpiled are extremely large in Cook County (noting that the city of Chicago is in Cook County).

[Figure 1 about here.]

[Figure 2 about here.]

[Table 8 about here.]

Finally, the average census tract with a coal-fired power plant within its boundaries has a lower median family income, a lower per-capita income, and a lower proportion of residents with a bachelor's degree relative to the average census tract without a coal-fired power plant; these tract-level averages, calculated using data from the 2000 U.S. Census, are displayed in Table 8. This suggests that the highly localized environmental impacts of coal storage and handling are borne disproportionately by economically disadvantaged communities.[29]

7 Conclusion and Policy Implications

Burning coal emits global pollutants which contribute to climate change as well as criteria air pollutants which affect the health of local populations. The trains, trucks, and barges transporting coal from mine to plant burn diesel or gasoline which also results in the emission of global and local pollutants. Finally, mining coal can cause acid mine drainage and the ecological impacts of mountaintop removal are significant. Seemingly every aspect of the coal supply chain is known to have environmental costs. This paper

[29] Our empirical evidence is corroborated by Davis (2011); Davis (2011) uses restricted-access census microdata from 1990 and 2000 to show that neighborhoods near fossil-fuel fired power plants have lower average household incomes and educational attainment.

demonstrates that storing and handling coal also emits local pollutants, arguing that environmental policy targeted specifically at the storage and handling of coal is especially important given the current and proposed policies that subsidize coal stockpiles held on-site at U.S. power plants. This section first summarizes our paper's primary empirical findings; we conclude by discussing the policy implications of these findings.

7.1 Summary of our Findings

Our paper proceeded in three steps. We first quantified the effect of coal stockpiles on local ambient air pollution. Next, this variation in local air pollution from coal stockpiles was used to identify the link between $PM_{2.5}$ and mortality rates. Finally, we combined our estimates of the effect of coal storage and handling on $PM_{2.5}$ and the effect of $PM_{2.5}$ on mortality rates, using the methodology employed by the USEPA in their regulatory impact analyses in order to calculate the per-ton external costs associated with storing and handling coal at U.S. power plants.

Summarizing our primary results, we find that a 10% increase in coal stockpiles results in a 0.09% increase in average $PM_{2.5}$ concentration levels within 25 miles of power plants. These $PM_{2.5}$ concentration increases are more severe for: 1) local populations downwind from coal-fired power plants, and 2) areas and months with less precipitation. We also show that there is **no** statistical effect of coal stockpiles on locally-monitored carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2), providing strong evidence that our estimated increase in $PM_{2.5}$ from coal storage and handling is not driven by combustion-based sources of $PM_{2.5}$ such as the coal burned by the power plant or the fuel burned by the trains, trucks, and barges delivering coal to the plant.

We next estimate the effect of $PM_{2.5}$ on adult (people ages 30 and older) and infant (children ages 0-4) mortality rates within an instrumental variables framework. Instrumenting $PM_{2.5}$ with monthly, plant-level coal stockpiles, we find a positive and statistically significant effect of $PM_{2.5}$ exposure on both adult and infant mortality rates; our estimated effect of $PM_{2.5}$ exposure on mortality rates is especially large for infants. As a placebo analysis, we show that there's no statistical effect of $PM_{2.5}$ on external mortality rates, which is comforting given that the number of deaths due to external causes such as accidents should not be affected by local $PM_{2.5}$ concentration levels. Our estimate of the semi-elasticity relating $PM_{2.5}$ in micrograms per cubic meter to overall adult mortality rates in percentage terms is roughly in line with the estimates used by policymakers such as the USEPA in order to assess the environmental costs of $PM_{2.5}$ exposure. Thus, we provide evidence that these analyses are not significantly over-stating or under-stating the benefits of policies aimed at reducing $PM_{2.5}$ concentration levels.

Finally, we combine our estimates of the average increase in $PM_{2.5}$ from coal stockpiles with the average increase in mortality rates from $PM_{2.5}$ in order to quantify the local air pollution costs of coal storage and handling. In particular, we calculate the total number of adult and infant deaths due to the $PM_{2.5}$ from coal stockpiles and use the Value of Statistical Life approach to monetize this increased mortality risk. Our results indicate that the local air pollution costs of coal storage and handling are roughly \$197 per ton of coal stockpiled. These local environmental costs are sizable given that U.S. power plants pay roughly \$48 per ton of coal on average and store an average of 212,781.6 tons of coal on-site during our 2002-2012 sample period. That being said, our external cost estimates are not unreasonably large given that the local damages from storing and handling coal (\$106 per MWh-equivalent of coal stockpiled) are substantially smaller than the local environmental cost associated with the $PM_{2.5}$ emitted when burning coal (roughly \$230 per MWh from Levy, Baxter and Schwartz (2009)).

7.2 Policy Implications

The economic costs of simple $PM_{2.5}$ mitigation strategies, such as covering coal stockpiles or the rail cars containing coal, are almost certainly small when compared to the environmental costs incurred by the economically disadvantaged communities living near coal-fired power plants and railroad tracks.[30] Moreover, a policy mandating that coal piles be covered does not require significant coordination across jurisdictions because the environmental impacts of coal storage and handling are highly local; the vast majority of the local environmental costs from a plant's coal storage and handling behavior are incurred within 25 miles of this plant. Thus, local environmental policies designed to mitigate the $PM_{2.5}$ emissions from coal storage and handling are likely easier to enact and implement relative to policies designed to mitigate global or regional pollutants such

[30] One potential alternative strategy to mitigate the $PM_{2.5}$ emissions associated with coal stockpiles is to frequently spray down these piles with water.

as CO_2 or SO_2.

The local environmental health costs of coal storage and handling apply more broadly than our examination of U.S. coal-fired power plants. For example, coal is also stored and handled at coal mines, rail yards, and coal export terminals; people living these mines, rail yards, and export terminals likely face a similar increase in $PM_{2.5}$ exposure due to the fugitive emissions from stationary coal piles, the volatile gases from these piles, and the on-site coal handling process. To demonstrate this, we combine our estimated effect of coal stockpiles on $PM_{2.5}$ with data from SNL Financial on the location of coal export terminals and the monthly quantity of coal stored at these terminals. Using the methodology described in Section 6, we estimate local environmental costs of roughly \$500 per ton of coal stored at a terminal. This value is considerably larger than the local air pollution costs of coal stored at power plants because export terminals are located in or near large cities while power plants tend to be in rural areas. The data report 36 terminals in the coterminous United States, each of which holds an average of 29,900 tons of coal. Based on this, our back-of-the-envelope estimate of the aggregate annual damage across all export terminals due to the $PM_{2.5}$ from coal storage and handling is roughly \$538 million. This exercise demonstrates that the adverse impacts on community health from coal export terminals are likely to be significant.

Finally, both current and proposed economic policies provide subsidies to U.S. power plants based on the level of coal stockpiles they hold on-site. For example, 74% of U.S. coal-fired electricity generation was produced under output price regulation in 2016; plants under output price regulation typically receive a "working capital allowance" from the regulator based on the level of coal stockpiles held on-site. Consistent with the intuition that this working capital allowance incentivizes plants under output price regulation to store more coal, Jha (2017*a*) shows using a matched difference-in-differences framework that regulated plants hold 13% more coal stockpiles on average relative to similar plants facing electricity market mechanisms. Our paper demonstrates that these regulatory distortions to coal storage behavior have environmental costs in addition to the economic costs documented in Jha (2017*a*). Namely, we estimate that a one ton increase in coal stockpiles has an air pollution cost of roughly \$197. For our 2002-2012 sample of plants within 25 miles of an air quality monitor, regulated plants on average hold roughly 262,000 tons of coal; there are 193 regulated plants in our data. Thus, the annual air pollution cost of the 13% increase in coal stockpiles due to the structure of output price regulation across all regulated plants is approximately $1.3 billion ($= 197 \times 193 \times 0.13 \times 261{,}999.8$). Jaramillo and Muller (2016) estimates that electric power generation in the United States (including both coal and natural gas plants) produced damages of roughly $230 billion in 2002 and $170 billion in 2011 (both magnitudes are in 2014 dollars). Hence, our aggregate estimates of the local air pollution cost of increases in coal storage and handling at power generation facilities due to the structure of output price regulation range from 0.57% to 0.76% of the aggregate damages from combustion at these power plants.

Plants facing electricity market mechanisms do not currently receive a subsidy for storing more coal. However, on September 9^{th}, 2017, the United States Department of Energy demanded that the Federal Energy Regulatory Commission provide subsidies to power plants holding at least 90 days of fuel on site. This proposed Grid Resiliency Pricing Rule (GRPR) would apply only to plants facing electricity market mechanisms; we thus calculate the local air pollution costs of this rule for the subset of coal-fired power plants that sold electricity in either the PJM (read: mid-Atlantic states) or ERCOT (read: Texas) wholesale markets in 2016. For this GRPR policy counterfactual, we increase a plant's end-of-month coal stockpiles to 90 days worth of coal if and only if this plant held less than 90 days worth of coal in that month; in order to translate coal stockpiles from tons to "days of burn", we approximate each plant's expected daily coal consumption in each month-of-sample using the rolling mean of the plant's average daily coal consumption (in tons) taken over the 36 previous months.[31] Put another way, we simply divide each plant's end-of-month coal stockpiles (in tons) by their expected daily coal consumption (in tons per day) in order to calculate the number of days of coal this plant has stored on-site. For the year 2014 (the latest year for which we have data), the average plant in the average month-of-sample stockpiles 444,208 more tons of coal under the GRPR policy relative to their observed coal stockpile levels. Aggregated across the 69 plants selling into PJM or ERCOT in 2016, we find that the local air pollution due to the increased coal storage and handling associated with the GRPR results in roughly 13 premature deaths. Using a value of statistical life of 9.85 million dollars (expressed in 2013 U.S. dollars), the local air pollution costs of the GRPR are roughly $128 million

[31] The Energy Information Administration (EIA) approximates each plant's expected daily consumption in each month-of-sample in the same way; for example, see https://www.eia.gov/todayinenergy/detail.php?id=18711.

dollars. Summarizing, the environmental costs of coal storage and handling must be considered when evaluating the benefits versus costs of economic policies such as the GRPR that subsidize on-site coal storage.

References

Anderson, Michael L. 2015. "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality."

Benmarhnia, Tarik, Prashant Bharadwaj, and Mauricio Romero. 2017. "Using Instrumental Variables under Partial Observability of Endogenous Variables for Assessing Effects of Air Pollution on Health."

Chay, Kenneth Y, and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics*, 1121–1167.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "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*, 110(32): 12936–12941.

Clay, Karen, Akshaya Jha, Nicholas Muller, and Randall Walsh. 2017. "The External Costs of Transporting Petroleum Products by Pipelines and Rail: Evidence From Shipments of Crude Oil from North Dakota." National Bureau of Economic Research.

Clay, Karen, Joshua Lewis, and Edson Severnini. 2015. "Canary in a Coal Mine: Impact of Mid-20th Century Air Pollution Induced by Coal-Fired Power Generation on Infant Mortality and Property Values." Working paper.

Currie, Janet, Matthew Neidell, and Johannes F Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of health economics*, 28(3): 688–703.

Currie, Janet, Matthew Neidell, et al. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *The Quarterly Journal of Economics*, 120(3): 1003–1030.

Davis, Lucas W. 2011. "The effect of power plants on local housing values and rents." *Review of Economics and Statistics*, 93(4): 1391–1402.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif. 2016. "The Effect of Pollution on Health and Health Care Utilization: Evidence from Changes in Wind Direction."

Dominici, Francesca, Michael Greenstone, and Cass R Sunstein. 2014. "Particulate matter matters." *Science*, 344(6181): 257–259.

EPA. 2010. "The Benefits and Costs of the Clean Air Act 1990 to 2020: EPA Report to Congress." United States Environmental Protection Agency, Office of Air and Radiation: Office of Policy in Washington, DC.

EPA, S. 1999. "The Benefits and Costs of the Clean Air Act: 1990 to 2010." EPA-410-R99-001. Washington, DC: US Environmental Protection Agency, Office of Air and Radiation.

Fuller, Thomas. 2016. "Oakland Votes to Block Large Shipments of Coal." *The New York Times*.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental regulations, air and water pollution, and infant mortality in India." *The American Economic Review*, 104(10): 3038–3072.

Herrnstadt, Evan, and Erich Muehlegger. 2015. "Air Pollution and Criminal Activity: Evidence from Chicago Microdata." National Bureau of Economic Research.

Jaffe, Daniel, Justin Putz, Greg Hof, Gordon Hof, Jonathan Hee, Dee Ann Lommers-Johnson, Francisco Gabela, Juliane L Fry, Benjamin Ayres, Makoto Kelp, et al. 2015. "Diesel particulate matter and coal dust from trains in the Columbia River Gorge, Washington State, USA." *Atmospheric Pollution Research*, 6(6): 946–952.

Jaramillo, Paulina, and Nicholas Z Muller. 2016. "Air pollution emissions and damages from energy production in the US: 2002–2011." *Energy Policy*, 90: 202–211.

Jha, Akshaya. 2017*a*. "Dynamic Regulatory Distortions: Coal Procurement at US Power Plants." Working Paper.

Jha, Akshaya. 2017*b*. "Regulatory Induced Risk Aversion: Coal Procurement at US Power Plants." mimeo.

Joskow, P.L. 1988. "Price Adjustments in Long-term Contracts: The Case of Coal." *JL & Econ.*, 31: 47.

Joskow, P.L. 1990. "The Performance of Long-Term Contracts: Further Evidence from Coal Markets." *The Rand Journal of Economics*, 251–274.

Krewski, Daniel, Michael Jerrett, Richard T Burnett, Renjun Ma, Edward Hughes, Yuanli Shi, Michelle C Turner, C Arden Pope III, George Thurston, Eugenia E Calle, et al. 2009. *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality.* Health Effects Institute Boston, MA.

Lepeule, Johanna, Francine Laden, Douglas Dockery, and Joel Schwartz. 2012. "Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009." *Environmental health perspectives*, 120(7): 965.

Levy, Jonathan I, Lisa K Baxter, and Joel Schwartz. 2009. "Uncertainty and variability in health-related damages from coal-fired power plants in the United States." *Risk Analysis*, 29(7): 1000–1014.

Muller, Nicholas Z. 2014. "Boosting GDP growth by accounting for the environment." *Science*, 345(6199): 873–874.

Muller, Nicholas Z, Robert Mendelsohn, and William Nordhaus. 2011. "Environmental accounting for pollution in the United States economy." *The American Economic Review*, 101(5): 1649–1675.

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

NRC and NAS. 2010. "Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use." National Research Council (US). Committee on Health, Environmental, and Other External Costs and Benefits of Energy Production and Consumption. National Academies Press.

Raman, Bipul. 2012. "Coal handling plant in a thermal power generating station." *Electrical Engineering Portal*.

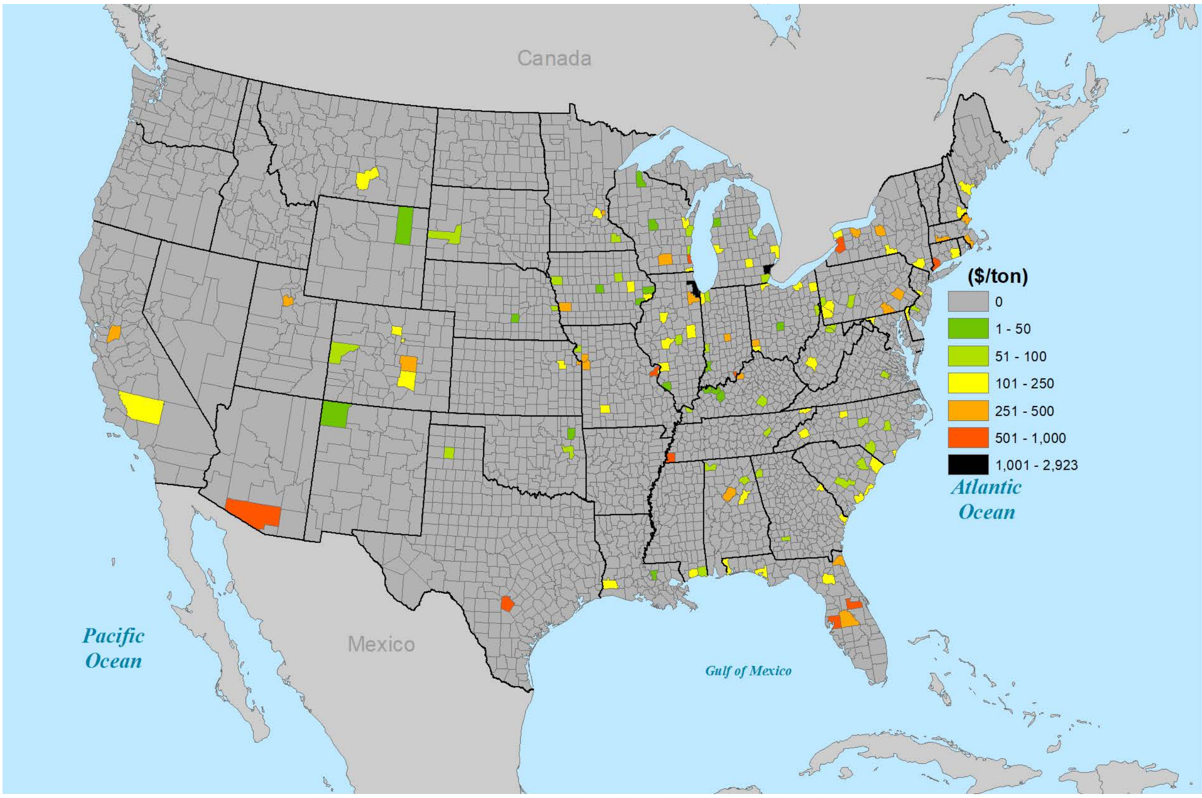
Viscusi, W Kip, and Joseph E Aldy. 2003. "The value of a statistical life: a critical review of market estimates throughout the world." *Journal of risk and uncertainty*, 27(1): 5–76.

Zhang, Xing. 2013. "Gaseous Emissions from Coal Stockpiles." *IEA Clean Coal Centre Working Paper CCC/213*.

List of Figures

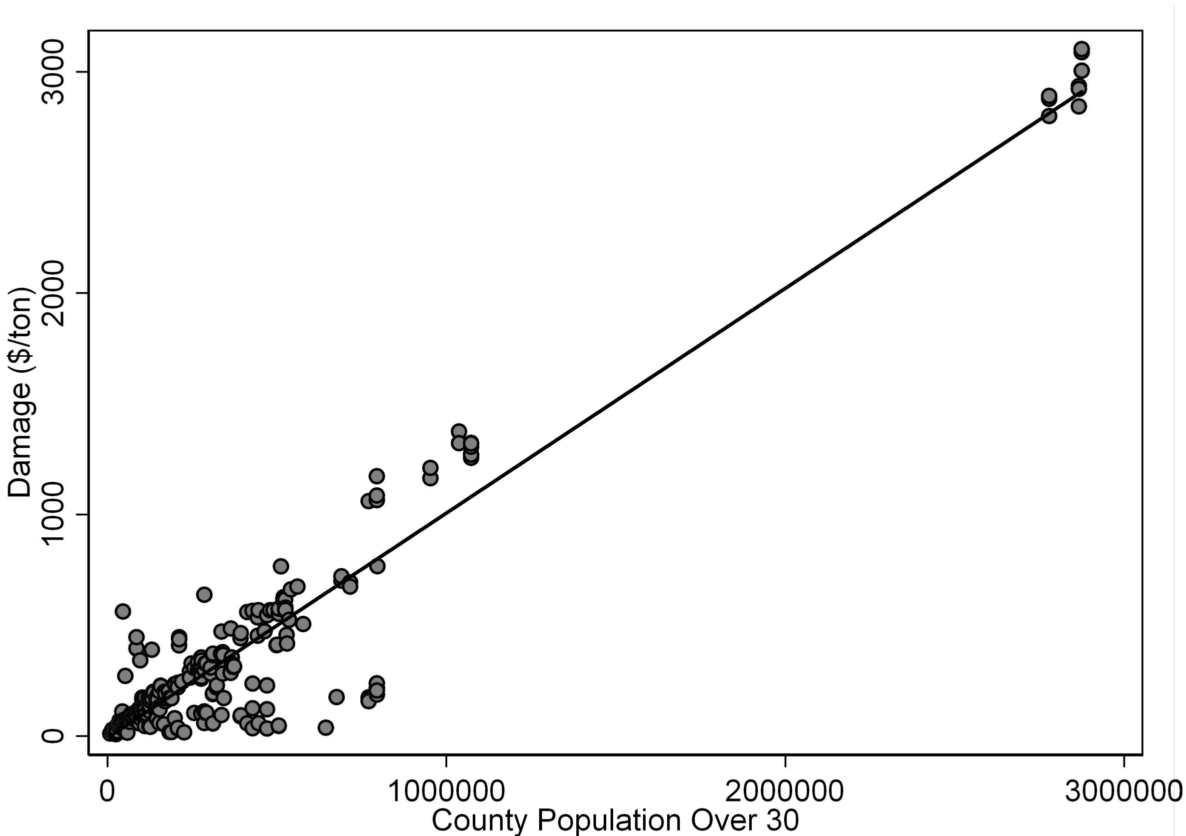
List of Tables

Figure 1: Damages Per Ton of Coal Stockpiled Across the United States



Notes: This figure displays the geographic dispersion across the United States of our plant-specific estimates of the local environmental cost of the increased mortality due to the $PM_{2.5}$ associated with a one ton increase in coal stockpiles. To calculate these costs, we use the partial effect of coal stockpiles on $PM_{2.5}$ implied by the estimates from our Log-Log specification allowing for different effects for plants upwind versus downwind from their matched air quality monitor (see the bottom panel of Table 2 in Section 4); we average over air quality monitors and months-of-sample in order to obtain plant-specific average partial effects. This calculation of costs also uses our own estimated semi-elasticity between $PM_{2.5}$ and overall, all-cause adult mortality rates from Section 5 (see Column 4 of Appendix Table A.6). We use a value of statistical life of 9.85 million dollars (in 2013 U.S. dollars) to monetize the increased mortality from the $PM_{2.5}$ due to coal storage and handling.

Figure 2: Damages Per Ton Stockpiled and Population



Notes: This scatterplot shows the relationship between county-level adult (read: people ages 30 and older) population and our plant-specific estimates of the local environmental cost of the increased mortality due to the $PM_{2.5}$ from a one ton increase in coal stockpiles. The adult population of the county where each coal-fired power plant is located comes from the 2010 U.S. Census. To calculate these costs, we use the partial effect of coal stockpiles on $PM_{2.5}$ implied by the estimates from our Log-Log specification allowing for different effects for plants upwind versus downwind from their matched air quality monitor (see the bottom panel of Table 2 in Section 4); we average over air quality monitors and months-of-sample in order to obtain plant-specific average partial effects. This calculation of costs also uses our own estimated semi-elasticity between $PM_{2.5}$ and overall, all-cause adult mortality rates from Section 5 (see Column 4 of Appendix Table A.6). We use a value of statistical life of 9.85 million dollars (in 2013 U.S. dollars) to monetize the increased mortality from the $PM_{2.5}$ due to coal storage and handling.

Table 1: Overall Effect of Coal Stockpiles on $PM_{2.5}$

Dependent Variable: PM2.5$_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t}$	1.00e-06**	-4.55e-08
	(3.93e-07)	(1.00e-07)
Number of Obs.	50,484	94,400
R-squared	0.691	0.717
Dependent Variable: log(PM2.5$_{i,t}$)		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t})$	0.0090*	0.0006
	(0.0054)	(0.0018)
Number of Obs.	49,171	92,090
R-squared	0.712	0.735
Average PM2.5$_{i,t}$ (in $\frac{ug}{m^3}$)	11.963	12.372
Std. Dev. of PM2.5$_{i,t}$ (in $\frac{ug}{m^3}$)	3.929	4.146
Average $CS_{p,t}$ (in tons)	228,391	331,854
Std. Dev. of $CS_{p,t}$ (in tons)	273,904	409,902

Notes: This table presents the regression results pertaining to the effect of coal stockpiles on $PM_{2.5}$ concentration levels. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter); the bottom panel of this table regresses the log of coal stockpiles on $PM_{2.5}$ in logs. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor in Column 1 and we restrict our sample to plants between 25-50 miles of their AQ monitor in Column 2. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, each plant's monthly total electricity generation, and the sum of the coal stockpiles held by other plants within either 25 miles (for Column 1) or 50 miles (for Column 2) from the air quality monitor. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

Table 2: Effect of Coal Stockpiles on $PM_{2.5}$: Upwind vs. Downwind

Dependent Variable: PM2.5$_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t} \times$		
Monitor Downwind from Plant	1.07e-06**	-8.27e-08
	(4.14e-07)	(1.34e-07)
Monitor Upwind from Plant	8.98e-07**	-2.86e-08
	(4.03e-07)	(1.12e-07)
Number of Obs.	50,484	94,400
R-squared	0.691	0.717
Dependent Variable: log(PM2.5$_{i,t}$)		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t}) \times$		
Monitor Downwind from Plant	0.0122*	-0.0025
	(0.0064)	(0.0026)
Monitor Upwind from Plant	0.0041	0.0040
	(0.0052)	(0.0029)
Number of Obs.	49,171	92,090
R-squared	0.712	0.736

Notes: This table presents the regression results regarding how the link between coal stockpiles and $PM_{2.5}$ concentration levels varies with wind direction. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter); the bottom panel of this table regresses the log of coal stockpiles on $PM_{2.5}$ in logs. In both panels, coal stockpiles is interacted with the relative bearing between power plant and air quality monitor. A relative bearing of 0° means that the wind is blowing directly from the power plant to the AQ monitor, while a relative bearing of 180° means that the wind is blowing directly from the AQ monitor to the power plant. For each plant/air quality (AQ) monitor pair in each month-of-sample, we code the AQ monitor as "downwind" from the plant if their relative bearing is less than 90° and code the AQ monitor as "upwind" from the plant if their relative bearing is greater than 90°. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor in Column 1 and we restrict our sample to plants between 25-50 miles of their AQ monitor in Column 2. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, each plant's monthly total electricity generation, and an indicator for whether the plant is upwind versus downwind from the AQ monitor in that month. We also control for the sum of the coal stockpiles held by other plants within either 25 miles (for Column 1) or 50 miles (for Column 2) from the air quality monitor; for this specification, we calculate this sum of coal stockpiles separately for plants upwind versus downwind from the AQ monitor, including both upwind and downwind sums as controls. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

Table 3: Effect of Coal Stockpiles on $PM_{2.5}$ Concentration: Interacted with 1(Precipitation>0)

Dependent Variable: $\text{PM2.5}_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t}\times$		
Constant	2.72e-06***	2.13e-07
	(9.58e-07)	(1.89e-07)
$1(\text{Precipitation}_{p,t} > 0)$	-1.87e-06**	-2.75e-07
	(8.55e-07)	(1.83e-07)
Number of Obs.	50,484	94,400
R-squared	0.691	0.717
Dependent Variable: $log(\text{PM2.5}_{i,t})$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t})\times$		
Constant	0.0444*	0.0251***
	(0.0229)	(0.0092)
$1(\text{Precipitation}_{p,t} > 0)$	-0.0365*	-0.0248***
	(0.0211)	(0.0094)
Number of Obs.	49,171	92,090
R-squared	0.712	0.736

Notes: This table presents the regression results regarding how the link between coal stockpiles and $PM_{2.5}$ concentration levels varies with precipitation. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter); the bottom panel of this table regresses the log of coal stockpiles on $PM_{2.5}$ in logs. In both panels, coal stockpiles are interacted with an indicator that takes on the value zero if monthly total precipitation is equal to zero and the value one if monthly total precipitation is strictly greater than zero; monthly total precipitation is measured by the set of meteorological monitors within M miles of the plant's corresponding air quality (AQ) monitor. We consider $M = 25$ for the within 25 mile bandwidth specifications and $M = 50$ for the 25-50 mile bandwidth specifications. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor in Column 1 and we restrict our sample to plants between 25-50 miles of their AQ monitor in Column 2. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, each plant's monthly total electricity generation, and the sum of the coal stockpiles held by other plants within either 25 miles (for Column 1) or 50 miles (for Column 2) from the air quality monitor. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

Table 4: Effect of Coal Stockpiles on CO, SO_2, and NO_2 Concentrations

Dep. Var.	$\text{CO}_{i,t}$	$\text{SO2}_{i,t}$	$\text{NO2}_{i,t}$
$CS_{p,t}$	-2.22e-08	9.81e-08	2.39e-07
	(2.50e-08)	(2.00e-07)	(4.14e-07)
Number of Obs.	16,331	32,764	16,274
R-squared	0.764	0.717	0.919
Dep. Var.	$log(\text{CO}_{i,t})$	$log(\text{SO2}_{i,t})$	$log(\text{NO2}_{i,t})$
$log(CS_{p,t})$	0.0040	-0.0023	0.0045
	(0.0147)	(0.0173)	(0.0075)
Number of Obs.	15,859	31,637	15,565
R-squared	0.690	0.735	0.922
Mean of Pollutant in Levels	0.385	3.262	12.889
Std. Dev. of Pollutant in Levels	0.215	2.652	6.664

Notes: This table presents the regression results regarding the link between coal stockpiles and carbon monoxide (CO) concentration levels (measured in parts per million), sulfur dioxide (SO_2) concentration levels (measured in parts per billion), and nitrogen dioxide (NO_2) concentration levels (measured in parts per billion). The top panel of this table regresses coal stockpiles (in tons) on CO, SO_2, and NO_2 in levels; the bottom panel of this table regresses the log of coal stockpiles on the concentration of these pollutants in logs. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. For all of the regressions in this table, we restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, each plant's monthly total electricity generation, and the sum of the coal stockpiles held by other plants within 25 miles of the air quality monitor. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

Table 5: Effect of $PM_{2.5}$ on Adult (Ages 30+) Mortality Rates

OLS Specification				
VARIABLES	log(Cardio MR)	log(Resp. MR)	log(Total MR)	log(External MR)
$log(PM_{2.5})$	-0.0009	-0.0030	-0.0058**	-0.0115*
	(0.0030)	(0.0050)	(0.0023)	(0.0060)
Observations	49,988	43,116	51,206	35,342
R^2	0.908	0.711	0.916	0.749
IV Specification				
VARIABLES	log(Cardio MR)	log(Resp. MR)	log(Total MR)	log(External MR)
$log(PM_{2.5})$	0.160***	0.284***	0.109***	-0.0412
	(0.0359)	(0.0708)	(0.0288)	(0.0805)
Observations	47,847	41,218	48,887	33,830
R^2	0.892	0.661	0.901	0.753

Notes: This table presents the ordinary least squares (OLS) and instrumental variables (IV) results regarding the link between $PM_{2.5}$ concentration levels and adult (people ages 30 and older) mortality rates. We run separate OLS and IV regressions for mortality rates associated with the cardiovascular system, the respiratory system, deaths due to any cause, and deaths due to external causes such as accidents. The top panel of this table regresses the log of $PM_{2.5}$ on the log of county/month-of-sample specific mortality rates (read: the OLS results), while the bottom panel of this table uses coal stockpiles as an instrument for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on mortality rates (read: the IV results). A unit of observation for these regressions is a plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% levels respectively. We control for the log of the total quantity of coal received by each plant in each month-of-sample, an indicator for whether the plant purchased any coal in that month, and each plant's monthly total electricity generation. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include county-of-AQ monitor/year-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.5 for the summary statistics associated with these regressions.

Table 6: Effect of $PM_{2.5}$ on Infant (Ages 0-4) Mortality Rates

OLS Specification	
VARIABLES	log(Total Infant MR)
$log(PM_{2.5})$	-0.0126
	(0.0142)
Observations	13,145
R^2	0.568

IV Specification	
VARIABLES	log(Total Infant MR)
$log(PM_{2.5})$	0.315***
	(0.0717)
Observations	12,704
R^2	0.501

Notes: This table presents the OLS and IV results regarding the link between $PM_{2.5}$ concentration levels and overall, all-cause infant (children ages 0-4) mortality rate. The top panel of this table regresses the log of $PM_{2.5}$ on the log of county/month-of-sample specific infant mortality rates (read: the OLS results), while the bottom panel of this table uses coal stockpiles as an instrument for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on infant mortality rates (read: the IV results). A unit of observation for these regressions is a plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% levels respectively. We control for the log of the total quantity of coal received by each plant in each month-of-sample, an indicator for whether the plant purchased any coal in that month, and each plant's monthly total electricity generation. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include county-of-AQ monitor/year-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.5 for the summary statistics associated with these regressions.

Table 7: Local Air Pollution Cost of Coal Stockpiles

	$\frac{Dollars}{Ton}$	$\frac{Dollars}{MWh}$	$\frac{\text{Life-Years}}{\text{100,000 Tons}}$
Adult: Krewski et al. (2009)	99.52	54.63	15.54
Adult: IV Regression	168.48	92.48	26.30
Infant: IV Regression	28.50	13.98	22.77

Notes: This table presents the local environmental costs of coal storage and handling for the *median* plant in our sample. To calculate these costs, we use the partial effect of coal stockpiles on $PM_{2.5}$ implied by the estimates from our Log-Log specification allowing for different effects for plants upwind versus downwind from their matched air quality monitor (see the bottom panel of Table 2 in Section 4); we average over air quality monitors and months-of-sample in order to obtain plant-specific average partial effects. Row 1 of this table uses the estimated semi-elasticity between $PM_{2.5}$ (in micrograms per cubic meter) and adult (ages 30+) mortality rates estimated in Krewski et al. (2009), Row 2 uses our own estimated semi-elasticity between $PM_{2.5}$ and overall, all-cause adult mortality rates from Section 5 (see Column 4 of Appendix Table A.6), and Row 3 uses our estimated semi-elasticity between $PM_{2.5}$ and infant (ages 0-4) mortality rates (see Column 3 of Appendix Table A.6). Column 1 of this table presents local air pollution costs per ton of coal stockpiled. Column 2 presents local air pollution costs per MWh-equivalent of coal stockpiled; we convert tons of coal to MWhs of electricity by taking plant-level total number of tons of coal burned and dividing by plant-level total electricity generated (in MWh). For the first two columns of this table, we use a value of statistical life of 9.85 million dollars (expressed in 2013 U.S. dollars) to monetize the cost of the increased mortality from the $PM_{2.5}$ due to coal storage and handling. The last column presents local air pollution costs in life-years per 100,000 tons of coal stockpiled.

Table 8: Census Tract Summary Statistics: With versus Without a Power Plant

	Tracts with a Plant		Tracts without a Plant	
	Average	Std. Dev.	Average	Std. Dev.
Prop. with a Bachelor's Degree	0.19	0.17	0.24	0.19
Per-Capita Income	18,536	6,380.70	21,070.9	11,530.13
Median Family Income	45,926.54	15,812.66	50,490.47	24,277.27

Notes: This table presents census-tract level summary statistics separately for the 632 tracts with a coal-fired power plant located within their borders versus the remaining 64,874 census tracts that do not have a coal-fired power plant located within their borders. The tract-level proportion of residents with a bachelor's degree, average per-capita income, and average median family income are collected from the 2000 U.S. Census. The geocoded locations (read: longitude and latitude) of U.S. coal-fired power plants are collected from the eGrid database constructed by the United States Environmental Protection Agency (USEPA).

List of Appendix Tables

A Additional Tables and Figures

A.1 Additional Tables and Figures: Coal Stockpiles and $PM_{2.5}$

Table A.1: Summary Statistics: Plants less than 25 Miles from AQ Monitor

Variable	Obs	Mean	Std. Dev.
$PM_{2.5}$ (in $^{ug}/_{m^3}$)	49,171	11.963	3.929
Stockpiles (in tons)	49,171	228,391	273,904
1(Quantity Received > 0)	49,171	0.878	0.327
Quantity Received (in tons)	49,171	121,885	146,406
Distance (in miles)	49,171	14.242	6.725
Relative Angle (in degrees)	49,171	88.217	55.607
Precipitation (in inches)	49,171	3.482	4.600
1(Precipitation¿0)	49,171	0.962	0.191
Dry Bulb Temp. (in degrees Fahrenheit)	49,171	54.355	17.252
Dew Point Temp. (in degrees Fahrenheit)	49,171	43.522	16.582
Wet Bulb Temp. (in degrees Fahrenheit)	49,171	48.944	15.658
Relative Humidity (percentage)	49,171	70.283	7.656
Station Pressure (in hundredths of an inch)	49,171	29.291	0.870
5% Wind Speed (in miles-per-hour)	49,171	0.340	0.852
95% Wind Speed (in miles-per-hour)	49,171	15.535	3.486
Number of Plants per AQ Monitor	479	1.539	0.788

Notes: This table presents the summary statistics for the regressions regarding the effect of coal stockpiles on $PM_{2.5}$ discussed in Section 4. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We weight observations by the inverse of the distance between air quality (AQ) monitor and power plant for these summary statistics. We restrict our sample to plants within 25 miles of their AQ monitor for this table.

Table A.2: Summary Statistics: Plants Between 25-50 Miles from AQ Monitor

Variable	Obs	Mean	Std. Dev.
$PM_{2.5}$ (in $^{ug}/_{m^3}$)	92,090	12.372	4.146
Stockpiles (in tons)	92,090	331,854	409,902
1(Quantity Received > 0)	92,090	0.890	0.313
Quantity Received (in tons)	92,090	179,145	216,899
Distance (in miles)	92,090	38.010	7.264
Relative Angle (in degrees)	92,090	88.524	53.859
Precipitation (in inches)	92,090	3.560	4.703
1(Precipitation¿0)	92,090	0.984	0.126
Dry Bulb Temp. (in degrees Fahrenheit)	92,090	54.577	16.875
Dew Point Temp. (in degrees Fahrenheit)	92,090	43.936	16.389
Wet Bulb Temp. (in degrees Fahrenheit)	92,090	49.261	15.415
Relative Humidity (percentage)	92,090	70.613	6.869
Station Pressure (in hundredths of an inch)	92,090	29.323	0.743
5% Wind Speed (in miles-per-hour)	92,090	0.310	0.700
95% Wind Speed (in miles-per-hour)	92,090	14.905	3.349
Number of Plants per AQ Monitor	606	1.865	1.326

Notes: This table presents the summary statistics for the regressions regarding the effect of coal stockpiles on $PM_{2.5}$ discussed in Section 4. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We weight observations by the inverse of the distance between air quality (AQ) monitor and power plant for these summary statistics. We restrict our sample to plants between 25-50 miles of their AQ monitor for this table.

Table A.3: Effect of Coal Stockpiles on $PM_{2.5}$ Concentration: Interacted with Precipitation

Dependent Variable: $\text{PM2.5}_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t} \times$		
Constant	1.28e-06***	-2.65e-08
	(4.66e-07)	(1.17e-07)
$log(\text{Precipitation} + 1)$	-2.08e-07*	-1.39e-08
	(1.14e-07)	(4.51e-08)
Number of Obs.	50,484	94,400
R-squared	0.691	0.717

Dependent Variable: $log(\text{PM2.5}_{i,t})$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t}) \times$		
Constant	0.0153**	0.0055**
	(0.0077)	(0.0027)
$log(\text{Precipitation} + 1)$	-0.0051	-0.0037**
	(0.0033)	(0.0015)
Number of Obs.	49,171	92,090
R-squared	0.712	0.736

Notes: This table presents the regression results regarding how the link between coal stockpiles and $PM_{2.5}$ concentration levels varies with precipitation. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter); the bottom panel of this table regresses the log of coal stockpiles on $PM_{2.5}$ in logs. In both panels, coal stockpiles is interacted with the log of total monthly precipitation (measured in inches); we include months-of-sample with zero precipitation by adding one to our total monthly precipitation variable. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor in Column 1 and we restrict our sample to plants between 25-50 miles of their AQ monitor in Column 2. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, each plant's monthly total electricity generation, and the sum of the coal stockpiles held by other plants within either 25 miles (for Column 1) or 50 miles (for Column 2) from the air quality monitor. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

Table A.4: Overall Effect of Coal Stockpiles on $PM_{2.5}$ Concentration: Additional Generation Controls

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t}$	9.90e-07**	-4.93e-08
	(3.97e-07)	(1.01e-07)
Observations	50,484	94,400
R-squared	0.691	0.717
Dependent Variable: $log(PM2.5_{i,t})$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t})$	0.0089*	0.0007
	(0.0054)	(0.0018)
Observations	49,171	92,090
R-squared	0.712	0.736

Notes: This table presents the regression results regarding the link between coal stockpiles and $PM_{2.5}$ concentration levels controlling flexibly for each plant's monthly electricity generation. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter); the bottom panel of this table regresses the log of coal stockpiles on $PM_{2.5}$ in logs. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor in Column 1 and we restrict our sample to plants between 25-50 miles of their AQ monitor in Column 2. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% level, respectively. We include a linear, quadratic, and cubic term for each plant's monthly electricity generation as well as logged monthly electricity generation as controls in these regressions. We also control for the log of the total quantity of coal (in tons) received by each plant in each month of sample, an indicator for whether the plant purchased any coal in that month, and the sum of the coal stockpiles held by other plants within either 25 miles (for Column 1) or 50 miles (for Column 2) from the air quality monitor. These regressions also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include facility code/AQ monitor fixed effects, county-of-AQ-monitor/year-of-sample fixed effects, and month-of-sample fixed effects. We weight observations by the inverse of the distance between air quality monitor and plant for these regressions. See Section 3.3 for more details regarding our regression specification. See Appendix Tables A.1 and A.2 for the summary statistics associated with these regressions.

A.2 Additional Figures/Tables: Mortality Rate Regressions

Table A.5: Mortality Rate Summary Statistics

Variable	Obs	Mean	Std. Dev.
Mortality Rate: Cardio (Per 100,000 People)	47,852	40.945	12.233
Mortality Rate: Respiratory (Per 100,000 People)	41,273	11.717	4.304
Mortality Rate: Total (Per 100,000 People)	48,887	119.778	26.350
Mortality Rate: External (Per 100,000 People)	33,902	7.019	2.585
Population	48,887	476,457	637,020
$PM_{2.5}$ (in $^{ug}/_{m^3}$)	48,887	11.974	3.928
Stockpiles (in Tons)	48,887	228,869	274,305
1(Quantity Received > 0)	48,887	0.879	0.326
Infant Mortality Rate: Total (Per 100,000 Infants)	12,704	20.534	6.322

Notes: This table presents the summary statistics for the OLS and IV regressions regarding the link between $PM_{2.5}$ concentration levels and mortality rates discussed in Section 5. A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We weight observations by the population of the county where the AQ monitor is located for these summary statistics. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor for this table. All of the mortality rates listed are for adults (ages 30+) except for the overall, all-cause mortality rate for infants (ages 0-4).

Table A.6: $PM_{2.5}$ on Mortality Rate: Semi-Elasticities

	log(Cardio MR)	log(Resp. MR)	log(Infant MR)	log(Total MR)	log(External MR)
$PM_{2.5}$	0.0104*	0.0313***	0.315***	0.0111**	-0.0080
	(0.0055)	(0.0112)	(0.0717)	(0.0045)	(0.0135)
Number of Obs.	49,002	42,270	12,704	50,176	34,715
R^2	0.897	0.607	0.501	0.889	0.746

Notes: This table presents the instrumental variables (IV) results regarding the link between $PM_{2.5}$ concentration levels and mortality rates. In particular, we use coal stockpiles as an instrument for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ in levels on mortality rates in logs. We run separate IV regressions for mortality rates associated with the cardiovascular system, the respiratory system, deaths due to any cause for infants (children ages 0-4), deaths due to any cause, and deaths due to external causes such as accidents; all of these mortality rates, except for the infant mortality rate, are calculated for adults (people ages 30 and older). A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% levels respectively. We control for the log of the total quantity of coal received by each plant in each month-of-sample, an indicator for whether the plant purchased any coal in that month, and each plant's monthly total electricity generation. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include county-of-AQ monitor/year-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.5 for the summary statistics associated with these regressions.

Table A.7: $PM_{2.5}$ on Mortality Rate: All Ages

	log(Cardio MR)	log(Resp. MR)	log(Total MR)	log(External MR)
$log(PM_{2.5})$	0.160***	0.274***	0.113***	-0.0290
	(0.0359)	(0.0696)	(0.0281)	(0.0693)
Number of Obs.	47,857	41,320	48,889	37,568
R^2	0.898	0.685	0.908	0.769

Notes: This table presents the instrumental variables (IV) results regarding the link between $PM_{2.5}$ concentration levels and all-age (read: people of all ages) mortality rates. We run separate IV regressions for mortality rates associated with the cardiovascular system, the respiratory system, deaths due to any cause, and deaths due to external causes such as accidents. In particular, we use coal stockpiles as an instrument for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on mortality rates. A unit of observation for these regressions is a plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% levels respectively. We control for the log of the total quantity of coal received by each plant in each month-of-sample, an indicator for whether the plant purchased any coal in that month, and each plant's monthly total electricity generation. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include county-of-AQ monitor/year-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.5 for the summary statistics associated with these regressions.

Table A.8: $PM_{2.5}$ on Mortality Rates:
Using Both Coal Stockpiles and Incidence of Delivery as Instruments

	log(Cardio MR)	log(Resp. MR)	log(Infant MR)	log(Total MR)	log(External MR)
$log(PM_{2.5})$	0.160***	0.284**	0.312**	0.109**	-0.0412
	(0.0607)	(0.125)	(0.125)	(0.0524)	(0.161)
Number of Obs.	47,847	41,218	12,704	48,887	33,830
R^2	0.892	0.661	0.586	0.901	0.753

Notes: This table presents the instrumental variables (IV) results regarding the link between $PM_{2.5}$ concentration levels and mortality rates. In particular, we use both monthly, plant-level coal stockpiles and an indicator for whether the plant purchased any coal in the month as instruments for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on mortality rates. We run separate IV regressions for mortality rates associated with the cardiovascular system, the respiratory system, deaths due to any cause for infants (children ages 0-4), deaths due to any cause, and deaths due to external causes such as accidents; all of these mortality rates, except for the infant mortality rate, are calculated for adults (people ages 30 and older). A unit of observation for these regressions is plant/air quality monitor/month-of-sample. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor. Standard errors are clustered by air quality monitor and are reported in parentheses. ***,**,* denote statistical significance at the 1%, 5%, and 10% levels respectively. We control for the log of the total quantity of coal received by each plant in each month-of-sample and each plant's monthly total electricity generation. We also include meteorological controls, smokestack emissions controls, and coal characteristics controls. Finally, we include county-of-AQ monitor/year-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.5 for the summary statistics associated with these regressions.

A.3 Additional Results: Quantification of Local Environmental Damages

Table A.9: Local Air Pollution Costs of Coal Stockpiles: Levels-Levels Specification

	$\frac{Dollars}{Ton}$	$\frac{Dollars}{MWh}$	$\frac{\text{Life-Years}}{\text{100,000 Tons}}$
Adult: Krewski et al. (2009)	90.45	43.51	11.73
Adult: IV Regression	153.13	73.65	19.86
Infant: IV Regression	20.97	9.87	16.75

Notes: This table presents the local environmental costs of coal storage and handling for the *median* plant in our sample. To calculate these costs, we use the partial effect of coal stockpiles on $PM_{2.5}$ implied by the estimates from our Levels-Levels specification allowing for different effects for plants upwind versus downwind from their matched air quality monitor (see the top panel of Table 2 in Section 4); we average over air quality monitors and months-of-sample in order to obtain plant-specific average partial effects. Row 1 of this table uses the estimated semi-elasticity between $PM_{2.5}$ (in micrograms per cubic meter) and adult (ages 30+) mortality rates estimated in Krewski et al. (2009), Row 2 uses our own estimated semi-elasticity between $PM_{2.5}$ and overall, all-cause adult mortality rates from Section 5 (see Column 4 of Appendix Table A.6), and Row 3 uses our estimated semi-elasticity between $PM_{2.5}$ and infant (ages 0-4) mortality rates (see Column 3 of Appendix Table A.6). Column 1 of this table presents local air pollution costs per ton of coal stockpiled. Column 2 presents local air pollution costs per MWh-equivalent of coal stockpiled; we convert tons of coal to MWhs of electricity by taking plant-level total number of tons of coal burned and dividing by plant-level total electricity generated (in MWh). For the first two columns of this table, we use a value of statistical life of 9.85 million dollars (expressed in 2013 U.S. dollars) to monetize the cost of the increased mortality from the $PM_{2.5}$ due to coal storage and handling. The last column presents local air pollution costs in life-years per 100,000 tons of coal stockpiled.

Table A.10: Local Environmental Costs of Coal Stockpiles in Life-Years Per GWh

	$\frac{\text{Life-Years}}{\text{GWh}}$
Adult: Krewski et al. (2009)	0.073
Adult: IV Regression	0.124
Infant: IV Regression	0.112

Notes: This table presents the local environmental costs of coal storage and handling for the *median* plant in our sample. To calculate these costs, we use the partial effect of coal stockpiles on $PM_{2.5}$ implied by the estimates from our Log-Log specification allowing for different effects for plants upwind versus downwind from their matched air quality monitor (see the bottom panel of Table 2 in Section 4); we average over air quality monitors and months-of-sample in order to obtain plant-specific average partial effects. This table presents external costs in life-years per GWh-equivalent of coal stockpiled; we convert tons of coal to GWhs of electricity by taking plant-level total number of tons of coal burned and dividing by plant-level total electricity generated (in GWh). Row 1 of this table uses the estimated semi-elasticity between $PM_{2.5}$ (in micrograms per cubic meter) and adult (ages 30+) mortality rates estimated in Krewski et al. (2009), Row 2 uses our own estimated semi-elasticity between $PM_{2.5}$ and overall, all-cause adult mortality rates from Section 5 (see Column 4 of Appendix Table A.6), and Row 3 uses our estimated semi-elasticity between $PM_{2.5}$ and infant (ages 0-4) mortality rates (see Column 3 of Appendix Table A.6).

B Data Appendix

This Appendix section describes the data used in this paper as well as the data construction process.

B.1 Data Sources: $PM_{2.5}$, CO, SO_2, and NO_2 Concentration Levels

We use the Air Quality System (AQS) data provided by the United States Environmental Protection Agency (USEPA). This publicly available database includes hourly readings of ambient $PM_{2.5}$ concentration levels (in micrograms per cubic meter), carbon monoxide (CO) concentration levels (measured in parts per million), sulfur dioxide (SO_2) concentration levels (measured in parts per billion), and nitrogen dioxide (NO_2) concentration levels (measured in parts per billion) at many different monitored sites across the contiguous United States. We aggregate these data to monthly average $PM_{2.5}$, CO, SO_2, and NO_2 concentration levels for each air quality monitor for the sample period 2002-2012. Importantly, the AQS database also provides the geocoded location (read: latitude and longitude) of each air quality monitor.

These hourly, monitor-level $PM_{2.5}$, CO, NO_2, and SO_2 data are available at: http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html.

B.2 Data Sources: Coal Purchase and Storage

We collect monthly, plant-level data on end-of-month coal stockpiles, total monthly coal consumption, and total monthly electricity generation from Forms EIA-906 (for 2002-2007) and EIA-923 (for 2008-2012).[32] Data on monthly, plant-level coal stockpiles for 2002-2012 are considered proprietary; we obtained a research contract with the Energy Information Administration (EIA) in order to use these data for our analysis.

We construct monthly, plant-level coal purchase quantities (in tons), average delivered coal prices (in dollars per mmBTU), average heat content (in mmBTUS), average sulfur content (in percentage of overall tonnage) and average ash content (in percentage

[32] Monthly, plant-level total coal consumption and monthly, plant-level total electricity generation are available at http://www.eia.gov/electricity/data/eia923/.

of overall tonnage) from Forms EIA-423 (2002-2007) and EIA-923 (2008-2012).[33] The variables in this purchase-level dataset include month of purchase, quantity purchased, delivered price, heat content, sulfur content, ash content, county of origin, and whether the delivery came from a long-term contract or the spot market. These data are at the "order level"; an "order" as defined by these forms based on the following criteria:

"Data on coal received under each purchase order or contract with a supplier should be reported separately. Aggregation of coal receipt data into a single line item is allowed if the coal is received under the same purchase order or contract and the purchase type, fuel, mine type, state of origin, county of origin, and supplier are identical for each delivery."

Due to the aggregation from deliveries to orders in the raw data, we focus on each plant's incidence of delivery in each month-of-sample (read: whether the plant received any coal in that month) rather than the plant's monthly number of coal deliveries (which we don't observe because of this aggregation).

We only consider electricity generation plants whose "primary business purpose is the sale of electricity to the public"[34]; this excludes plants that also sell significant quantities of heat ("combined heat and power plants") as well as commercial and industrial plants that generate electricity for their own use.

The EPA eGrid database provides the geocoded location (read: latitude and longitude) of each coal-fired power plant; this database is located at: http://www.epa.gov/energy/egrid.

B.3 Data Sources: Meteorological Variables

Our meteorological data come from the quality controlled local climatological data (QCLCD) collected by the National Climatic Data Center (NCDC); these data include hourly wind speed and direction, dry bulb temperature, wet bulb temperature, dew-point temperature, relative humidity, station pressure, and precipitation at approximately 1,600 U.S. locations. We aggregate these data to the meteorological monitor/month-

[33]This purchase-level dataset is available at: http://www.eia.gov/electricity/data/eia423/ for pre-2008 data and http://www.eia.gov/electricity/data/eia923/ for post-2008 data.

[34]This quotation is from the EIA Form 923 data dictionary.

of-sample level by taking time-weighted averages over hours-of-sample of dry bulb temperature, wet bulb temperature, dew-point temperature, relative humidity, and station pressure; we use the meteorological monitor/month-of-sample level sum of hourly precipitation. Wind speed is an important factor in determining both how much $PM_{2.5}$ is generated from various sources as well as how this $PM_{2.5}$ is dispersed. Thus, we also control for the (5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95) hourly percentiles of wind speed, calculated over all hours-of-sample for each meteorological monitor/month-of-sample. Finally, we use the wind-speed weighted monthly average wind direction measured at each meteorological monitor. Importantly, the NCDC database also provides the geocoded location (read: latitude and longitude) of each meteorological monitor.

These hourly, monitor-level, meteorological data are available at: http://www.ncdc.noaa.gov/orders/qclcd/.

B.4 Data Sources: SO_2, CO_2, and NO_x Emissions from Coal Combustion

The EPA's Continuous Monitoring Emissions System (CEMS) collects hourly data for each plant on SO_2, CO_2, and NO_x emissions (in tons) resulting from coal burned; we sum these hourly data to the monthly level and control for the total SO_2, CO_2, and NO_x emissions from each plant's smokestack in each month-of-sample.

These hourly, plant-level data on the smokestack emissions associated with burning fuel are available at: http://ampd.epa.gov/ampd/.

B.5 Data Merge

We merge each air quality monitor i to meteorological monitors and coal-fired power plants as follows:

1. For each month-of-sample, we find all meteorological monitors within M miles of air quality monitor i. We take th weighted average of the meteorological data (for example, wind speed and wind direction) across these meteorological monitors for each air quality monitor i, where we weight by the inverse of the distance between the air quality monitor and the meteorological monitor.

2. If $M = 25$ miles, we consider all coal-fired power plants less than 25 miles away from air quality monitor i. If $M = 50$ miles, we consider all coal-fired power plants between 25 miles and 50 miles away from air quality monitor i.

Thus, our unit of observation is an air quality monitor/power plant pair for each month-of-sample, emphasizing that each air quality monitor can be linked to multiple power plants for a given month-of-sample. We examine how the effects of coal stockpiles and number of deliveries on $PM_{2.5}$ concentration levels decay with distance by separately estimating these effects for plants within 25 miles of their corresponding air quality monitor versus plants between 25 miles and 50 miles away from their corresponding air quality monitor.