The Local Environmental Costs of Coal Procurement at U.S. Power Plants

Akshaya Jha and Nicholas Z. Muller[*][†]

April 6, 2017

Abstract

Burning coal is known to have environmental costs; this paper quantifies the local environmental costs of transporting and storing coal at U.S. power plants for the sample period 2002-2012. We first demonstrate that a 10% increase in coal stockpiles (number of deliveries) results in a 0.07% (0.16%) increase in the average concentration of fine particulates ($PM_{2.5}$) for locations up to 25 miles away from, and downwind from, plants. We next assess the impacts of $PM_{2.5}$ on average adult and infant mortality rates using coal stockpiles and deliveries as instruments for $PM_{2.5}$. Our findings within this instrumental variables framework indicate that a 10% increase in $PM_{2.5}$ leads to a 1.1% (6.6%) increase in average adult (infant) mortality rates; comfortingly, these causal estimates are similar in magnitude to the epidemiological estimates used by the USEPA in their regulatory impact analyses. Our estimated increase in mortality rates implies local environmental costs of \$183 (\$203) per ton of coal stockpiled (delivered); to put this in perspective, the average power plant paid roughly \$48 per ton for coal during our sample period. These sizable but highly localized environmental costs of coal transportation and storage disproportionately impact the economically disadvantaged communities living near coal-fired 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: Department of Economics, Middlebury College and NBER, 303 College Street, Middlebury, VT 05753. 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

The reliance on coal as an energy source has a multitude of well-known environmental consequences.[1] 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. As a result of these different environmental impacts, coal is subject to a plethora of regulatory constraints. 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 (USEPA). Finally, the combustion of coal for power generation and manufacturing is regulated by the Clean Air Act. Seemingly every stage of the supply chain from coal production at mines to coal burned by power plants is subject to a policy constraint.

Despite this thicket of environmental regulations relevant to coal, our paper uncovers an as yet unstudied dimension of coal use that we argue requires policy intervention: the environmental consequences of the coal purchase and storage behavior of U.S. power plants. We first demonstrate that coal deliveries to power plants and the level of coal stockpiles held at these plants result in statistically significant increases in the concentrations of fine particulates ($PM_{2.5}$) within 25 miles of these plants.[2] This highly local but sizable effect of coal procurement on ambient $PM_{2.5}$ levels provides local, state, and federal regulators with a new lever they can use to comply with the National Ambient Air Quality Standards (NAAQS) established by the Clean Air Act. Moreover, large quantities of coal are also stored, transported, and handled at coal mines and coal export terminals. Coal export terminals in particular are located in densely populated coastal cities such as Los Angeles, Houston, and Baltimore. An empirical connection between coal delivery, storage, and ambient $PM_{2.5}$ can inform the ongoing debate regarding the potential adverse health impacts of coal terminals on local communities.[3]

---

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

[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.

[3] There is staunch local opposition to two proposed coal export terminals in Washington (near Bellingham and Longview) as well as a proposed terminal in Oakland due in part to concerns surrounding the

We also test for a causal relationship between mortality rates and exposure to $PM_{2.5}$ using coal deliveries and stockpiles as instruments for $PM_{2.5}$ exposure. Few empirical relationships have a more significant bearing on environmental policy evaluation than the link between $PM_{2.5}$ and mortality. For example, the USEPA has repeatedly found that the vast majority of the benefits from the Clean Air Act come from reductions in mortality risk due to decreased $PM_{2.5}$ exposure ((EPA, 1999); (EPA, 2010)). Thus, the estimated magnitude of the benefits from any policy designed to reduce air pollution is quite sensitive to which $PM_{2.5}$/mortality estimate is used (EPA, 2010). We demonstrate this point by quantifying the monetary social costs of plants' coal purchase and storage behavior using the same approach as the USEPA's benefit-cost analyses of the Clean Air Act. This calculation relies on our estimated relationship between coal procurement and $PM_{2.5}$, the relationship between $PM_{2.5}$ and mortality rates, as well as standard valuation techniques to monetize changes in mortality. We consider two different estimates of the link between $PM_{2.5}$ and mortality rates: our estimate versus the epidemiological estimate currently used by the USEPA. Though our $PM_{2.5}$/mortality-rate relationship is not markedly different in magnitude from existing epidemiological estimates, the resulting local environmental costs of coal procurement differ substantially based on which of the two estimates is used.

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 (EIA). These coal procurement 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 (AQS) database maintained by the Environmental Protection Agency (EPA). Weather conditions play a large role in dictating $PM_{2.5}$ levels; moreover, wind speed, wind direction, and precipitation are important determinants of the extent to which coal stockpiles and coal deliveries 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 (NCDC). 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

---

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 transport and storage of large coal shipments within their jurisdiction; a recent New York Times article (Fuller, 2016) provides more details regarding this ban.

(CDC Wonder, 2016).

To estimate the impact of coal stockpiles and coal deliveries on local $PM_{2.5}$ concentration levels, we specify a spatial econometric framework that matches air quality (AQ) monitoring stations to coal-fired power plants within a given distance radius. This framework allows us to test for the distance at which plants' coal purchase and storage behavior ceases to affect pollution levels. We find that a 10% increase in coal stockpiles (number of deliveries) results in a 0.06% (0.12%) increase in average $PM_{2.5}$ concentration levels for coal-fired power plants within 25 miles of their matched air quality monitor. In contrast, there is no statistical link between coal procurement (for either coal stockpiles or deliveries) and $PM_{2.5}$ levels for plants beyond 25 miles of their matched air quality monitor; the effect of coal procurement on $PM_{2.5}$ concentration levels is highly localized.[4] This is intuitive, given that the emissions from coal stockpiles, coal handling, and trains delivering coal all occur at ground level and are therefore not likely to be entrained in upper-level winds. In contrast, the effects of coal combustion on $PM_{2.5}$ are regional because the emissions from coal combustion are typically released from tall smokestacks.

Importantly, we find no link between locally monitored carbon monoxide and either coal stored or delivered using the same empirical framework; this provides strong evidence that our estimated relationship between coal procurement 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 effects of coal procurement on $PM_{2.5}$ are higher: 1) for areas downwind from a plant, and 2) for lower levels of precipitation.[5] These empirical results using spatiotemporal variation in weather, our placebo tests considering carbon monoxide, as well as our battery of controls give us confidence that we have causally identified how coal stored and delivered affects $PM_{2.5}$.

We next estimate the impact of $PM_{2.5}$ concentration levels on mortality rates using the coal stored at, and obtained by, power plants as instruments for $PM_{2.5}$ exposure.

---

[4] 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; they similarly find that the local air pollution costs from coal-fired power plants were concentrated primarily within 30 miles of these plants.

[5] 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.

This instrumental variables (IV) strategy relies on the assumption that a plant's coal procurement behavior is unlikely to affect local mortality rates except through its effect on $PM_{2.5}$ exposure. We argue that this identifying assumption is likely to hold, thus asserting that our paper presents a novel identification strategy for the causal relationship between $PM_{2.5}$ and mortality rates. Our analysis consider mortality rates associated with different causes of death: deaths due to the cardiovascular system, deaths due to the respiratory system, and deaths due to the nervous system. We also examine total number of deaths of any type, both for people over 30 years old ("adults") and children aged 0 to 4 ("infants"). Finally, deaths due to external causes such as accidents are considered as a placebo test. Our specifications include a bevy of controls as well as county-year fixed effects. We find an economically small (and sometimes negative) association between $PM_{2.5}$ and mortality rates if we simply estimate this relationship using ordinary least squares. However, the 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 and the monthly number of coal deliveries from nearby power plants. Of particular note, we find that a 10% increase in $PM_{2.5}$ leads to a 1.1% (6.6%) 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% increase in adult mortality. The findings from the two epidemiological studies most commonly used by the EPA indicate a mortality response between 0.6% and 1.3% per unit of $PM_{2.5}$ (Krewski et al. (2009); Lepeule et al. (2012)). Previous empirical research reports that the damages from air pollution are significant in magnitude and that these damages are mostly due to increased mortality risk from exposure to $PM_{2.5}$.[6] However, recent work in economics has challenged the causal basis for the link between exposure to PM and the adverse effects on mortality risk reported in epidemiological studies.[7] 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 EPA to quantify the impacts of $PM_{2.5}$ emissions. Our IV results also provide further empirical evidence that there is a link

---

[6] This empirical literature includes NRC (2010), EPA (2010), Muller, Mendelsohn and Nordhaus (2011), and Muller (2014).

[7] This work includes (among others) Currie, Neidell et al. (2005), Currie, Neidell and Schmieder (2009), and Chay and Greenstone (2003).

between $PM_{2.5}$ concentration levels and plants' coal procurement behavior.

Finally, we combine our estimates of the effect of coal procurement on $PM_{2.5}$ with our IV estimates of the effect of $PM_{2.5}$ on mortality rates in order to calculate the increased number of deaths from coal deliveries and storage. These health risks are monetized using the value of statistical life (VSL) approach that is standard both for academic studies and federal regulatory impact analyses ((EPA, 1999); (EPA, 2010)). We calculate that the local environmental cost of $PM_{2.5}$ increases from coal stockpiles is \$182.67 per ton of coal stockpiled; the local air pollution cost per ton of coal delivered is \$202.51. Roughly 75% of this social cost is due to adult mortality while 25% 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, stockpiles 212,781.6 tons of coal and has 106,235 tons of coal delivered to it each month. However, if we translate tons of coal to MWh of electricity, our local air pollution costs are \$95.51 per MWh-equivalent of coal stored and \$103.12 per MWh-equivalent of coal delivered. Levy, Baxter and Schwartz (2009) finds that the social cost of increased $PM_{2.5}$ exposure from coal being burned is roughly \$230 per MWh. Thus, our empirical estimates of the local environmental damages from purchasing and 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 local air pollution costs of transporting and storing coal are borne primarily by people living within 25 miles of coal-fired power plants. Given that people living in census tracts with power plants have lower per-capita incomes and educational attainment on average relative to residents of census tracts without power plants, the highly localized environmental costs of coal procurement disproportionately affect economically disadvantaged communities.[8]

Summarizing, our paper has two primary contributions. First, we inform the policy debate surrounding the environmental impacts of coal transportation and storage at coal mines, power plants, and export terminals. Second, we quantify the relationship between $PM_{2.5}$ and mortality rates; this estimated relationship is arguably the most important component of any cost-benefit analysis of any environmental policy designed

---

[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.

to improve air quality. Moreover, many purely economic policies also have the unintended consequence of affecting air quality. For example, Jha (2017) argues that U.S. coal-fired electricity generation plants facing output price regulation hold larger stockpiles of coal on-site for the same level of coal consumption and receive more deliveries per month for the same quantity of coal purchased relative to plants facing electricity market mechanisms. The results from our paper indicate that this regulatory-induced increase in coal stockpiles and number of deliveries has environmental costs in addition to the economic costs documented in Jha (2017). This example demonstrates that the debate surrounding even purely economic policy questions, such as whether electricity generation should be provided under regulation versus markets, must account for the environmental impacts of the different policies.

The remainder of the paper proceeds as follows. In Section 2, we elaborate on the physical process underlying how plants' coal purchase and storage behavior can result in higher $PM_{2.5}$ concentration levels. Section 3 describes the data sources and the methodology used to estimate the relationship between plants' coal procurement behavior and $PM_{2.5}$ concentration levels, while Section 4 presents the empirical results demonstrating that increases in coal purchase and storage behavior lead to 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 $PM_{2.5}$ emissions from coal purchased and stored. Finally, we conclude in Section 7 by illustrating how our findings can be applied to a variety of other policy contexts, such as the environmental costs of coal dust at mines and export terminals as well as the environmental impacts of distortions to plant-level coal purchase and stockpiling behavior due to the structure of output price regulation.

2 Economic and Environmental Background

2.1 Input Coal Procurement Behavior of U.S. Power Plants

Coal-fired power plants burn coal in order to heat water into steam that drives the turbines used to generate electricity. These power plants typically run continuously;

their electricity generation does not vary substantially day-to-day. Plants inject their electricity into the transmission grid where it is distributed regionally to where electricity is needed; thus, short-term fluctuations in electricity demand near a coal-fired power plant do not map directly into its electricity generation. Jha (2017) describes the factors underlying the coal purchase and storage behavior of U.S. coal-fired power plants; these factors include regional electricity demand, spot coal prices, and natural gas prices. All of these factors vary regionally rather than local to the power plant. We argue based on this that plant-level electricity demand (and thus plant-level coal procurement) is unlikely to be affected by local economic conditions. This assertion is important because local mortality rates are undoubtedly correlated with unobserved local economic conditions, which is why we identify increases in mortality rates due to increases in $PM_{2.5}$ using coal delivered and stored as instruments.

Power plants purchase the majority of their input coal from long-term contracts with coal suppliers, purchasing the remainder from spot markets. They also store large quantities of coal on-site to hedge against coal price and supply risks; these coal inventories are very rarely resold in practice.[9] Coal is not homogeneous; coal mined from different regions of the United States differs in heat content, sulfur content, ash content, distance from mine to plant, etc. Coal-fired power plants primarily value the heat content of coal; burning a ton of coal with a higher heat content generates a larger amount of output electricity. Sulfur content and ash content are important primarily for their environmental impacts. Finally, roughly 67 percent of coal is transported via rail. The remaining coal is shipped via barge (12%), trucks (10%) and various other modes of transportation used primarily for short distances (ex: conveyors, pipelines, etc.).[10] The local environmental costs of coal delivered to a power plant can differ substantially based on the mode of transportation.

---

[9] This lack of resale is due primarily to the fact that coal transportation infrastructure is designed to bring coal *to* power plants rather than away from them; therefore, coal becomes very costly to resell once it's in a plant's stockpile.

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

2.2 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. 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 emissions from coal stockpiles and coal deliveries. We posit two mechanisms for emissions from coal stockpiles. The first is wind erosion. 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.

Roughly 67 percent of coal is delivered by train; the coal carried by these trains is typically in uncovered freight cars. Barges and trucks carrying coal are also typically uncovered. Finally, there is a significant coal handling process at the power plant any time coal is delivered, which includes moving the coal from the train to different places within the plant site, separating "light dust" from the coal, and crushing the coal in order to make it suitable for burning.[11] Thus, we expect coal deliveries to result in increased $PM_{2.5}$ levels due to wind erosion (dust entrainment), gaseous discharge and the movement of coal piles around the plant. Additionally, trains run on diesel fuel; burning diesel is also associated with increased $PM_{2.5}$ concentration levels. Thus, the burning of diesel may be another factor determining the local environmental costs from coal deliveries.

Our empirical strategy does not distinguish between the two types of passive emissions (dust entrainment and gaseous discharges) from coal stockpiles; $PM_{2.5}$ emissions resulting

---

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

from both of these mechanisms increase mortality rates in local populations and thus contribute to local environmental health costs. Similarly, we do not disentangle how much of the $PM_{2.5}$ increase from coal deliveries comes from passive $PM_{2.5}$ emissions from coal piles versus emissions from the coal handling process when coal is delivered, as both of these factors contribute to the local environmental costs of transporting coal. However, we provide empirical evidence that our effect of coal procurement (either stockpiles or deliveries) on $PM_{2.5}$ does not result from the combustion of *any* fuel; our effect is not driven by either the coal burned by the power plant or the diesel or gasoline burned by the trains, barges and trucks transporting coal. To be clear, these combustion-based sources do emit $PM_{2.5}$, but we focus on the local environmental costs at the power plant from purchasing and storing coal in this paper.

3 Local Environmental Impacts of Coal Procurement: Data and Methodology

This section describes the data used to estimate the link between plant-level coal procurement behavior (coal stockpiles and number of deliveries) and $PM_{2.5}$ concentration levels at nearby air quality monitors. We also present the empirical framework used to measure this relationship, which includes specifying our set of controls for alternative sources of $PM_{2.5}$, such as burning coal, as well as other factors that increase or decrease $PM_{2.5}$ for 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).[12] 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. Finally, we only consider electricity generation plants whose primary business purpose is the sale of electricity to the public; this excludes plants that also sell signif-

---

[12] The data regarding fuel inventories for 2002-2012 are confidential; we obtained a research contract from the EIA in order to use these restricted-access data.

icant 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 Environmental Protection Agency (EPA). 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 U.S. locations. We aggregate these data to the meteorological monitor/month-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; the results presented below differ very little if we instead take an un-weighted average.

We additionally control for a variety of other factors that can affect the $PM_{2.5}$ concentration reading at an air quality monitor in each month. First, 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 for each plant in each month-of-sample. We also control for the total monthly quantity of coal purchased by each plant as well as the total monthly electricity generation produced by each plant; both 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 (as well as wind direction in each month-of-sample) for each meteorological monitor, and the EPA eGrid database provides the latitude and longitude for each coal-fired power plant. We use these variables in order 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.[13]
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. We also show how the effect of coal procurement 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.

---

[13] Our empirical results do not differ substantially if we instead take an un-weighted average of the meteorological data.

3.3 Empirical Framework

We are interested in estimating the effect of coal stockpiles held by coal-fired power plants on local $PM_{2.5}$ concentration levels. We consider both "Log-Log" and "Levels-Levels" specifications. Our Log-Log specification for how coal stockpiles impact $PM_{2.5}$ levels is:

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

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.

We control for the total level of coal stockpiles across all plants $k \neq p$ 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 $PM_{2.5}$ in our estimate of $\gamma_{i,p}$. We also control for a wide variety of alternative factors that affect local $PM_{2.5}$ levels, such as annual sulfur content and ash content of coal purchased by the plant, monthly quantity of coal received by the plant, monthly number of coal deliveries to the plant, the plant's monthly total electricity generation, 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. Finally, we include fixed effects for each county-of-air-quality-monitor/month-of-sample ($\alpha_{c,t}$) and air quality monitor/power plant ($\theta_{i,p}$). We report a specification considering one overall coefficient ($\gamma_{i,p} \equiv \gamma$), as well as specifications allowing $\gamma_{i,p}$ to vary by relative wind direction and precipitation.[14]

Our goal is to estimate how the *level* of $PM_{2.5}$ measured by a given air quality monitor is affected by a one ton increase in coal stockpiles at a given plant, controlling for all of the other factors listed above. The Log-Log specification in Equation 1 implies that this

---

[14] 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. The empirical results are similar if we 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 $log(CS_{p,t})$.

partial effect in levels for observation (i, p, t) is:

$$\frac{dP\hat{M}2.5_{i,t}}{dCS_{p,t}} = \hat{\gamma_{i,p}}(\frac{PM2.5_{i,t}+1}{CS_{p,t}})$$

We also estimate a Levels-Levels specification in order to report partial effects in levels. This Levels-Levels specification, presented below, additionally serves as a sensitivity check of the functional form relating coal stockpiles to $PM_{2.5}$:

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

This specification includes the same set of controls as listed above; the partial effect associated with coal stockpiles is simply $\hat{\gamma_{i,p}}$ in the Levels-Levels case.

Finally, we also estimate both the Log-Log and Levels-Levels specifications replacing coal stockpiles ($CS_{p,t}$) with the number of coal deliveries ($ND_{p,t}$) arriving at plant p in month-of-sample t.[15] We control for the same factors as described above, with the obvious exception of the number of deliveries (our dependent variable); our number of deliveries specifications instead control for coal stockpiles. We estimate these specifications both controlling for the total monthly quantity received and not controlling for quantity received in order to explore the extent to which $PM_{2.5}$ emissions stem from an additional ton of coal delivered versus an additional delivery (no matter how much coal is delivered).

For all specifications, our standard errors are clustered at the air quality monitor level and we weight our regressions by the inverse distance between air quality monitor and power plant.[16]

---

[15] We consider $log(ND_{p,t}+1)$ for the Log-Log specifications given that there are a non-trivial number of plant/month-of-sample observations where there are zero deliveries (read: the plant did not purchase coal in that month-of-sample).

[16] 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 for both of these alternative weighting schemes; these results are available upon request.

4 Local Environmental Impacts of Coal Procurement: Empirical Findings

In this section, we present our regression results regarding the link between coal purchase/stockpiling behavior and $PM_{2.5}$ concentration levels. We find that a 10% increase in coal stockpiles (number of deliveries) results in a 0.06% (0.12%) increase in average $PM_{2.5}$ concentration levels for populations within 25 miles of power plants. We find no average effect of either number of deliveries or 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 wind erosion and volatile emissions from coal piles at or near ground level can only travel so far. The decay of our effect after 25 miles also makes clear that we are primarily capturing the effect of coal deliveries at the power plant rather than all along the route from origin mine to destination power plant.

We should also expect more severe increases in $PM_{2.5}$ concentrations for local populations downwind from coal piles and train deliveries (read: the wind is blowing from the coal pile to the local population). Consistent with this intuition, we estimate that the average $PM_{2.5}$ increase from coal purchase and storage behavior 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 $PM_{2.5}$ increases from both coal stockpiles and number of deliveries are lower for higher levels of precipitation. Finally, we consider a variety of robustness checks in order to demonstrate that our empirical results are not driven by the emissions from the coal burned by power plants or the fuel burned by trains, barges, or trucks delivering coal to the power plant. For example, carbon monoxide (CO) is emitted when any fuel is burned; we demonstrate that there is no effect of either coal stockpiles or number of deliveries on CO at any distance bandwidth. Summarizing, we have confidence that we've identified a causal relationship between coal procurement (coal stockpiles and deliveries) and $PM_{2.5}$ for three reasons. First, we find no effect of coal procurement on CO as expected. 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 coal procurement results in higher levels of $PM_{2.5}$ for AQ monitors downwind of power plants and observations with low levels of precipitation.

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

The top (bottom) panel of Table 1 displays the results from the Levels-Levels (Log-Log) specification with the number of coal deliveries arriving at each plant in each month-of-sample as the covariate of interest. The summary statistics corresponding to all regressions discussed in this section are in Appendix Tables A.1 and A.2. The unit of observation for all regressions in this section is an air quality monitor/power plant pair in a given month-of-sample.

We find that an additional delivery results in an increase in monthly average $PM_{2.5}$ levels of 0.018 micrograms per cubic meter ($^{ug}/_{m^3}$) when we restrict our sample to plants within 25 miles of their corresponding air quality (AQ) monitor. This regression model explains nearly 85% of the variation in ambient $PM_{2.5}$ readings. Similarly, the bottom panel of Table 1 indicates that a 10% increase in number of deliveries is associated with a 0.12% increase in average $PM_{2.5}$ concentration levels at AQ monitor/power plant pairs within 25 miles of each other. The effect of coal deliveries on $PM_{2.5}$ is statistically significant for the 25 mile bandwidth for both the Levels-Levels and Log-Log specifications. However, we find no statistical link between average $PM_{2.5}$ concentration levels and number of deliveries for either the Levels-Levels or Log-Log specifications when focusing on plants between 25 and 50 miles from their corresponding AQ monitors.[17] This result indicates that we are capturing the effect of coal being delivered to the power plant rather than the $PM_{2.5}$ emissions from coal piles in open train hoppers all along the route from origin coal mine to destination power plant.

Coal deliveries may increase $PM_{2.5}$ concentration levels in a number of ways, including wind erosion and gaseous emissions from coal transported in uncovered piles, coal handling at the power plant, and the fuel burned by the trains, barges, and trucks delivering coal. However, the combustion of any fuel emits carbon monoxide as well as $PM_{2.5}$, and we show below in Section 4.4 that there is no statistical link between carbon monoxide (CO) and number of deliveries. This rules out any combustion-based explanation for our effect such as the burning of coal by power plants or the burning of

---

[17] We also ran all of our regressions for plants between 50-100 miles and 100-200 miles away from their AQ monitors. We do not find a statistically significant effect for these distance bandwidths for most specifications, which is unsurprising given that we demonstrate that our effect of coal procurement on $PM_{2.5}$ decays even for distances farther than 25 miles away.

Table 1: Overall Effect of Number of Deliveries on $PM_{2.5}$ Concentration

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t}$	0.0180**	0.0011
	(0.0085)	(0.0048)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1)$	0.0124**	0.0034
	(0.0053)	(0.0034)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding the link between number of deliveries and $PM_{2.5}$ concentration levels. The top panel of this table regresses number of deliveries on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses number of deliveries 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 and 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 quantity received, coal stockpiles, and sum of deliveries from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

fuel by the trains, barges, and trucks delivering coal. Also, the effect of the quantity of coal delivered on $PM_{2.5}$ is not statistically significant in either the Levels-Levels or the Log-Log specifications presented in Table 1. This provides evidence that it is the act of delivering coal rather than the quantity of coal delivered that generates the majority of our estimated increase in local $PM_{2.5}$ concentration levels.[18] As described in Section 2.2, the coal handling and preparation process at power plants is quite involved. $PM_{2.5}$ emissions can be generated during many stages of this process, including when power

---

[18] We also present the effect of number of deliveries on $PM_{2.5}$ emissions without controlling for total quantity of coal delivered in Appendix Table A.3. We see from this table that there is a statistically significant and positive effect of number of deliveries on $PM_{2.5}$ for plants within 25 miles of their AQ monitor. Moreover, this effect has similar size whether or not we control for quantity delivered (0.0180 in Table 1 versus 0.0146 in Appendix Table A.3 for the Levels-Levels specification). This provides further evidence that the majority of our effect of coal deliveries on $PM_{2.5}$ comes from an additional shipment rather than more coal being delivered in a given shipment.

plant managers extract light dust from coal, crush coal to make is suitable to burn, or simply move coal around on-site. Based on the empirical evidence provided above, our effect of the number of deliveries on $PM_{2.5}$ is likely driven by the handling of coal at power plants rather than wind erosion/gaseous emissions from open coal piles on trains, barges, and trucks or the fuel burned by these modes of transportation.

The top panel of Table 2 reports the results from the Levels-Levels specification where the size (in tons) of the coal stockpiles held at each power plant in each month-of-sample is the covariate of interest. A one ton increase in coal stockpiles increases ambient $PM_{2.5}$ levels by 1.02e-06 $^{ug}/_{m^3}$ when only plants within 25 miles of their corresponding air quality (AQ) monitor are included. However, we find no statistical link between coal stockpiles and $PM_{2.5}$ when focusing on plants between 25 and 50 miles of their AQ monitor. We should expect this given that the $PM_{2.5}$ emissions from the wind erosion of stationary coal stockpiles at ground level and the volatile emissions from these coal piles can only travel so far. The bottom panel of Table 2 shows the estimated elasticities from the Log-Log specification for the coal stockpile models. We see that a 10% increase in coal stockpiles results in a 0.06% increase in average $PM_{2.5}$ concentration levels for plants within 25 miles of their AQ monitor, but there is no statistical link between coal stockpiles and $PM_{2.5}$ for plants between 25 and 50 miles of their AQ monitors.

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

In this subsection, we interact the covariate of interest (number of deliveries in Table 3 and coal stockpiles in Table 4) with the relative bearing from coal-fired power plant to 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°.

The top panel of Table 3 shows that, for the sample restricted only to plants within 25 miles of each AQ monitor, the partial effect of number of deliveries on $PM_{2.5}$ is statis-

Table 2: Overall Effect of Coal Stockpiles on $PM_{2.5}$ Concentration

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t}$	1.02e-06***	3.18e-08
	(2.26e-07)	(1.06e-07)
Observations	48,521	92,886
R-squared	0.847	0.824
Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t})$	0.0061**	-0.0016
	(0.0031)	(0.0018)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding the link between coal stockpiles and $PM_{2.5}$ concentration levels. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses 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 and 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% levels respectively. We control for quantity received, number of deliveries, and sum of coal stockpiles from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

tically significant and positive for monitors downwind from plants, but not statistically significant for monitors upwind of plants. Moreover, the effect for the 25 mile bandwidth for downwind monitors in Table 3 (0.0244) is larger than the overall average effect across both upwind and downwind AQ monitors reported in Table 1 (0.0180). We should expect more severe increases in $PM_{2.5}$ concentrations for local populations downwind from coal deliveries given that $PM_{2.5}$ particulates are carried from the power plant to the AQ monitor by these downwind currents. Finally, we find no link between the number of coal deliveries and $PM_{2.5}$ for plants between 25 and 50 miles away from their AQ monitor, regardless of whether they are upwind or downwind from their AQ monitor.

The bottom panel of Table 3 shows the results for the directional models expressed in terms of elasticities between number of deliveries and $PM_{2.5}$. As with the Levels-Levels specification, we find a positive and statistically significant elasticity only for AQ

Table 3: Effect of Number of Deliveries on $PM_{2.5}$ Concentration: Upwind vs. Downwind

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t} \times$		
Monitor Downwind from Plant	0.0244**	0.0039
	(0.0100)	(0.0047)
Monitor Upwind from Plant	0.0110	-0.0021
	(0.0105)	(0.0066)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1) \times$		
Monitor Downwind from Plant	0.0163***	0.0042
	(0.0056)	(0.0036)
Monitor Upwind from Plant	0.0089	0.0026
	(0.0058)	(0.0036)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding how the link between number of deliveries and $PM_{2.5}$ concentration levels varies with wind direction. The top panel of this table regresses number of deliveries on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses number of deliveries on $PM_{2.5}$ in logs. In both panels, number of deliveries 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, 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 and 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 quantity received, coal stockpiles, and sum of deliveries from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

monitors that are both within 25 miles of their plant and downwind from their plant. We find no effect of number of deliveries on $PM_{2.5}$ for: 1) plants between 25 and 50 miles from their AQ monitor, or 2) AQ monitors upwind from their plant. As expected, the downwind elasticity reported for the 25 mile bandwidth in Table 3 (0.0163) is larger than the corresponding overall estimate reported in the bottom panel of Table 1 (0.0124).

The top panel of Table 4 displays the results for the directional models in which the size of coal stockpiles (in tons) is the independent variable of interest. We see a similar pattern in empirical results as with the number of deliveries regressions. First, the interaction terms between coal stockpiles and the dummy variable for (relative) angles between 0° to 90° is positive and statistically significant for plants less than 25 miles away from their AQ monitor for both the Levels-Levels (top panel) and Log-Log (bottom panel) specifications. The estimated coefficients for downwind AQ monitors are larger than the corresponding estimates for upwind AQ monitors for both the less than 25 miles (Column 1) and between 25-50 miles (Column 2) distance bandwidths. However, neither the upwind nor downwind coefficient estimates are statistically significant for plants between 25 and 50 miles from their AQ monitor. Finally, the estimated coefficient for downwind monitors in Table 4 (1.14e-06) is larger than the corresponding overall estimate reported in Table 2 (1.02e-06). We draw the exact same conclusions when examining the Log-Log specification presented in the bottom panel of Table 4 as described above for the Levels-Levels specification. $PM_{2.5}$ 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 fact that these $PM_{2.5}$ particulates are transported from power plant to AQ monitor via downwind currents.

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

In this subsection, we interact monthly, plant-level number of deliveries and coal stockpiles with the logarithm of 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 data on precipitation for each meteorological monitor and then take the inverse-distance weighted average over all four of

Table 4: Effect of Coal Stockpiles on $PM_{2.5}$ Concentration: Upwind vs. Downwind

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	(≤ 25 Miles)	(25-50 Miles)
$CS_{p,t} \times$		
Monitor Downwind from Plant	1.14e-06***	6.53e-08
	(2.17e-07)	(1.23e-07)
Monitor Upwind from Plant	9.18e-07***	9.82e-09
	(2.81e-07)	(1.06e-07)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	(≤ 25 Miles)	(25-50 Miles)
$log(CS_{p,t}) \times$		
Monitor Downwind from Plant	0.0067**	-0.0015
	(0.0031)	(0.0018)
Monitor Upwind from Plant	0.0056*	-0.0017
	(0.0030)	(0.0018)
Observations	47,169	90,535
R-squared	0.849	0.829

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) in levels; the bottom panel of this table regresses coal stockpiles on $PM_{2.5}$ in logs. In both panels, coal stockpiles are 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, 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 and 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% levels respectively. We control for quantity received, number of deliveries, and sum of coal stockpiles from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

these monitors for each month-of-sample.[19] We consider $M = 25$ for the within 25 mile

---

[19] We weight by the inverse of the distance between the AQ monitor and each meteorological monitor; the empirical results are very similar if we take an unweighted average instead.

bandwidth specifications and $M = 50$ for the 25-50 mile bandwidth specifications.

The top panel of Table 5 shows that an increased number of coal deliveries generates less additional $PM_{2.5}$ at higher levels of precipitation. This 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 transporting and delivering coal generates a given amount of $PM_{2.5}$, less of this $PM_{2.5}$ remains in the atmosphere for higher levels of monthly total precipitation. It is especially striking that there is a positive and statistically significant effect of number of deliveries on $PM_{2.5}$ even for plants between 25-50 miles from their AQ monitor when there is zero monthly total precipitation; $PM_{2.5}$ can remain in the air and thus travel greater distances if there is no rainfall to bring these fine particulates to the ground.[20] We draw exactly the same conclusions as described above for the Levels-Levels specification when examining the results from the Log-Log specification presented in the bottom panel of Table 5.

Table 6 displays the results interacting monthly total precipitation with coal stockpiles. As with the number of deliveries, if a one ton increase in coal stockpiles translates into a given level of $PM_{2.5}$, we should expect less of this $PM_{2.5}$ to remain in the atmosphere for higher levels of precipitation due to wet deposition. We find that this intuition holds for both the top panel (Levels-Levels specification) and bottom panel (Log-Log specification) of Table 6. Namely, we see from Table 6 that the interaction term between coal stockpiles and monthly precipitation is statistically significant and negative for both the less than 25 miles and the 25-50 mile distance bandwidths. As before, we see that there is a positive and statistically significant effect of coal stockpiles on $PM_{2.5}$ even for plants between 25-50 miles away from their AQ monitor when there is zero monthly total precipitation; this highlights the importance of rainfall in reducing the level of local ambient $PM_{2.5}$ concentration levels.

4.4 Robustness Checks and Sensitivity Analyses

One potential concern with our empirical specification is that our estimates capture the $PM_{2.5}$ increases from burning coal rather than purchasing and storing it. We consider two

---

[20] We find positive and statistically significant increases in $PM_{2.5}$ from coal procurement (both number of deliveries and coal stockpiles) even for plants between 100 and 200 miles from their AQ monitors if there is no precipitation in the entire month; these results are available upon request.

Table 5: Effect of Number of Deliveries on $PM_{2.5}$ Concentration: Interacted with Precipitation

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t} \times$		
Constant	0.0593***	0.0298***
	(0.0141)	(0.0065)
$log(Precipitation + 1)$	-0.0303***	-0.0217***
	(0.0086)	(0.0042)
Observations	48,521	92,886
R-squared	0.847	0.824
Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1) \times$		
Constant	0.0286***	0.0218***
	(0.0066)	(0.0037)
$log(Precipitation + 1)$	-0.0124***	-0.0135***
	(0.0037)	(0.0020)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding how the link between number of deliveries and $PM_{2.5}$ concentration levels varies with precipitation. The top panel of this table regresses number of deliveries on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses number of deliveries on $PM_{2.5}$ in logs. In both panels, number of deliveries 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 and 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 quantity received, coal stockpiles, and sum of deliveries from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

robustness checks in order to rule out this concern. First, we include more flexible controls for the monthly total thermal generation at each coal-fired power plant; in particular, we control for linear, quadratic, and cubic terms for monthly thermal generation as well

Table 6: Effect of Coal Stockpiles on $PM_{2.5}$ Concentration: Interacted with Precipitation

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t} \times$		
Constant	1.65e-06***	4.59e-07***
	(2.59e-07)	(1.02e-07)
$log(Precipitation + 1)$	-4.67e-07***	-3.07e-07***
	(1.29e-07)	(5.57e-08)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t}) \times$		
Constant	0.0116***	0.0045***
	(0.0030)	(0.0016)
$log(Precipitation + 1)$	-0.0048***	-0.0048***
	(0.0009)	(0.0004)
Observations	47,169	90,535
R-squared	0.850	0.830

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) in levels; the bottom panel of this table regresses coal stockpiles on $PM_{2.5}$ in logs. In both panels, coal stockpiles are 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 and 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% levels respectively. We control for quantity received, number of deliveries, and sum of coal stockpiles from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

as logged monthly thermal generation. Our empirical results including these controls (presented in Appendix Tables A.4 and A.5) 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 procurement (read: coal stockpiles and number of deliveries).

We also consider carbon monoxide (CO) as our dependent variable rather than $PM_{2.5}$. Combustion of any fuel (be it coal, diesel, or gasoline) emits both $PM_{2.5}$ and CO. Thus, we would expect to see a positive association between coal procurement and CO if our empirical findings are due to combustion-based sources such as the burning of coal or the burning of diesel or gasoline from the trains, barges and trucks carrying coal. Instead, we find in Table 7 that there is *no* statistically significant link between changes in number of deliveries and local $PM_{2.5}$ concentration levels for either the within 25 miles or the 25-50 mile distance bandwidths. Similarly, Table 8 demonstrates that there is *not* a statistically significant, positive association between coal stockpiles and carbon monoxide concentration levels.[21] Tables 7 and 8 provide strong evidence that our effect of coal stockpiles and number of coal deliveries on $PM_{2.5}$ is not coming from any alternative source based on combustion. This rules out many of the potential confounding sources of our findings, including the burning of coal or the burning of fuel by trains, barges, or trucks delivering coal.

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-ran 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 very similar results for both of these alternative weighting schemes; these empirical results are available upon request.

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

The previous section estimated the impact of coal procurement on $PM_{2.5}$. This section quantifies the impact of $PM_{2.5}$ concentration levels on mortality rates using coal stockpiles and number of deliveries as instruments for $PM_{2.5}$. We first describe our data on mortality rates. We next specify the ordinary least squares (OLS) and instrumental

---

[21] The coefficient estimate on coal stockpiles is statistically significant at the 10% level for the 25-50 mile bandwidth for the Levels-Levels specification (top panel, Column 2). However, this coefficient estimate is negative rather than positive (increased coal stockpiles lead to *lower* CO) and its magnitude is not economically significant.

Table 7: Overall Effect of Number of Deliveries on CO Concentration

Dependent Variable: $CO_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t}$	-0.0004	0.0006
	(0.0019)	(0.0010)
Observations	14,190	27,461
R-squared	0.865	0.809

Dependent Variable: $log(CO_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1)$	-0.0003	-0.0032
	(0.0045)	(0.0054)
Observations	13,812	26,709
R-squared	0.866	0.824

Notes: This table presents the regression results regarding the link between number of deliveries and carbon monoxide (CO) concentration levels. The top panel of this table regresses number of deliveries on carbon monoxide (in parts per million) in levels; the bottom panel of this table regresses number of deliveries on carbon monoxide 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 and 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 quantity received, coal stockpiles, and sum of deliveries from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

variables (IV) regression frameworks used to measure the impact of $PM_{2.5}$ on mortality rates due to different causes. 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 using monthly plant-level coal stockpiles and number of deliveries. Among other results, we find that a 10% increase in $PM_{2.5}$ leads to a 1.1% (6.6%) increase in average overall adult (infant) mortality rates.

One potential criticism of our coal procurement instruments is that power plants respond to electricity demand, and electricity demand is correlated with local economic outcomes and therefore local health outcomes. However, coal-fired power plants largely

Table 8: Overall Effect of Coal Stockpiles on CO Concentration

Dependent Variable: $CO_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$CS_{p,t}$	1.26e-08	-5.23e-08*
	(3.60e-08)	(2.81e-08)
Observations	14,190	27,461
R-squared	0.865	0.810

Dependent Variable: $log(CO_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t} + 1)$	-0.0002	-0.0023
	(0.0041)	(0.0030)
Observations	13,812	26,709
R-squared	0.866	0.824

Notes: This table presents the regression results regarding the link between coal stockpiles and carbon monoxide (CO) concentration levels. The top panel of this table regresses coal stockpiles (in tons) on carbon monoxide (in parts per million) in levels; the bottom panel of this table regresses coal stockpiles on carbon monoxide 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 and 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% levels respectively. We control for quantity received, number of deliveries, and sum of coal stockpiles from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.

run continuously during our sample period, with little month-to-month variation in their electricity generation. Moreover, the electricity produced by power plants goes into the electricity transmission grid; this electricity generation is distributed regionally. Hence, coal-fired power plants typically do not respond to month-to-month variation in local electricity demand. This in turn makes it unlikely that power plants' coal purchase and storage behavior affects local health outcomes through the channel of local economic outcomes correlated with electricity demand. As such, we argue that this setting provides a novel strategy for identifying the causal relationship between $PM_{2.5}$ and mortality rates.

5.1 Data Sources For Mortality Rate Regressions

We collect monthly, county-level total number of deaths by category from the Centers for Disease Control and Prevention (CDC); these categories are deaths related to the cardiovascular, respiratory, or nervous systems, deaths due to external causes, and total number of deaths by age group. We present results considering only people who are older than 30 years old ("adults"), as this is the sub-population considered in epidemiological studies such as Krewski et al. (2009). We also provide the empirical results counting deaths of people of all ages in Appendix A; these all-age results are very similar to those presented below for the ages 30+ subpopulation. Finally, our empirical results demonstrate that infants (children ages 0-4) are especially vulnerable to the mortality risk associated with $PM_{2.5}$ increases.

We also use annual, county-level population data by age group from the Survey of Epidemiology and End Results (SEER) collected by the National Bureau of Economic Research (NBER) website.

5.2 Empirical Framework

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

$$log(\frac{Deaths_{c,t}}{Pop_{c,y}}) = \alpha_{c,y} + log(PM2.5_{i,t} + 1)\gamma + X_{i,p,t}\beta + \epsilon_{i,p,t} \tag{2}$$

where c indexes the county where air quality monitor (AQ) 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, nervous, external, total, and infant).

The first-stage regression for our instrumental variables (IV) specification is:

$$log(PM2.5_{i,t} + 1) = \delta_{c,y} + log(CS_{p,t} + 1)\theta_1 + log(ND_{p,t} + 1)\theta_2 + X_{i,p,t}\eta + \epsilon_{i,p,t}$$

In particular, we instrument for $log(PM_{2.5} + 1)$ in Equation 2 with both $log(CS_{p,t})$ and $log(ND_{p,t} + 1)$. As with the OLS framework, we estimate this IV specification separately for each type of mortality rate. Our standard errors are clustered by air quality monitor and we weight by population for both OLS and IV specifications.[22] Finally, we focus on plants within 25 miles of their corresponding air quality monitor, as we found that there's a statistical link between coal procurement and $PM_{2.5}$ for this distance bandwidth but not the 25-50 mile distance bandwidth in Section 4.

5.3 Empirical Results

The top panel of Table 9 lists the empirical results when regressing mortality rate on $PM_{2.5}$ concentrations using the OLS framework described in the previous subsection. We run separate regressions for the mortality rates associated with the cardiovascular system, the respiratory system, the nervous system, all adult deaths, and external causes such as accidents. The summary statistics for the regressions in this subsection are relegated to Appendix Section A. Interpreting the first column of Table 9, we find that a 10% increase in $PM_{2.5}$ concentration levels corresponds to a 0.09% increase in the cardiovascular mortality rate on average. Column 2 indicates that a 10% increase in $PM_{2.5}$ is associated with a 0.16% average increase in the rate of deaths due to the respiratory system. The OLS coefficient estimates in Columns 1 and 2 are statistically significant but economically small. However, Column 3 shows a *negative* and statistically significant association between $PM_{2.5}$ exposure and the nervous system mortality rate, while the coefficient estimate on $log(PM_{2.5}+1)$ is not statistically significant in the total mortality rate regression (Column 4). The mortality rate due to external causes such as accidents should not be related to $PM_{2.5}$ concentration levels; however, the OLS results in Column 6 show a negative and statistically significant association between the external mortality rate and $PM_{2.5}$.

In contrast, the bottom panel of Table 9 provides the empirical results relating $PM_{2.5}$ concentration levels to mortality rates for the instrumental variables (IV) specification. As with the OLS results, we find a positive and statistically significant link between $PM_{2.5}$ concentration levels and both cardiovascular and respiratory mortality rates. However, the estimated coefficients in Columns 1 and 2 are much larger for the IV specification relative to the OLS specification; the IV findings indicate that a 10% increase in $PM_{2.5}$ concentration levels results in a 1.3% (4.7%) average increase in the cardiovascular (res-

---

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

Table 9: $PM_{2.5}$ on Mortality Rate: 25 Mile Bandwidth for Adults

		OLS Specification			
VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.009**	0.016***	-0.018**	0.003	-0.024***
	(0.004)	(0.006)	(0.009)	(0.003)	(0.007)
Observations	42,988	36,961	29,122	44,021	30,255
R^2	0.911	0.717	0.790	0.919	0.753
		IV Specification			
VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.134***	0.469***	0.239**	0.109***	-0.104
	(0.039)	(0.094)	(0.097)	(0.027)	(0.086)
Observations	42,093	36,203	28,558	42,965	29,741
R^2	0.903	0.609	0.769	0.908	0.752

Notes: This table presents the OLS and IV results regarding the link between $PM_{2.5}$ concentration levels and adult (ages 30+) mortality rates. The top panel of this table regresses the log of $PM_{2.5}$ on county/month specific mortality rates, while the bottom panel of this table uses both coal stockpiles and number of deliveries as instruments for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on mortality rates. 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 quantity received and also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. We include county-of-AQ monitor/month-of-sample fixed effects as well. We weight observations by the population of the county where the AQ monitor is located for these regressions. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.6 for the summary statistics associated with these regressions.

piratory) mortality rate. Moreover, the IV coefficient estimates in Columns 3 and 4 are positive and statistically significant as expected; we find a positive link between $PM_{2.5}$ exposure and both nervous system mortality rates as well as total mortality rates. Using the IV framework, there is *not* a statistically significant relationship between the external mortality rate and $PM_{2.5}$ exposure in Column 5; 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, this 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.[23]

---

[23] We obtain similar results if we only use coal stockpiles as an instrument rather than both coal

We focus on the overall mortality rate for the subsample of people over 30 years old because the USEPA typically uses the association between $PM_{2.5}$ levels and overall, post-30 mortality rates reported by Krewski et al. (2009) or Lepeule et al. (2012) in its regulatory impact analyses for $PM_{2.5}$.[24] For example, Lepeule et al. (2012) reports that a 1 $^{ug}/_{m^3}$ increase in $PM_{2.5}$ is associated with a 1.4% increase in overall, post-30 mortality rates. The annual mean $PM_{2.5}$ reading in 2014 was 8.4 $^{ug}/_{m^3}$; at this mean $PM_{2.5}$ level, a 1 $^{ug}/_{m^3}$ increase in $PM_{2.5}$ amounts to a 11.9% increase in $PM_{2.5}$ levels. This suggests an elasticity of $\frac{0.014}{0.119} = 0.12$ based on the coefficient estimate reported by Lepeule et al. (2012). The estimate of the all-cause, post-30 mortality elasticity with respect to $PM_{2.5}$ implied by Lepeule et al. (2012) (0.12) is quite similar to our elasticity estimate (0.11). However, $PM_{2.5}$ levels have fallen since the Lepeule et al. (2012) study was published. For example, the annual average $PM_{2.5}$ level was 12.47 $^{ug}/_{m^3}$ in 2008; a 1 unit change in $PM_{2.5}$ from the 2008 annual average level implies an elasticity of 0.18. However, we obtain an elasticity of $\frac{0.006}{0.119} = 0.05$ if we instead do the same calculation using the results from Krewski et al. (2009). Summarizing, our IV estimate of the all-cause, post-30 mortality elasticity with respect to $PM_{2.5}$ (which in 0.11) is slightly smaller than the Lepeule et al. (2012) estimate (either 0.12 based on the 2014 average or 0.18 based on the 2008 average) but larger than the Krewski et al. (2009) estimate (which is 0.05 based on the 2008 average). The fact that our estimated elasticity of mortality rate with respect to $PM_{2.5}$ is of the same magnitude as these epidemiological studies lends credence to the USEPA analyses which make use of these studies.

Finally, we focus on the effect of $PM_{2.5}$ on the mortality rates for children ages 0 to 4 ("infants") in Table 10. Infants are known to be especially vulnerable to the negative health effects from $PM_{2.5}$. Consistent with this intuition, we see from the IV results in the bottom panel of Table 10 that the average increase in infant mortality from a 1% increase in $PM_{2.5}$ (0.66%) is much larger than the corresponding average increase in adult mortality rates (0.11%). Finally, we see from the top panel of Table 10 that ordinary least squares does not identify a positive link between $PM_{2.5}$ and infant mortality, highlighting both the necessity of our instrumental variables identification strategy based on coal procurement as well as the strength of the statistical link between coal procurement and $PM_{2.5}$ concentration levels.

---

stockpiles and number of coal deliveries (see Appendix Table A.8).

[24] The empirical results for the full population (see Appendix Table A.7) are quite similar to the results described above for the over-30 subpopulation.

Table 10: $PM_{2.5}$ on Infant Mortality Rate: 25 Mile Bandwidth for Ages 0-4

OLS Specification	
VARIABLES	Total Infant MR
$log(PM_{2.5} + 1)$	-0.012
	(0.016)
Observations	11,291
R^2	0.648

IV Specification	
VARIABLES	Total Infant MR
$log(PM_{2.5} + 1)$	0.657***
	(0.030)
Observations	11,085
R^2	0.442

Notes: This table presents the OLS and IV results regarding the link between $PM_{2.5}$ concentration levels and total infant (ages 0-4) mortality rate. The top panel of this table regresses the log of $PM_{2.5}$ on county/month specific total infant mortality rate, while the bottom panel of this table uses both coal stockpiles and number of deliveries as instruments for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on infant mortality rate. 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 quantity received and also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. We include county-of-AQ monitor/month-of-sample fixed effects as well. We weight observations by the population of the county where the AQ monitor is located for these regressions. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.6 for the summary statistics associated with these regressions.

6 Local Environmental Health Costs of Coal Procurement

We estimated the effect of coal procurement on $PM_{2.5}$ in Section 4. The previous section measured the effect of local $PM_{2.5}$ concentration levels on mortality rates. This section combines those two estimates; we translate the increases in $PM_{2.5}$ from coal stockpiles and deliveries into total number of deaths and ultimately calculate the social cost from those deaths in dollars per ton using the Value of Statistical Life (VSL) approach. We also present two alternative methods of calculating social costs as sensitivity analyses. First, we use a concentration-response relationship from the epidemiological literature in order to translate our estimated $PM_{2.5}$ increase from coal procurement into total adult deaths. This alternative approach does not use our estimates of the effect of $PM_{2.5}$ on mortality rates. We also present social costs in life-years lost per ton of coal stockpiled or delivered, allowing the reader to choose their own VSL when calculating social costs in dollars per ton. Summarizing our primary findings, we calculate that a one ton increase in coal stockpiles has local environmental costs of \$182.67 while a one ton increase in the quantity of coal delivered has local environmental costs of \$202.51. Roughly 75% (25%) of these social costs come from adult (infant) mortality due to the greater number of adults exposed to the $PM_{2.5}$ increases relative to infants.

6.1 Translating Partial Effects to Local Environmental Costs: Methodology

In Section 4, we measured the extent to which fugitive and gaseous emissions from coal deliveries and coal stockpiles have an effect on ambient $PM_{2.5}$ levels. We quantify the local health costs of such emissions in this section. We focus exclusively on mortality risk because prior research has shown that the majority of damage from $PM_{2.5}$ exposure is due to elevated mortality risk (EPA (1999); Muller, Mendelsohn and Nordhaus (2011)). This aspect of the analysis relies on plant-level average partial effects linking coal stockpiles and deliveries to $PM_{2.5}$. We denote these partial effects $\Delta PM2.5_p^{CS}$ for coal stockpiles and $\Delta PM2.5_p^{ND}$ for number of deliveries, which we obtain by averaging our air quality monitor/plant/month-of-sample level partial effects from Section 4 over air quality monitors and months-of-sample for each power plant p. We divide the partial effect from an additional delivery by the number of tons in the delivery in order to derive a comparable marginal emissions figure for a ton of coal delivered versus a ton of coal stockpiled. Summarizing, these partial effects translate a ton of coal stored or delivered into $PM_{2.5}$ increases in micrograms per cubic meter. We focus on $\Delta PM2.5_p^{CS}$ in the description below purely for ease of exposition.

We also rely on our estimated impact of $PM_{2.5}$ on average adult and infant mortality rates from Section 5; we denote these impacts ΔMR^{adult} and ΔMR^{infant}. ΔMR^{adult} and ΔMR^{infant} are based on estimated semi-elasticities converting $PM_{2.5}$ in micrograms per cubic meter into percentage changes in average adult and infant mortality rates. The following equation formally describes how we calculate the total number of adult deaths from coal procurement induced increases in $PM_{2.5}$ concentration levels:

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

where $MR_{c,a}$ is the mortality rate of age group a in the county c where power plant p is located. We collect data on baseline mortality rates from the Centers for Disease Control and Prevention (CDC) for the year 2011; these mortality rate data are provided in 13 age groups. For $POP_{c,a}$, we use 2010 U.S. census block level population data in order to find the total number of people in age group a living in census blocks in county c whose centroid is within 25 miles of power plant p.[25] We partially aggregate the Census population data in order to match the age group designations used by the CDC for the mortality rate data.

The result of multiplying population times mortality rate gives us total number of deaths for each age group; we use the partial effects we estimated in Section 4 ($\Delta PM2.5_p^{CS}$) as well as the semi-elasticities we estimated in Section 5 in order to calculate the total number of adult deaths (M_p) attributable to increased $PM_{2.5}$ exposure stemming from coal stockpiles.

The equation describing how to calculate total number of infant deaths is:

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

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.

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

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

---

[25] In practice, 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 (2010 Census).

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 the same as defined above for Equation 3.

The next step is quantifying the cost in dollars of these deaths. We employ the Value of a Statistical Life (VSL) methodology for this conversion (Viscusi and Aldy, 2003), using a VSL of \$9.85 million based on the regulatory impact analyses for air pollution conducted by the USEPA. The monetary marginal damage due to emissions from an additional ton of coal stockpiled at plant p is given by:

$$MD_p^{CS} = M_p^{CS} VSL$$

We also calculate the life-years lost due to coal procurement induced increases in $PM_{2.5}$ concentration levels. To do this, we employ data on conditional life expectancy by age group from the National Vital Statistics Reports (U.S. 2011 Life Tables). In our VSL-based approach, we assume that an additional death in a given age group due to increases in $PM_{2.5}$ results in the loss of full life expectancy for that age group. In contrast, life-years lost are 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 at a given location 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.

6.2 Translating Partial Effects to Local Environmental Costs: Empirical Results

We calculate the local environmental health costs due to increases in $PM_{2.5}$ emissions from coal stockpiles and number of deliveries using the methodology discussed above. We report social cost estimates based on the Log-Log specifications that account for the wind direction between air quality monitor and coal-fired power plant;[26] we focus on the partial effects for plants within 25 miles of their corresponding air quality (AQ) monitor. Table

---

[26] Our social cost estimates are similar in magnitude if we instead use the Levels-Levels specifications (Appendix Table A.9) or if we instead use the Log-Log specifications that fail to account for wind direction (Appendix Table A.10).

11 shows the results from these social cost calculations for the *median* plant; we present the median social cost rather than the average social cost as the plant-level distribution of these social costs is right-skewed. Column 1 of this table uses the link between $PM_{2.5}$ and mortality estimated in Krewski et al. (2009) for adults (ages 30+), Column 2 of this table uses our estimate of the effect of $PM_{2.5}$ on overall adult mortality rates, and Column 3 uses our estimate of the effect of $PM_{2.5}$ on overall infant (ages 0-4) mortality rates.

We see from Column 2 of the top panel of Table 11 that the monetary damage per ton of coal stockpiled (delivered) due to increased adult mortality rates is \$141.17 (\$159.73). The additional social cost from coal stockpiles (number of deliveries) due to infant mortality is \$41.50 (\$42.78) per ton. Though a 1% increase in $PM_{2.5}$ results in a higher percentage increase in infant mortality rates relative to adult mortality rates, the health cost for infants is smaller than the health cost for adults because the population of adults is substantially larger than the population of infants. Our social cost estimates for adults are substantially larger when we use our estimated effect of $PM_{2.5}$ on adult mortality rate rather than the effect estimated in Krewski et al. (2009), which makes sense given that we estimate a larger effect of $PM_{2.5}$ on adult mortality than Krewski et al. (2009). However, our social cost estimates are slightly smaller than those found in Lepeule et al. (2012), indicating that our estimates are not unreasonably large when compared to the epidemiological literature. This highlights the importance of the relationship between $PM_{2.5}$ and mortality rates in quantifying the local environmental costs of air pollution; even small differences in the $PM_{2.5}$/mortality rate relationship used to calculate social costs result in large differences in these social costs.

These per-ton local environmental damages from coal stockpiles and number of deliveries are quite large given that the average plant pays roughly \$48 per ton for coal, stockpiles 212,781.6 tons of coal and has 106,235 tons of coal delivered to it. However, the middle panel of Table 11 presents our social 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 per-MWh social costs summing across both adults and infants from storing (delivering) coal are \$95.51 (\$103.12). As a basis for comparison, Levy, Baxter and Schwartz (2009) reports a median social cost (across roughly 400 plants) of about \$140 per MWh for $PM_{2.5}$ increases due to the combustion of coal; this translates to roughly \$230 per MWh if we adjust for the value of statistical life (VSL) used in our paper versus the VSL used in Levy, Baxter and Schwartz (2009). Burning coal emits significantly more $PM_{2.5}$ than storing and transporting coal; it is thus comforting that our local environmental health cost estimates for coal purchased and stored are substantially smaller than the local environmental health costs estimated for coal burned.

Table 11: Local Environmental Costs of Coal Procurement

Air Pollution Costs: Dollars Per Ton

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	79.90	141.17	41.50
$ND_{p,t}$	90.41	159.73	42.78

Air Pollution Costs: Dollars Per MWh

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	42.16	74.51	21.00
$ND_{p,t}$	45.08	79.64	23.48

Air Pollution Costs: Life-Years Per 10,000 Tons

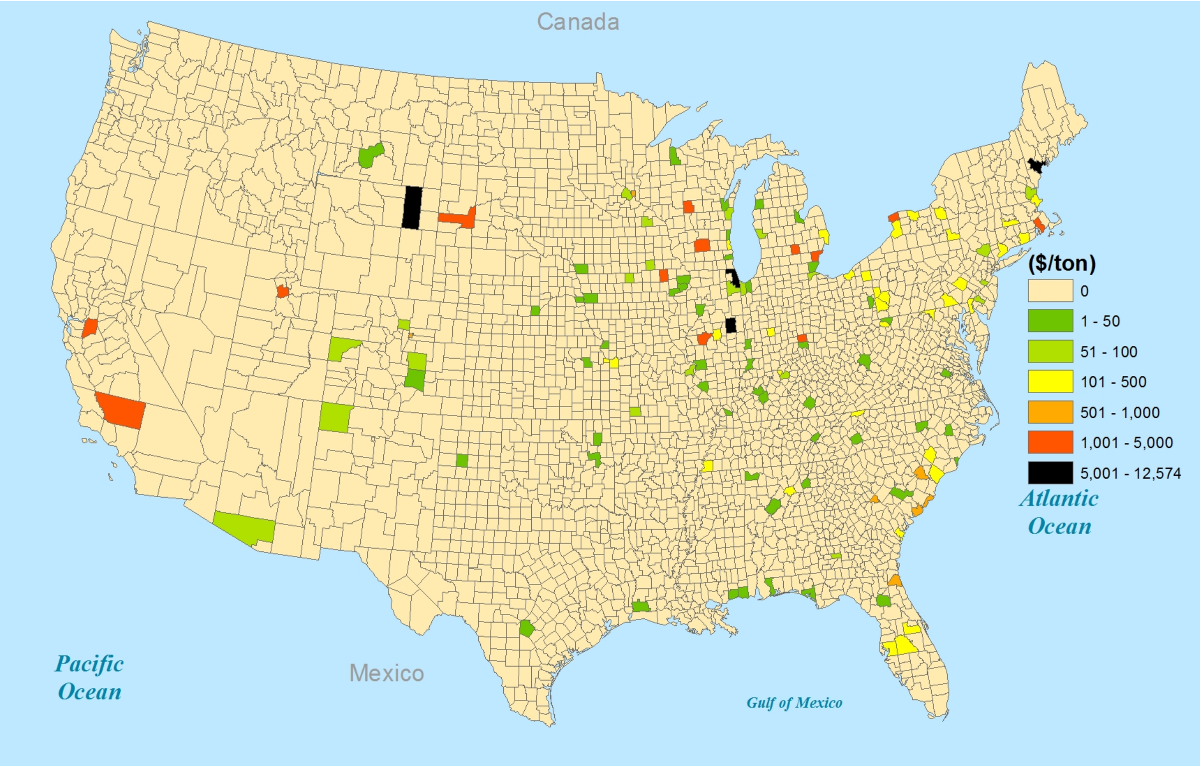
	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	1.08	1.90	3.32
$ND_{p,t}$	1.27	2.24	3.42

Notes: This table presents the local environmental costs of coal stockpiles and number of deliveries for the *median* plant in our sample. We use the partial effects of coal stockpiles and number of deliveries on $PM_{2.5}$ based on the Log-Log specification accounting for the relative bearing between power plant and air quality monitor from Section 4; we average over air quality monitors and months-of-sample to obtain plant-specific partial effects. Column 1 of each panel uses the link between $PM_{2.5}$ estimated in Krewski et al. (2009) for adults (ages 30+), Column 2 of each panel uses our own estimated link between $PM_{2.5}$ and adult mortality from Section 5, and Column 3 uses our estimated link between $PM_{2.5}$ and infant (ages 0-4) mortality. The top panel of this table presents local environmental costs per ton of coal stockpiled and delivered. The middle panel of this table presents social costs per MWh-equivalent of coal stockpiled and delivered; we convert tons of coal to MWh of electricity by taking plant-specific total number of tons of coal burned and dividing by total electricity generated. We use a value of statistical life of 9.85 million dollars to quantify the costs in dollars of the increased mortality from coal procurement based $PM_{2.5}$ for these top two panels. Finally, the bottom panel presents social costs in life-years per 10,000 tons stockpiled and delivered.

We also see from Table 11 that the local environmental damages from storing coal are smaller than the local environmental costs of delivering coal. Our effect of number of deliveries on $PM_{2.5}$ primarily captures the emissions from coal handling and preparation at the power plant; the effect of coal stockpiles on $PM_{2.5}$ is due to the wind erosion and gaseous emissions from stationary stockpiles. Thus, our empirical findings are consistent with the notion that the air pollution costs from displacing coal should be larger than the air pollution costs from a stationary coal pile.

The bottom panel of Table 11 presents our local environmental cost estimates in life-years per-10,000 tons. Focusing on Column 2, we see that a 10,000 ton increase in coal stockpiled (delivered) results in roughly 1.90 (2.24) life-years lost for adults.[27] Table 11 tells us that a 10,000 ton increase in coal stockpiles results in 0.1433 ($= \frac{141.17}{9850000} \times 10,000$) deaths; this means that roughly 13.26 ($= \frac{1.90}{0.1433}$) life-years lost equals a death due to the increased $PM_{2.5}$ exposure from coal stockpiles. Thus, the reader can pick either a different value of statistical life (VSL) or a different conversion rate between life-years lost and deaths in order to scale up or scale down our local environmental cost estimates.

Figure 1: Damages Per Ton of Coal Stockpiled Across the U.S.



Notes: This figure displays the geographic dispersion across the United States of our plant-specific estimates of the local environmental damages from increased $PM_{2.5}$ exposure due to a one ton increase in coal stockpiles. We use the effect of coal stockpiles on $PM_{2.5}$ based on the Log-Log specification accounting for the relative bearing between power plant and air quality monitor from Section 4; we average over air quality monitors and months-of-sample to obtain a plant-specific partial effect. We calculate damages based on our own estimated link between $PM_{2.5}$ and adult mortality from Section 5. We use a value of statistical life of 9.85 million dollars to quantify the costs in dollars of the increased mortality from coal procurement based $PM_{2.5}$.

Figure 1 displays the geographic dispersion across the United States of our plant-

---

[27] We draw exactly the same conclusions if we present our findings on a per-GWh basis rather than a per-10,000 ton basis (see Appendix Table A.11).

Table 12: 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 within their borders. The tract-level proportion of residents with a bachelor's degree, per-capita income, and median family income are collected from the 2000 U.S. Census. The locations of coal-fired power plants are collected from the eGrid database constructed by the USEPA.

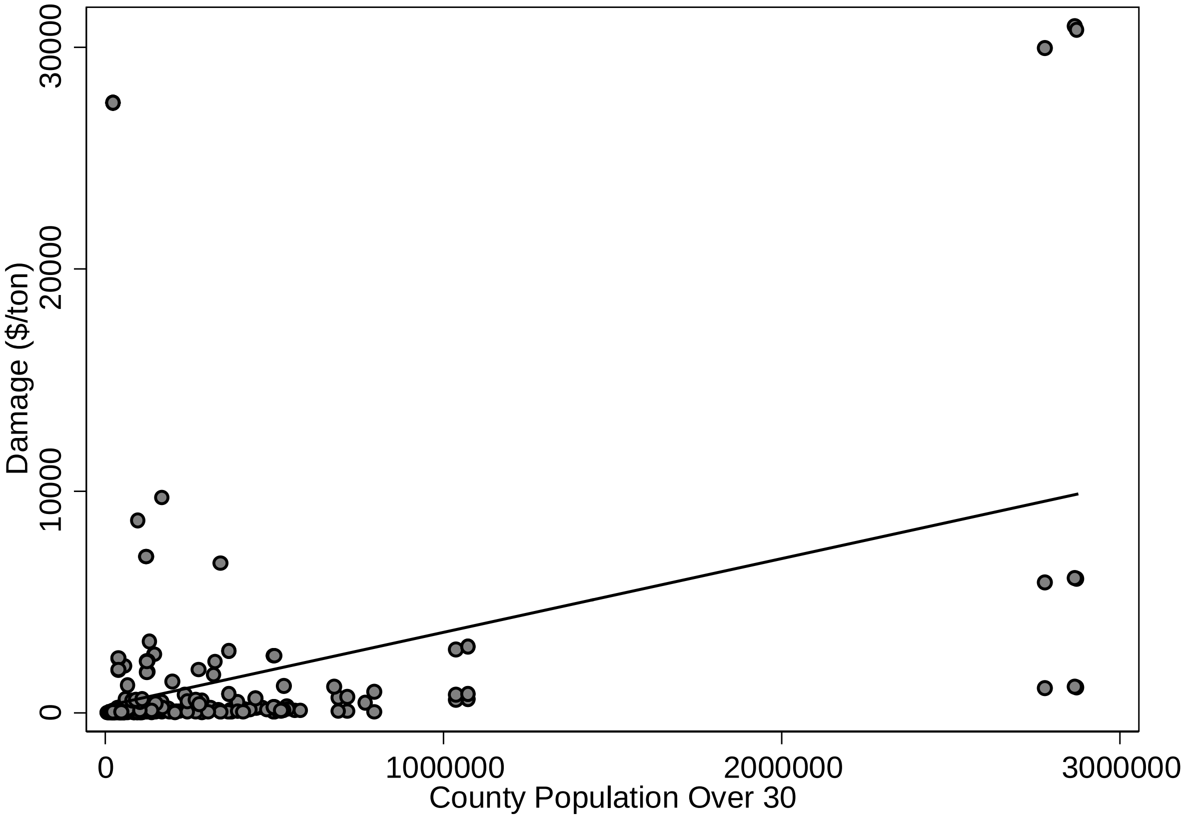
specific estimates of the local environmental damages from increased $PM_{2.5}$ exposure due to a one ton increase in coal stockpiles. This figure highlights that the environmental costs of coal stockpiling are highly local; most counties do not have a power plant located in them, and thus do not incur any local environmental costs from plants stockpiling coal. Figure 1 also shows that local environmental damages are higher in areas with higher populations; for example, we see particularly high health costs associated with coal stockpiles in the Northeastern region of the United States. Figure 2 demonstrates directly that our estimates of the per-ton local environmental costs from coal stockpiling increase with total county-level adult population. This figure also shows that the distribution of local environmental health costs are right-skewed; for example, we estimate extremely large air pollution costs per ton of coal stockpiled in Cook County (which contains Chicago).

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 12. This evidence suggests that the highly localized environmental impacts of coal transportation and storage are borne disproportionately by economically disadvantaged communities.[28]

---

[28] Our empirical evidence is corroborated by Davis (2011), which finds that neighborhoods near fossil-fuel fired power plants have lower average household incomes and educational attainment using restricted-access census microdata from 1990 and 2000.

Figure 2: Damages Per Ton Stockpiled versus Total Population



Notes: This scatterplot shows the relationship between county-level adult (ages 30+) population and our plant-specific estimates of the local environmental damages from increased $PM_{2.5}$ exposure due to 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. We use the effect of coal stockpiles on $PM_{2.5}$ based on the Log-Log specification accounting for the relative bearing between power plant and air quality monitor from Section 4; we average over air quality monitors and months-of-sample to obtain a plant-specific partial effect. Damages are calculated based on our own estimated link between $PM_{2.5}$ and adult mortality from Section 5. We use a value of statistical life of 9.85 million dollars to quantify the costs in dollars of the increased mortality from coal procurement based $PM_{2.5}$.

7 Conclusion and Policy Implications

Reliance on coal as an energy source is known to have environmental consequences all along the supply chain from coal produced at mines to coal burned at power plants. This paper uncovers a new dimension of coal use, the coal purchase and storage behavior of U.S. power plants, that requires environmental policy intervention. We first quantify the effect of coal deliveries and coal stockpiles on local ambient air pollution. Next, this variation in local air pollution from coal procurement is used to identify the link between $PM_{2.5}$ and mortality rates. Finally, we combine our estimates with the methodology utilized by the USEPA in their regulatory impact analyses in order to calculate the per-ton social costs from plants' coal purchase and storage behavior.

This paper demonstrates that the storage and conveyance of coal contributes to ambient $PM_{2.5}$; thus, local, state, and federal air quality regulators have a new policy lever they can use to comply with the National Ambient Air Quality Standards. Our analysis also aids in the design of policies that manage the local environmental consequences of any large scale coal transport or storage decision. As will be demonstrated below, this is especially important for coal export terminals given that these terminals are located in densely populated areas and transport and store large quantities of coal. Finally, our results can be used to assess the local environmental consequences of even purely economic policies. As an example of this, power plants subject to output price regulation hold larger quantities of coal on site on average relative to plants facing electricity market mechanisms (Jha, 2017); our findings indicate that this results in higher local $PM_{2.5}$ concentration levels around regulated power plants. We argue that this environmental cost should be included in assessments of utility rate regulation.

Summarizing our primary results, we find that a 10% increase in coal stockpiles (number of deliveries) results in a 0.06% (0.12%) increase in $PM_{2.5}$ concentration levels on average. These $PM_{2.5}$ concentration increases are more severe for: 1) local populations downwind from coal-fired power plants, and 2) areas/months with less precipitation. However, the combustion of any fuel, whether it's coal burned by the power plant or the fuel burned by the trains, barges, and trucks carrying coal, emits both $PM_{2.5}$ and carbon monoxide (CO). We show that there is no statistical effect of either coal stockpiles or number of deliveries on carbon monoxide (CO), providing strong evidence that our estimated increase in $PM_{2.5}$ from coal procurement is not driven by combustion-based sources of $PM_{2.5}$.

We next estimate the effect of $PM_{2.5}$ on mortality rates, separately for mortality rates based on cardiovascular system deaths, respiratory system deaths, nervous system deaths, overall adult (people ages 30+) deaths, overall infant (children ages 0-4) deaths, and deaths from external causes such as accidents. Ordinary least squares (OLS) regressions of $PM_{2.5}$ exposure on mortality rates yield economically small and sometimes negative coefficient estimates. In contrast, we see positive, statistically significant, and economically significant effects of $PM_{2.5}$ exposure on mortality rates (excepting the external cause mortality rate) if we instrument for $PM_{2.5}$ with coal stockpiles and number of coal deliveries. The positive effect of $PM_{2.5}$ exposure on mortality rates is especially large for infants. Our elasticity estimate of the overall adult mortality rate with respect to $PM_{2.5}$ based on our new identification strategy is roughly in line with those used by policymakers such as the USEPA to assess the environmental costs of $PM_{2.5}$ exposure. Thus, we provide evidence that these policy analyses are not significantly over-stating or under-stating the effects of $PM_{2.5}$ on mortality rates.

Finally, we combine our estimates of the average increase in $PM_{2.5}$ from coal procurement and the average increase in mortality rates from $PM_{2.5}$ in order to quantify the social costs of coal stockpiles and deliveries. In particular, we calculate the total number of adult and infant deaths from coal procurement based $PM_{2.5}$ emissions and use the Value of a Statistical Life approach to monetize this mortality. Our results indicate that health costs are roughly \$183 (\$203) per ton of coal stockpiled (delivered). These local environmental costs are sizable given that U.S. coal-fired power plants pay roughly \$48 per ton for coal on average. However, our estimates are not unreasonably large given that the *per-MWh* local damages from delivering and storing coal (\$95.51 per MWh-equivalent for storing coal and \$103.12 per MWh-equivalent for delivering coal) are substantially smaller than the local environmental damages from burning coal (roughly \$230 per MWh from Levy, Baxter and Schwartz (2009)).

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. Moreover, a policy requiring that coal piles be covered does not require significant coordination across jurisdictions because the environmental impacts of coal purchase and storage behavior are highly local; for example, the vast majority of the local environmental costs from a plant's coal storage and handling behavior are incurred in the jurisdiction where this plant is located. This stands in direct contrast to proposed policy interventions for global pollutants such as the CO_2 emissions generated when coal is burned. Thus, local environmental policies designed to mitigate the $PM_{2.5}$ emissions from coal procurement are likely easier to enact and implement relative to policies designed to mitigate global or regional pollutants such as CO_2 or SO_2. Finally policies targeting the $PM_{2.5}$ emissions from coal procurement may be a low-cost method to comply with the National Ambient Air Quality Standards.

The local environmental health costs from coal handling and storage apply more broadly than our examination of U.S. coal-fired power plants. The terminals exporting coal, the railroads used to transport coal, and the mines that produce coal are also likely to be associated with increased $PM_{2.5}$ exposure from dust emissions. To demonstrate this, we combine our estimates of the effect of coal storage on local $PM_{2.5}$ with data from SNL Financial on the location of coal export terminals and the monthly quantity of coal stored at the terminals in order to provide a rough estimate of the social costs from these facilities. Using the methodology described in the previous section, we find local environmental costs of \$325 per ton of coal stored at a terminal. This value is considerably larger than the effect from coal stockpiles at power plants primarily 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 hold an average of roughly 29,900 tons of coal. Our back-of-the-envelope estimate of the monthly damage from export terminals due to $PM_{2.5}$ exposure is \$350 million which amounts to \$4.2 billion annually. This exercise demonstrates that the adverse impacts on community health from coal export terminals are likely to be significant.

Finally, even purely economic policies may have environmental consequences; thus, the environmental impacts of different policy options must be included in any cost-benefit analysis. For example, 75% of U.S. coal-fired electricity generation was produced under output price regulation in 2015. Jha (2017) finds within a matched difference-in-differences framework that, relative to plants facing electricity market mechanisms, regulated plants: 1) hold roughly 13% more coal stockpiles controlling for coal consumption, and 2) receive roughly 13% more deliveries per month controlling for quantity purchased.[29] Our paper demonstrates that these regulatory distortions to coal procurement behavior have environmental costs in addition to the economic costs documented in Jha (2017). Namely, we estimate that a one ton increase in coal stockpiles has an air pollution cost of \$182.67, while coal deliveries generate social costs of \$202.51 per ton. For our 2002-2012 sample of plants within 25 miles of an air quality monitor, regulated plants on average hold 261,999.8 tons of coal and receive 135,436.5 tons of coal per month; there are 193 regulated plants in our data. Thus, the monthly air pollution costs across all regulated plants of the 13% increase in coal stockpiles due to the structure of output price regulation is approximately \$1.2 billion ($= 182.67 \times 193 \times 0.13 \times 261,999.8$). The equivalent calculation for coal deliveries yields an estimate of monthly damages of roughly \$700 million ($= 202.51 \times 193 \times 0.13 \times 135,436.5$). Annualizing this figure, the total damages from coal stockpiles (coal deliveries) are \$14 billion (\$8 billion). 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 air pollution damages from increases in coal procurement at power generation facilities due to output price regulation range between 10% and 13% of the aggregate damages from combustion at these power plants.

References

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

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.

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.

---

[29] Jha (2017) argues that these distortions to coal procurement occur due to the structure of output price regulation, which provides a working capital allowance for coal stockpiles held on-site and partially passes through coal purchase costs into the output price paid by consumers.

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."

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*.

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. "Dynamic Regulatory Distortions: Coal Procurement at US Power Plants." Working Paper.

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.

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.

NRC, NAS. 2010. *Hidden Costs of Energy: unpriced consequences of energy production and use.* National Academy of Sciences, National Research Council. National Academy Press, Washington, D.C., USA.

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 Tables

List of Figures

List of Appendix Tables

List of Appendix Figures

A Additional Tables and Figures

A.1 Additional Tables and Figures: Coal Procurement 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 $^{ug}/_{m^3}$)	47,169	11.760	3.934
Stockpiles (in tons)	47,169	212,781.6	244,681.8
Number of Deliveries	47,169	2.899	3.628
Quantity Received (in tons)	47,169	106,235	123,238.9
Distance (in miles)	47,169	8.678	6.953
Relative Angle (in degrees)	47,169	87.915	55.813
Precipitation (in inches)	47,169	3.442	3.838
Dry Bulb Temp. (in degrees Fahrenheit)	47,169	53.334	17.491
Dew Point Temp. (in degrees Fahrenheit)	47,169	42.612	16.773
Wet Bulb Temp. (in degrees Fahrenheit)	47,169	48.036	15.873
Relative Humidity (percentage)	47,169	70.448	7.780
Station Pressure (in hundredths of an inch)	47,169	29.256	1.012
5% Wind Speed (in miles-per-hour)	47,169	0.328	0.842
95% Wind Speed (in miles-per-hour)	47,169	15.874	3.541
Num. of Plants per AQ Monitor	29,453	1.602	0.841

Notes: This table presents the summary statistics for the regressions regarding the effect of coal procurement (coal stockpiles and number of coal deliveries) 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 AQ monitor and power plant for these summary statistics. We restrict our sample to plants within 25 miles of their air quality (AQ) monitor for this table.

Table A.2: Summary Statistics: Plants between 25 and 50 Miles from AQ Monitor

Variable	Obs	Mean	Std. Dev.
$PM_{2.5}$ (in $^{ug}/_{m^3}$)	90,535	12.412	4.150
Stockpiles (in tons)	90,535	327,299	402,177.1
Number of Deliveries	90,535	4.529	5.205
Quantity Received (in tons)	90,535	177,000.3	213,660.9
Distance (in miles)	90,535	36.528	7.276
Relative Angle (in degrees)	90,535	88.438	53.912
Precipitation (in inches)	90,535	3.539	4.616
Dry Bulb Temp. (in degrees Fahrenheit)	90,535	54.554	16.876
Dew Point Temp. (in degrees Fahrenheit)	90,535	43.920	16.389
Wet Bulb Temp. (in degrees Fahrenheit)	90,535	49.243	15.416
Relative Humidity (percentage)	90,535	70.619	6.860
Station Pressure (in hundredths of an inch)	90,535	29.322	0.744
5% Wind Speed (in miles-per-hour)	90,535	0.311	0.705
95% Wind Speed (in miles-per-hour)	90,535	14.907	3.345
Num. of Plants per AQ Monitor	45,283	1.999	1.412

Notes: This table presents the summary statistics for the regressions regarding the effect of coal procurement (coal stockpiles and number of coal deliveries) 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 AQ monitor and power plant for these summary statistics. We restrict our sample to plants between 25 and 50 miles of their air quality (AQ) monitor for this table.

Table A.3: Number of Deliveries on $PM_{2.5}$: No Quantity Received Control

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t}$	0.0146*	-0.0062
	(0.0077)	(0.0049)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1)$	0.0069*	-0.0053**
	(0.0041)	(0.0026)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding the link between number of deliveries and $PM_{2.5}$ concentration levels not controlling for the total monthly quantity of coal delivered to each power plant. The top panel of this table regresses number of deliveries on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses number of deliveries 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 and 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 coal stockpiles, and sum of deliveries from other plants. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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 Number of Deliveries on $PM_{2.5}$ Concentration: Additional Generation Controls

Dependent Variable: $PM2.5_{i,t}$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$ND_{p,t}$	0.0178**	0.0012
	(0.0074)	(0.0044)
Observations	48,521	92,886
R-squared	0.847	0.824
Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(ND_{p,t} + 1)$	0.0123**	0.0034
	(0.0048)	(0.0030)
Observations	47,169	90,535
R-squared	0.849	0.829

Standard errors clustered by air quality monitor in parentheses
Facility Code/AQ Monitor Fixed Effects Included
AQ County/Month-of-Sample Fixed Effects Included
Sum of Deliveries from Other Plants Included
Meteorological and EPA Emissions Controls Included
Coal Quality, Quantity Received, and Thermal Generation Controls Included
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the regression results regarding the link between number of deliveries and $PM_{2.5}$ concentration levels controlling flexibly for total monthly thermal generation. The top panel of this table regresses number of deliveries on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses number of deliveries 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 and 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 monthly thermal generation as well as logged monthly thermal generation as controls in these regressions. We also control for quantity received, coal stockpiles, and sum of deliveries from other plants. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.5: 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)
$CS_{p,t}$	1.02e-06***	3.13e-08
	(2.05e-07)	(8.32e-08)
Observations	48,521	92,886
R-squared	0.847	0.824

Dependent Variable: $log(PM2.5_{i,t}) + 1$		
Dist. Bandwidth	($\leq$ 25 Miles)	(25-50 Miles)
$log(CS_{p,t})$	0.0061**	-0.0016
	(0.0029)	(0.0016)
Observations	47,169	90,535
R-squared	0.849	0.829

Notes: This table presents the regression results regarding the link between coal stockpiles and $PM_{2.5}$ concentration levels controlling flexibly for total monthly thermal generation. The top panel of this table regresses coal stockpiles (in tons) on $PM_{2.5}$ (in micrograms per cubic meter) in levels; the bottom panel of this table regresses 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 and 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 monthly thermal generation as well as logged monthly thermal generation as controls in these regressions. We also control for quantity received, number of deliveries, and sum of deliveries from other plants. Finally, we include facility code/AQ monitor fixed effects and county-of-AQ-monitor/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.6: Mortality Rate Summary Statistics: 25 Mile Bandwidth

Variable	Obs	Mean	Std. Dev.
Mortality Rate: Cardio (Per 1,000 People)	42,093	0.413	0.106
Mortality Rate: Respiratory (Per 1,000 People)	36,203	0.106	0.031
Mortality Rate: Nervous (Per 1,000 People)	28,558	0.054	0.018
Mortality Rate: Total (Per 1,000 People)	42,965	1.167	0.223
Mortality Rate: External (Per 1,000 People)	29,741	0.063	0.021
Population	42,965	1,282,459	1,119,970
$PM_{2.5}$ (in $^{ug}/_{m^3}$)	42,965	12.704	3.959
Stockpiles (in Tons)	42,965	173,513.4	210,923.6
Number of Deliveries	42,965	3.479	3.114
Infant Mortality Rate: Total (Per 1,000 People)	11,085	0.021	0.006

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 4. 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 total mortality rate for infants (ages 0-4).

Table A.7: $PM_{2.5}$ on Mortality Rate: All Ages/25 Mile Bandwidth

VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.0088**	0.0138**	-0.0163*	0.0023	-0.0257***
	(0.0036)	(0.0055)	(0.0086)	(0.0028)	(0.0065)
Observations	42,996	37,054	29,489	44,027	33,708
R^2	0.916	0.735	0.800	0.926	0.773

Standard errors clustered by air quality monitor in parentheses
County of AQ Monitor/Year-of-Sample Fixed Effects Included
Meteorological and EPA Emissions Controls Included
Coal Quality, Quantity Received, and Thermal Generation Controls Included
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.134***	0.461***	0.196**	0.117***	-0.0739
	(0.0377)	(0.0887)	(0.0850)	(0.0286)	(0.0648)
Observations	42,101	36,293	28,921	42,965	33,079
R^2	0.908	0.636	0.787	0.914	0.773

Standard errors clustered by air quality monitor in parentheses
County of AQ Monitor/Year-of-Sample Fixed Effects Included
Meteorological and EPA Emissions Controls Included
Coal Quality, Quantity Received, and Thermal Generation Controls Included
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table presents the OLS and IV results regarding the link between $PM_{2.5}$ concentration levels and mortality rates considering people of all ages. The top panel of this table regresses the log of $PM_{2.5}$ on county/month specific mortality rates, while the bottom panel of this table uses both coal stockpiles and number of deliveries as instruments for $PM_{2.5}$ when assessing the impact of $PM_{2.5}$ on mortality rates. 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 quantity received and also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. We include county-of-AQ-monitor/month-of-sample fixed effects as well. We weight observations by the population of the county where the AQ monitor is located for these regressions. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.6 for the summary statistics associated with these regressions.

Table A.8: $PM_{2.5}$ on Adult Mortality Rates Using only $CS_{p,t}$ as an Instrument

VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.0086**	0.0156***	-0.0179**	0.0027	-0.0234***
	(0.0036)	(0.0055)	(0.0088)	(0.0027)	(0.0070)
Observations	42,988	36,961	29,122	44,021	30,255
R^2	0.911	0.717	0.790	0.919	0.753

VARIABLES	Cardio MR	Resp. MR	Nervous MR	Total MR	External MR
$log(PM_{2.5}+1)$	0.230***	0.473***	0.0721	0.182***	0.0088
	(0.0465)	(0.0984)	(0.0875)	(0.0359)	(0.0811)
Observations	42,093	36,203	28,558	42,965	29,741
R^2	0.885	0.607	0.788	0.886	0.754

Notes: This table presents the OLS and IV results regarding the link between $PM_{2.5}$ concentration levels and adult (ages 30+) mortality rates. The top panel of this table regresses the log of $PM_{2.5}$ on county/month specific mortality rates, 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. 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 number of deliveries, and quantity received. We also include meteorological controls, EPA emissions controls, coal quality controls, and thermal generation controls. Finally, we include county-of-AQ-monitor/month-of-sample fixed effects. We weight observations by the population of the county where the AQ monitor is located for these regressions. See Section 5.2 for more details regarding our regression specification. See Appendix Table A.6 for the summary statistics associated with these regressions.

A.3 Additional Results: Quantification of Local Environmental Damages

Table A.9: Local Environmental Costs of Coal Procurement: Levels-Levels Specification

Air Pollution Costs: Dollars Per Ton

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	95.04	167.91	39.42
$ND_{p,t}$	33.50	59.19	14.82

Air Pollution Costs: Dollars Per MWh

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	46.48	82.12	19.49
$ND_{p,t}$	18.87	33.34	8.03

Air Pollution Costs: Life-Years Per 10,000 Tons

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	1.23	2.17	3.15
$ND_{p,t}$	0.46	0.81	1.18

Notes: This table presents the local environmental costs of coal stockpiles and number of deliveries for the *median* plant in our sample. We use the effect of coal stockpiles (number of deliveries) on $PM_{2.5}$ based on the Levels-Levels specification accounting for the relative bearing between power plant and air quality monitor; we average over air quality monitors and months-of-sample to obtain a plant-specific partial effect. Column 1 of each panel uses the link between $PM_{2.5}$ estimated in Krewski et al. (2009) for adults (ages 30+), Column 2 of each panel uses our own estimated link between $PM_{2.5}$ and adult mortality from Section 5, and Column 3 uses our estimated link between $PM_{2.5}$ and infant (ages 0-4) mortality. The top panel of this table presents local environmental costs per ton of coal stockpiled and delivered. The middle panel of this table presents social costs per MWh-equivalent of coal stockpiled and delivered; we convert tons of coal to MWh of electricity by taking plant-specific total number of tons of coal burned and dividing by total electricity generated. We use a value of statistical life of 9.85 million dollars to quantify the costs in dollars of the increased mortality from coal procurement based $PM_{2.5}$ for these top two panels. Finally, the bottom panel presents social costs in life-years per 10,000 tons stockpiled and delivered.

Table A.10: Local Environmental Costs of Coal Procurement Not Accounting For Wind Direction

Air Pollution Costs: Dollars Per Ton

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	74.94	132.40	39.37
$ND_{p,t}$	89.81	158.66	42.81

Air Pollution Costs: Dollars Per MWh

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	39.23	69.31	19.68
$ND_{p,t}$	45.36	80.13	21.05

Air Pollution Costs: Life-Years Per 10,000 Tons

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	1.03	1.81	3.15
$ND_{p,t}$	1.16	2.04	3.42

Notes: This table presents the local environmental costs of coal stockpiles and number of deliveries for the *median* plant in our sample. We use the effect of coal stockpiles (number of deliveries) on $PM_{2.5}$ based on the Log-Log specification that does not account for the relative bearing between power plant and air quality monitor; we average over air quality monitors and months-of-sample to obtain a plant-specific partial effect. Column 1 of each panel uses the link between $PM_{2.5}$ estimated in Krewski et al. (2009) for adults (ages 30+), Column 2 of each panel uses our own estimated link between $PM_{2.5}$ and adult mortality from Section 5, and Column 3 uses our estimated link between $PM_{2.5}$ and infant (ages 0-4) mortality. The top panel of this table presents local environmental costs per ton of coal stockpiled and delivered. The middle panel of this table presents social costs per MWh-equivalent of coal stockpiled and delivered; we convert tons of coal to MWh of electricity by taking plant-specific total number of tons of coal burned and dividing by total electricity generated. We use a value of statistical life of 9.85 million dollars to quantify the costs in dollars of the increased mortality from coal procurement based $PM_{2.5}$ for these top two panels. Finally, the bottom panel presents social costs in life-years per 10,000 tons stockpiled and delivered.

Table A.11: Local Environmental Costs of Coal Procurement In Life-Years: Per GWh

	Adult: Krewski et al. (2009)	Adult: IV Regression	Infant: IV Regression
$CS_{p,t}$	0.06	0.11	0.17
$ND_{p,t}$	0.06	0.11	0.19

Notes: This table presents the local environmental costs of coal stockpiles and number of deliveries for the *median* plant in our sample. We use the partial effects of coal stockpiles and number of deliveries on $PM_{2.5}$ based on the Log-Log specification accounting for the relative bearing between power plant and air quality monitor from Section 4; we average over air quality monitors and months-of-sample to obtain plant-specific partial effects. This table presents social costs in life-years per GWh-equivalent of coal stockpiled (delivered); we convert tons of coal to GWh of electricity by taking plant-specific total number of tons of coal burned and dividing by total electricity generated. Column 1 of each panel uses the link between $PM_{2.5}$ estimated in Krewski et al. (2009) for adults (ages 30+), Column 2 of each panel uses our own estimated link between $PM_{2.5}$ and adult mortality from Section 5, and Column 3 uses our estimated link between $PM_{2.5}$ and infant (ages 0-4) mortality.

B Data Appendix

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

B.1 Data Sources: $PM_{2.5}$ and CO 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 $PM_{2.5}$ concentrations at roughly 1,000 monitored sites across the contiguous United States. We aggregate these data to monthly average $PM_{2.5}$ levels for each air quality monitor for the sample period 2002-2012. Similarly, the AQS data includes hourly readings of carbon monoxide (CO) concentrations at roughly 700 monitored sites across the contiguous United States. As with $PM_{2.5}$, we aggregate these data to monthly average CO levels for each air quality monitor for the regressions in Section 4.4. Importantly, the AQS database also provides the latitude and longitude for each $PM_{2.5}$ monitor and CO monitor.

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

B.2 Data Sources: Coal Procurement

We collect monthly, plant-level data on end-of-month coal inventories, total monthly coal consumption, and total monthly generation from Forms EIA-906 (for 2002-2007) and EIA-923 (for 2008-2012).[30] Data on coal stocks 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 coal purchase quantities, prices, number of deliveries, sulfur content and ash content from Forms EIA-423 (2002-2007) and EIA-923 (2008-2012).[31] The variables in this purchase dataset include month of purchase, quantity purchased, delivered price,

---

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

[31] This purchase 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.

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 is 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."

In some specifications, we consider the "number of deliveries" to a plant in each month, as measured by the number of orders each plant reports on the form for each month-of-sample. Thus, our measure for the number of deliveries may be under-estimated to the extent that deliveries under the same purchase order are aggregated as described in the above quotation.

We only consider electricity generation plants whose "primary business purpose is the sale of electricity to the public"[32]; 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 latitude and longitude for each coal-fired power plant; this database is located at: http://www.epa.gov/energy/egrid.

B.3 Data Sources: Meteorological Variables

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 U.S. locations. We aggregate these data to the meteorological monitor/month-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.

---

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

Wind speed is of primary importance to quantifying the local environmental costs associated with coal procurement as wind blowing over coal stockpiles and coal deliveries generates $PM_{2.5}$. 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 latitude and longitude for 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 for each plant in each month-of-sample.

These hourly, plant-level data on emissions from coal burned 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 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 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.