Coal Use, Air Pollution, and Student Performance*

Valentina Duque
University of Sydney

Michael Gilraine
New York University

October 2020

Abstract

Coal is both the primary source of global energy and of air pollution. This paper presents the first causal evidence of the impact of pollution due to coal power plant emissions on cognitive outcomes. Our approach combines rich longitudinal student data with a design leveraging year-to-year coal plant emissions, persistent wind patterns, and also plant closures. We find that every one million megawatt hours of coal-fired power production decreases student performance in schools within ten kilometers by 0.02σ. Gas-fired plants exhibit no such relationship. Our analysis indicates that declining coal use has affected student performance and test score inequality substantially.

Keywords: Air Pollution; Coal Power; Education; Cognitive Outcomes; Health.
JEL codes: Q53, II4, I24

---

*Gilraine: Department of Economics, New York University, New York, NY 10003 (email: mike.gilraine@nyu.edu); Duque: Department of Economics, University of Sydney, Camperdown, NSW, Australia 2006 (email: valentina.duque@sydney.edu.au). We would like to thank Joshua Murphy, Robert McMillan, Alfia Karimova, Joshua Lewis, Alastair Fraser, and Janet Currie for helpful discussions. Timothy Grey provided excellent research assistance. Financial support from the University of Sydney FASS Small Grant is gratefully acknowledged. All remaining errors are our own.

1 Introduction

Coal is the largest fuel source for electricity production worldwide, and contributes to the bulk of pollution emissions related to energy production. In the United States, for example, coal-fired power plants account for a third of power production, but 60% of all sulfur dioxide, 50% of mercury, 60% of arsenic, and 13% of nitrogen oxide emissions (US Environmental Protection Agency, 2018). As the Clean Air Task Force puts it: "Among all industrial sources of air pollution, none poses greater risks to human health and the environment than coal-fired power plants."[1]

The recent trends, especially in the United States, are as striking as the high levels of coal usage and emissions. Over the past decade, U.S. coal use has declined dramatically, as cheaper and cleaner alternatives have replaced it, especially natural gas.[2] Given the considerable emissions of coal-fired power plants, the precipitous decline in coal usage has substantially improved air quality (Currie and Walker, 2019), which provides a unique opportunity to learn about the effects of air pollution caused by electricity generation. While a large body of research has shown air pollution negatively affects health (see, for instance, Currie, Zivin, Mullins, and Neidell 2014) and student achievement (Ebenstein, Lavy, and Roth, 2016; Persico and Venator, 2019; Heissel, Persico, and Simon, forthcoming), little research has focused on the effects of pollution derived from the largest sources of electricity generation – coal and natural gas power plants – on human capital. Given that the burning of coal is responsible for a third of carbon dioxide emissions worldwide,[3] the *global* benefits of reducing coal emissions on human capital are potentially first-order.

Speaking to these broader benefits, this paper provides the first causal evidence of the impact of cleaner air via reduced coal usage on cognitive development. To

[1] Link: https://www.catf.us/resource/the-toll-from-coal/.

[2] According to the United States Energy Information Administration, in 2006, coal-fired power plants accounted for 49% of total electricity generation. By 2018, coal's share of electricity generation decreased to 27%. During the same period, the share of natural gas-fired electricity generation nearly doubled, from 20% in 2006 to 36% in 2018.

[3] Source: International Energy Agency (IEA) (2019).

do so, it leverages detailed data on students from the state of North Carolina along with three sources of quasi-experimental variation. In our first design, we use the fact that power production varies year-to-year to compare student performance in nearby schools in high-production relative to low-production years. Because our methodology relies on year-to-year production variation, it alleviates concerns about individuals sorting into more/less polluted areas.[4] Second, we use the fact that emissions from power plants can be carried long distances by wind (Schneider and Banks, 2010; Zhang et al., 2017). This fact allows us to augment our first source of variation by comparing production-induced performance changes in schools downwind relative to upwind of a given plant. Third, we conduct an event study that leverages ten coal-fired power plant closures.

We implement our methods using administrative records from the universe of third to eighth grade public school students in North Carolina from 2000-01 through 2016-17. Along with detailed information about student-level demographic characteristics and test scores, the administrative data contain student identifiers, which allow us to track students over time. We link these data to monthly power plant production data from the U.S. Energy Information Administration, which enable us to calculate energy production levels by fuel type for all academic years. These data are then connected to persistent wind patterns using information from eighty-one meteorological stations throughout North Carolina to ascertain whether schools are 'downwind' or 'upwind' from a power plant.

All three empirical designs indicate that each one million megawatts-hour (Mwh) increase in coal-fired power production – about a third of the average yearly production of a coal plant in our sample – lowers student performance in schools within $10km$ by 0.02σ and 0.01σ in mathematics and English, respectively. The entirety of this effect is concentrated in schools downwind from the power plant. We find no such relationship for gas-fired power production. To put these estimates into perspective, they imply that the average annual production of a coal plant in North

[4] We confirm that individuals do not sort based on observables according to year-to-year production variation as shown in Table 2.

Carolina equates to the test score impacts of a $\$1,000$ decrease in household income for all households within $10km$ of the plant.[5]

Next, we use school-level proficiency data from the U.S. Department of Education to confirm that our estimates from North Carolina extend across the U.S. To do so, we leverage year-to-year production variation in power plants across the country and estimate near-identical effects in the national data to those found in North Carolina. With the relationship between student achievement and coal and gas-fired power production supported nationally, we then calculate the change in coal exposure in schools across the United States over the past decade. Multiplying these changes in exposure by our estimated impact of coal power production allows us to estimate the impact of the decline in coal use on student performance.

Nationwide, coal use has declined by forty-five percent over the last decade, yet the national impact obscures considerable spatial variation. In particular, the West (e.g., California) has seen little change in coal exposure since it never relied heavily on coal for electricity generation, while students in the Midwest[6] – who were exposed to the highest emissions from coal-fired plants in the country – have seen substantial reductions. During the school months of September-May, the average Midwestern student in 2006-07 was exposed to 408,000 Mwh of coal-fired power production occurring within $10km$ of their school. By 2017-18, this number had dropped to 128,000 Mwh, a 69 percent reduction.

Given the estimated impact of coal-fired power production on student achievement, this decline indicates that the average Midwestern student who experiences average 2017-18 relative to 2006-07 coal use for *one year* scores 0.006σ (=0.02*(0.408-0.128)) higher. Using fade out estimates from Chetty et al. (2014) that allow us to determine cumulative exposure effects, we calculate that the average Midwestern student experiencing 2017-18 coal levels scores a full 0.016σ higher than one facing

[5]Prior research used variation from the Earned Income Tax Credit (Dahl and Lochner, 2012), the Canadian Child Benefit (Milligan and Stabile, 2011), and childcare subsidies (Black et al., 2007) and found that a $\$1,000$ increase in household income raised mathematics scores by 0.04-0.07σ.

[6]The Midwest is defined as the states of Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin.

2007-08 coal use. The decline in coal use also reduced inequality, since underprivileged students are more likely to attend schools in polluted regions (Currie, Voorheis, and Walker, 2020). For instance, we calculate that the decline in coal use over the past decade reduced the black-white test score gap by 0.023σ in the Midwest. In general, we treat our estimates as conservative since we only consider the impact of coal-fired production on schools within $10km$ of the plant; prior research has found that coal-fired plants can affect air quality 20-40 miles away (Yang and Chou, 2015; Jha and Muller, 2017).

The analysis draws attention to large potential gains that remain from the complete elimination of coal. We identify six states (West Virginia, Nebraska, Wisconsin, Wyoming, Colorado, and Missouri) that can still achieve average test score gains of over 0.01σ from the elimination of coal-fired power production. This gives policymakers a set of schools to focus on to improve test scores via mitigation efforts (e.g., closing coal plants, moving schools, or providing air filters), along with also reducing test score gaps given that disadvantaged children disproportionately attend schools near coal plants. The share of electricity generated from coal remains high in many developed and developing countries – Australia (75%), China (59%), India (60%), Indonesia (60%), and South Korea (40%) to name a few – indicating the vast potential of mitigation and reduced coal use on human capital globally.

Contribution to Literature: Our paper makes several contributions. First, we provide the first causal estimates of the relationship between fossil fuel power production and student performance, to the best of our knowledge. While power plants can increase local economic activity (Greenstone et al., 2010), our research adds to the growing literature demonstrating that these plants negatively affect nearby children (Currie et al., 2015; Yang and Chou, 2015; Murphy, 2017).[7] In turn, this evidence points to the need for mitigation or efficient environmental regulations that equate the marginal damage of pollution to marginal abatement costs of, say,

[7]Research in epidemiology has also shown that child's neurodevelopment is associated with exposure to coal-fired power plant emissions, primarily particulate matter and polyaromatic hydrocarbon exposure (Amster and Lew Levy, 2019).

locating power plants away from schools (Muller and Mendelsohn, 2009).

Second, we add to an emerging body of research that demonstrates the adverse effects of air pollution on cognition (Persico and Venator, 2019; Gilraine, 2020; Heissel et al., forthcoming). By showing the adverse effects of air-pollution experienced over the whole academic year on school-aged children's test scores, we differentiate our findings from studies showing short-term associations between changes in air-pollution on the day of the exam and test scores (Ebenstein et al., 2016; Roth, 2019) and the literature examining the long-term impacts of in-utero exposures and children's test scores (e.g., Bharadwaj et al. (2017) and Sanders (2012)). Third, we place in context the nationwide benefit of cleaner air, showing that transitioning from coal to cleaner fuel sources leads to a non-negligible increase in student performance and helps reduce inequality.

The rest of the paper is organized as follows: The next section describes the recent findings regarding the effects of pollution exposure on health and education along with a brief history of power production in the United States. Section 3 then sets out our empirical methodology and introduces the North Carolina data. These are followed by our results in Section 4, which are confirmed in the national data and extrapolated nationwide in Section 5. Section 6 concludes.

2 Background

This section reviews the literature related to the effects of air pollution on child health and cognition, along with some of the potential mechanisms examined by prior research. We then describe power generation in the United States, paying particular attention to recent trends in coal-fired power production.

2.1 Related Literature

Effects on children's health: An extensive literature in economics links air pollution to children's health, with the first quasi-experimental evidence coming from

Ransom and Pope (1995). Focusing on a labour strike in Utah that forced the closure of a steel mill – a major local source of particulate matter – the authors show that child hospitalizations fell when the mill closed. The well-known study by Chay and Greenstone (2003a) exploits changes in environmental regulations arising from the Clean Air Act of 1970 and find that a one-unit decline in particulates led to 8 fewer infant deaths per 100,000 live births. Recent studies echo these findings using pollution variation arising from cars (Currie et al., 2009b; Currie and Walker, 2011; Knittel et al., 2016; Simeonova et al., forthcoming), airplanes (Schlenker and Walker, 2016), local weather conditions (Deryugina et al., 2019; Heft-Neal et al., 2019), and industrial plant closures (Chay and Greenstone, 2003b; Currie et al., 2015). Given the strong association between early life conditions and future outcomes (see Almond et al. (2018) for a recent survey), several studies also link early-life pollution exposure to reduced health and productivity in adulthood (Isen et al., 2017; Grönqvist et al., 2017; Barreca et al., 2017).

In addition, several papers specifically connect coal use to infant mortality. Cesur et al. (2013) find that the adoption of natural gas by households for space heating and cooking purposes in Turkey, which supplanted coal use, caused a large decline in infant mortality. Similarly, Beach and Hanlon (2018) exploit wind patterns and show that surging coal use in mid-19th century Britain significantly increased infant mortality. In terms of trade-offs, Clay et al. (2016) investigate the mid-20th century United States and find that coal-fired power plants significantly raised the child mortality rate, although these negative effects were offset by the health benefits of electrification in counties without previous access to electricity. Likewise, Muller et al. (2011) evaluate trade-offs and calculate that the air pollution damages of coal-fired power production are greater than their net contribution to economic output. Finally, Deschenes et al. (2017) show that declines in nitrogen oxides pollution (mostly due to coal plants) reduce medication purchases in addition to mortality.

Effects on children's test scores: Compared to the substantial body of research on the air pollution effects on health outcomes, far fewer studies examine the effects on child cognition. However, measuring air pollution effects beyond health is important, as previous work has linked test score improvements with long-term gains in outcomes (Garces et al., 2002; Chetty et al., 2011).

Zweig et al. (2009) link changes in outdoor air pollution near Los Angeles schools with changes in students' test scores and find that a 10 percent decrease in outdoor fine particulate matter ($PM_{2.5}$) and nitrogen oxides (NO_x) raises mathematics test scores by 0.34 percent and 0.18 percent, respectively.[8] Similarly, Ebenstein et al. (2016) use variation in test day particulate pollution in Israel and find that every 10 percent increase in $PM_{2.5}$ on the day of the test reduces test scores by 0.02 standard deviations. Using data from Texas, Sanders (2012) shows that a one-standard deviation decrease in total suspended particles in a student's year of birth leads to a 0.06 standard deviation increase in high school test scores; and Bharadwaj et al. (2017), using pollution-variation from Santiago, Chile, show that prenatal exposure to carbon monoxide reduces fourth grade test scores by 0.04 standard deviations. Persico and Venator (2019) find that students attending schools within one mile of a Toxic Release Inventory site experience a 0.024 standard deviations increase in test scores once the site closed relative to those attending schools between 1 and 2 miles away. Likewise, Heissel et al. (forthcoming) show that students attending a school where the prevailing winds place it 'downwind' of a highway score 0.04 standard deviations below comparable students attending 'upwind' schools.

Potential mechanisms involving air pollution: Pollution can affect students' test scores through several mechanisms. First, air pollution can lower the availability of oxygen in the environment thereby limiting the oxygen required for the appropriate functioning of the brain. Studies have shown that short-term exposure to air pollution is associated with inflammation and oxidative stress in the brain (Kleinman and Campbell, 2014), cerebro-vascular dysfunction, and alterations in the blood-brain barrier of the central nervous system (Genc et al., 2012), negatively

[8] Most of the scientific literature on the health effects of air pollution focuses on fine particulate matter ($PM_{2.5}$) as it can penetrate lung tissue and get into the bloodstream (Pope and Dockery, 2006; Dominici et al., 2014; Deryugina et al., 2020).

affecting an individual's concentration and decision-making (Heyes et al., 2016). Second, air pollution is associated with contemporaneous health conditions such as eye or nose irritation or child's asthma (Neidell, 2004; Lleras-Muney, 2010; Ward, 2015; Alexander and Currie, 2017; Simeonova et al., forthcoming), which may impede a child's learning in school. Third, polluted air can also increase students' (and teachers') school absences (Currie et al., 2009a; Persico and Venator, 2019), and these absences can in turn cause lower grades.

2.2 Energy Production in the United States

Coal has been used to generate electricity in the United States since the beginning of electrification: Pearl Street Station, the first commercial power plant, used coal for electricity generation when it began operating in Manhattan in 1882 (Bennion, 1940). Coal maintained its dominance as the leading fuel for electricity generation as the country electrified, and by the end of the 20th century, coal generated over half the country's electricity.

The start of the 21st century, however, has seen a precipitous decline in coal-fired electricity generation: the average share of electricity generated from coal in the United States dropped from 49% in 2006 to 27% in 2018. The leading causes are the emergence of shale gas and increased environmental regulation (Schiavo and Mendelsohn, 2019). In particular, the emergence of shale gas due to advances in hydraulic fracturing extraction techniques caused gas prices to fall, making natural gas a cheaper alternative to coal (Joskow, 2013; Knittel et al., 2015; Fell and Kaffine, 2018). In turn, the share of electricity produced by natural gas nearly doubled from 20% in 2006 to 36% in 2018. The decline in coal was also abetted by increased environmental regulation, namely the Acid Rain Program, the Cross State Air Pollution Rule, and the Mercury and Air Toxics Standards (Burtraw et al., 2012; US Environmental Protection Agency, 2015; Chan et al., 2018).

Coal plants are considered a 'dirty' fossil fuel, releasing large amounts of airborne pollutants such as particulates, carbon dioxide, carbon monoxide, sulfur dioxide, and nitrogen oxides, as well as airborne toxins in the form of mercury, lead, and various other heavy metals (e.g., arsenic, cadmium, and cobalt). In contrast, natural gas is known as a 'clean' fossil fuel, since it emits roughly half the carbon dioxide and one-quarter the nitrogen oxides of coal and almost no sulfur dioxide, carbon monoxide, black carbon, particulates, and mercury (Nature, 2009). Given coal's substantial emissions of airborne pollutants, research has linked the decline in coal use to significant improvements in airborne pollution across the United States over the last fifteen years, including decreased mercury, carbon dioxide, nitrogen oxides, and sulfur dioxides (Venkatesh et al., 2012; De Gouw et al., 2014; Zhang et al., 2016; Holland et al., forthcoming).

Figure A.1 reports the number of large coal and gas-fired power plants in the United States from 2004 to 2018, along with their emissions intensity for three pollutants: carbon dioxide, sulfur dioxide, and nitrogen oxides. In line with the previous discussion, we see a large drop in the number of coal plants, and a corresponding rise in natural gas plants over the last fifteen years. The emissions intensity of coal-fired power plants has seen extraordinary improvements over the last fifteen years as the emissions per unit of power produced for nitrogen oxides and sulfur dioxide have declined by a factor of two and four, respectively. In terms of emissions intensity, gas-fired plants release virtually no sulfur dioxide and substantially less nitrogen oxides and carbon dioxide emissions than coal-fired plants; further, their emissions intensity have been remarkably stable.

Driving the improvements for coal-fired pollution is a combination of the retirement of less efficient coal plants, increased use of low-sulfur coal, and the installation of pollution abatement technologies, such as low NO_x burners and flue-gas desulfurization scrubbers. Flue gas desulfurization scrubbers are the key driver in the reduction of sulfur dioxide emissions, with the large emissions decline from 2007-2011 coinciding with the installation of these scrubbers in over 110 coal plants at a cost of $30 billion. At the start of our time period, nearly all large coal plants in the United States had particulate matter control systems (e.g., electrostatic precipitators) and mercury control systems so we expect that the emissions intensity for these pollutants to be relatively constant during our time period, although reports do indicate some improvements for these as well.[9]

North Carolina's energy sector mimics national trends in energy production. Coal provided roughly sixty percent of North Carolina's electricity generation until 2010, when natural gas generation began to rise rapidly. As in the rest of the country, this was driven by the decline in natural gas prices due to advances in hydraulic fracturing increasing gas production in states such as Texas, North Dakota, Colorado, and Wyoming. From 2010 to 2017, coal's share of electricity generation dropped from 60% to 27%, while the share of electricity generation from natural gas increased six-fold, from 5% in 2010 to 30% in 2017. During this time, nuclear power remained the only other major source of power and its share of electricity generation was remarkably stable, representing 32% of electricity generation in 2001 and 33% in 2017.

Similar to nationwide trends, North Carolina's coal-fired power plants have seen improvements in emissions intensity (see Table A.2). The improvements for sulfur dioxide emissions are especially pronounced, likely because the 2007-2014 period saw the installation of flue-gas desulfurization scrubbers in nearly every boiler in the state or the closure of boilers that lacked these scrubbers. Emissions intensity for nitrogen oxides and carbon dioxide are otherwise very similar to the rest of the country.

An additional catalyst for the decline in coal use in North Carolina comes from the fact that its largest energy producer, Duke Energy, was subject to an ongoing lawsuit from 2000-2015 by the United States Justice Department on behalf of the Environmental Protection Agency. The lawsuit claimed that Duke Energy violated federal clean air laws by modifying thirteen coal-fired power generators without the required equipment to control air pollution. During the fifteen-year lawsuit, Duke

[9] For instance, sulfur dioxide and nitrogen oxides can be chemically transformed in the atmosphere into fine particulate matter so reduction in those pollutants will also generate a corresponding reduction in particulate matter.

Energy shut down eleven of these units by closing four power plants and eventually settled in 2015 by agreeing to pay a civil penalty and shut down the remaining two offending units by 2024.[10]

3 Methods and Data

Our empirical strategy for estimating the effect of air pollution caused by coal-fired power production on student achievement relies on three 'quasi-experiments' that leverage separate sources of variation. This section describes each of these and then introduces the data set that we have assembled.

3.1 Empirical Strategies

Quasi-experiment 1: Production Variation. Our first quasi-experiment uses the fact that power production at individual plants varies significantly year-to-year, allowing us to compare student performance in schools nearby during high production years to low production years. Since higher power production is associated with increased plant-level pollution, we expect student performance in nearby schools to suffer in years of high power production.

We investigate the relationship between power production and student performance through the following regression:

$$y_{igs(p)t} = \gamma_{s(p)} + \beta prod_{s(p)t} + \phi Z_{igs(p)t} + \theta_i + \lambda_{gt} + \epsilon_{igs(p)t}\,, \quad for\ X_{s(p)} \leq \bar{X}\,, \tag{1}$$

where $y_{igs(p)t}$ is the test score of student i in grade g at the end of school year t who attends school s that is nearest to power plant p, $prod_{s(p)t}$ represents the annual power production (in millions of Mwh) of plant p near school s during school year t, and $Z_{igs(p)t}$ includes student demographic controls (e.g., socioeconomically disadvantaged status, English learner status, disability, and gifted status). θ_i is an

[10]The Washington Post, September 2015, Link: https://www.washingtonpost.com/national/duke-energy-to-settle-federal-lawsuit-over-claims-of-clean-air-law-violations/2015/09/10/953adeca-57fc-11e5-8bb1-b488d231bba2_story.html

individual fixed effect to account for time-invariant student-level characteristics, $\gamma_{s(p)}$ is a school fixed effect which absorbs time-invariant school-level characteristics such as school location, and λ_{gt} are grade-by-year fixed effects, accounting for temporal shocks that are common to all students in a particular grade-year (e.g., state-level curriculum reform). $X_{s(p)}$ denotes the distance of school s to power plant p and $\bar{X}$ is our 'radius of interest' whereby we restrict our regression to only include schools within distance $\bar{X}$ of a power plant. We bracket the p subscript to indicate that schools are always assigned to one power plant (and so school fixed effects subsume power plant fixed effects). Standard errors are clustered at the school level to account for within-school serial correlation in the observations.[11] β is our coefficient of interest and represents the effect of a one million Mwh increase in annual power plant production on student test scores.

Although the prior literature has indicated that coal-fired power plants can affect air quality 20-40 miles away (Yang and Chou, 2015; Jha and Muller, 2017), we set $\bar{X} = 10km$ (6.2 miles) throughout so that we focus on regions near the power plant that are likely to face similar economic shocks (or other conditions) in a given year. We also establish the impact of power production on student performance as a function of distance to the power plant by plotting the relationship between power production and student performance for various $2.5km$ distance bins. Doing so, we find significant declines in student performance among bins below $10km$, with no significant declines in performance observed in bins above $10km$, in line with expectations that the impact of power plant emissions diminishes with distance.[12]

Quasi-experiment 2: Wind *and* Production Variation. Our second quasi-experiment introduces an additional level of variation coming from the fact that wind carries air pollutants downwind from power plants. This second fact allows us to combine persistent wind direction with the year-to-year power production variation used in our first quasi-experiment, alleviating concerns that would arise if

[11] Clustering at the school level also leads to the most conservative standard errors.

[12] Given the spatial distribution of power plants in North Carolina, setting $\bar{X} = 10km$ has the additional benefit that schools are not treated by more than one power plant.

we *solely* used power production, as it might be correlated with other local factors (e.g., employment or wages) that could also affect student achievement.[13]

Intuitively, wind variation allows us to compare student performance in years of high and low power production for schools that are relatively 'downwind' versus 'upwind' of the power plant. Since both 'downwind' and 'upwind' schools are near to the power plant, any unobserved variables that might affect student achievement and are correlated with power production (e.g., income shocks, weather, etc.) should similarly affect students attending either type of school. Therefore, any differential effect of the increased power production between 'downwind' relative to 'upwind' schools can be attributed to the fact that pollution travels downwind (as in Deryugina et al. (2019); Anderson (forthcoming); Heissel et al. (forthcoming)), and so students in downwind schools are more affected by the higher pollution levels caused by increased power production.

We leverage both wind and year-to-year power production variation by estimating the following regression:

$$\begin{aligned} y_{igs(p)t} = \gamma_{s(p)} + \beta prod_{s(p)t} * downwind_{s(p)} + \delta prod_{s(p)t} \\ + \phi Z_{igs(p)t} + \theta_i + \lambda_{gt} + \epsilon_{igs(p)t}\,,\ for\ X_{s(p)} \leq \bar{X}\,, \end{aligned} \tag{2}$$

where all the terms are the same as in equation (1), with the additional term $downwind_{s(p)}$ denoting our measure of how 'downwind' school s is from plant p (formally defined below in equation (6)). β is our coefficient of interest and represents the marginal effect of a one million Mwh increase in annual power plant production in downwind relative to upwind schools on student test scores.

Quasi-experiment 3: Coal Plant Closures. Our third identification strategy

[13] Alternatively, one could simply contrast student performance in downwind relative to upwind schools. This identification strategy would mirror Anderson (forthcoming), who compares health outcomes in individuals living upwind versus downwind of highways. While Anderson finds no evidence of sorting, that design may be ill-suited to our context as power plants are likely more salient than highways. Indeed, we detect differences in observables across upwind and downwind schools (see Table A.2), necessitating that we augment wind variation with production variation given we find no sorting on observables based on production variation (see Table 2).

leverages coal plant closures in North Carolina, using six coal plant closures and four coal plant conversions into natural gas plants as natural experiments. Given that we do not find any relation between natural gas energy production and student achievement (as we show later in Tables 3 and A.3), we do not distinguish between coal plant conversions or closures, leaving us with ten events. Of these, four occurred in 2010-11, making 2011-12 the first year with no nearby coal-fired power production for almost half our sample.

We use an event-study approach to estimate the effects of coal plant closures on students' test scores using the following equation:

$$y_{igs(p)t} = \gamma_{s(p)} + \beta post_{s(p)t} + \delta pre{*}closed_{s(p)t} + \phi Z_{igs(p)t} + \theta_i + \epsilon_{igs(p)t}, \ for\ X_{s(p)} \leq \bar{X}, \quad (3)$$

where $closed_{s(p)t}$ denotes time in years relative to the plant closure (i.e., event time), $post_{s(p)t} \equiv \mathbb{1}\{closed_{s(p)t} \geq 0\}$ indicates that plant p nearest to school s has closed and $pre{*}closed_{s(p)t} \equiv (1 - post_{s(p)t}){*}closed_{s(p)t}$ controls for the (linear) trend in test scores in the years leading up to a plant closure. All other terms are the same as those in equation (1).

We also assess the validity of the event study by plotting event-time coefficients using the following regression:

$$\begin{aligned} y_{igs(p)t} = {} & \gamma_{s(p)} + \sum_{j=-4}^{-2} \gamma_j \cdot \mathbb{1}\{Plant\ Closure_{s(p)t} = j\} + \sum_{j=0}^{3} \beta_j \cdot \mathbb{1}\{Plant\ Closure_{s(p)t} = j\} \\ & + \phi Z_{igs(p)t} + \theta_i + \epsilon_{igs(p)t}\ , \ for\ X_{s(p)} \leq \bar{X}\ , \end{aligned} \quad (4)$$

where $\mathbb{1}\{Plant\ Closure_{s(p)t} = j\}$ is an indicator variable equal to one if school s near plant p is in event time j, and equal to zero otherwise. This specification is a non-parametric version of equation (3) and allows for non-parametric estimation of the time path of the estimated effects of the coal plant closure at different periods. The event-time coefficients γ_j show how student test scores evolve in the years leading up to the plant closure and β_j display the evolution of test scores after the plant closure. Period $j = -1$ in event-time (i.e., the year before the plant closes) is the omitted category. Given the inclusion of student fixed effects, the sample is restricted to four periods pre- and post-closure to ensure that we observe the same student both before and after the plant closure.

3.2 Data

We now describe the data we use for estimation. Our data combines three distinct data sources to construct a student-level data set that captures students' exposure to nearby power plants both in relation to power production levels and wind direction.

Power Plant Data: Our three sources of quasi-experimental variation require information on power plant location and production. These data come from the U.S. Energy Information Administration (EIA) Monthly Generation and Fuel Consumption Time Series File (EIA-923) and its pre-2008 predecessor (EIA-906).[14] These data report monthly power production by plant and type of fuel for all large power plants across the United States for the calendar years 2001-2018, which we use as our analysis period.

The EIA data cover all power-producing locations, including those relying on renewable energy sources (e.g., hydroelectric) as well as small power producing locations (e.g., crematoriums, factories, mills) that are unlikely to affect individuals significantly by themselves. We thus restrict our sample to large coal and natural gas fired power plants with average power production levels of over 250,000 Mwh during the academic year over their period of operation. The 250,000 Mwh threshold restricts the majority of our sample to large power stations, although our data also includes the world's largest tobacco factory (at the time).

To match the academic calendar, we sum power production generated from September to May[15] for each plant by fossil fuel type. Each plant-year is classified as being coal or natural gas, based on the majority fuel type used in that academic

[14] Both available from United States Energy Information Administration (EIA) (2020).

[15] School calendars vary across school districts, but the academic year usually runs from the last week of August to the first week of June.

year. These designations are clear in the data: coal plants on average generate 98 percent of their electricity using coal, with the remainder being produced using natural gas, oil, or biomass. Similarly, natural gas plants on average generate 97 percent of their electricity using natural gas. Given this, total power production for each plant for each academic year is defined as the power produced at that plant coming from its main fossil fuel (i.e., coal or natural gas) from September-May.

We are left with a sample of twenty-one coal and natural gas plants, of which (as four coal plants converted to natural gas) seventeen operated as coal plants and eight operated as gas plants at some point during our period of interest (2001-2017). Of the remaining thirteen coal plants that did not convert to natural gas, six closed during our analysis period. Figure 1 shows the location of all twenty-one power plants across the state of North Carolina along with their production type, and Table A.1 lists all power plants along with their location, dates of operation, production characteristics, and distance to the nearest wind station (see below).

Wind Data: With our sample of power plants in hand, we now construct a measure of wind direction for each power plant. To do so, we obtain information on wind patterns from the National Oceanic and Atmospheric Administration's Surface Hourly Global data.[16] These data report wind speed and direction every five minutes from eighty-one meteorological stations across the state of North Carolina. Because the number of wind stations has been increasing over time and wind patterns are consistent year-to-year, we proxy wind direction for the earlier years (i.e., 2001-2013) with data from 2014-2017. We then restrict the wind observations to those: (i) that occur during school hours,[17] and (ii) with non-negligible wind speeds to ensure that the wind is pushing pollutants in a given direction.[18]

The data are then collapsed by angle of wind according to an 8-directional wind-

[16] Available from National Oceanic and Atmospheric Administration (2018).

[17] We assume that school hours run from 8am through 3pm, although school hours do vary somewhat by district.

[18] Specifically, only observations where the wind speed is 5mph or greater are included. We also omit observations where the wind direction cannot be determined as it is bi-directional (i.e., 'swirling').

rose. The wind data thus represent the proportion of time wind blows *toward* one of four cardinal directions (N, E, S, W) plus the four intercardinal directions (NE, SE, SW, NW) during school hours. Formally, given N observations per wind station w, the proportion of time the wind blows *toward* direction D during school hours is expressed as follows:

$$wind_{w(p)D} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\{wind_i = D\}\,,\ D \in \{N, NE, E, SE, S, SW, W, NW\}\,. \quad (5)$$

Power plants are then matched to the nearest wind station according to physical distance. The closest wind station is usually within twenty kilometers of the power plant (see column "Km to Wind Station" in Table A.1), which should adequately capture the direction wind is blowing at the power plant given that wind direction is highly spatially correlated. In fact, prevailing wind directions are fairly consistent across North Carolina, with the wind generally blowing toward the North and, to a lesser extent, the South. Only a small portion of the state has wind predominantly blowing from the East or West.

School Location Data: We combine our plant-level production and wind direction data with geocoded location data on every public school in North Carolina (including charter schools). Using these data, we first calculate the distance of a school relative to the nearest power plant, giving us a unique school-power plant combination.[19] We then use the coordinates of both the school and its nearby power plant to calculate the *direction* of the school relative to the power plant. Once again, we discretize the direction the school is located relative to the power plant into one of four cardinal directions (N, E, S, W) plus the four intercardinal directions (NE, SE, SW, NW).

Next, we calculate the proportion of time the wind is blowing from a given power plant toward a given school during school hours. There is no clear-cut way to calculate this measure. We choose one method we believe is reasonable, and then

[19]Given the spacing of power plants (see Figure 1), there are no cases where schools are within 10km of two power plants and so the school-power plant combination is unique.

show robustness to this measure (see Figure A.5). Specifically, denote D' as the directions adjacent to a given direction D (e.g., if $D = N$, then $D' = \{NE, NW\}$). We define the degree 'downwind' school s is from plant p as:

$$\begin{aligned} downwind_{s(p)} =& \mathbb{1}\{School\ Location_{s(p)} = D\} * wind_{w(p)D} \\ &+ \frac{1}{2}\mathbb{1}\{School\ Location_{s(p)} = D\} * wind_{w(p)D'}\,. \end{aligned} \tag{6}$$

Effectively, this measure captures the proportion of time the wind is blowing from power plant p toward school s either directly or from an adjacent direction, with adjacent directions given half the weight.

Figure 2 provides an example of a typical wind rose – a diagram often used by meteorologists to summarize the distribution of the wind direction – to provide intuition for our wind measure. The circular format of the wind rose allows one to represent the direction the wind is blowing *toward*[20] and the length of each circle segment shows the proportion of time the wind blows toward that particular direction during school hours. For instance, the dark segment in Figure 2 shows that for 27.5 percent of observations the wind blows toward the North 27.5.

As is typical in North Carolina, Figure 2 indicates wind generally blows *toward* the North, with little wind blowing East or West. The 'downwind' measure for a school located North of the power plant will incorporate all three shaded segments (N, NW, NE), with the lighter segments (NW,NE) being given half weight. A school located due North will have a 'downwind' measure of 0.4375 ($= 0.275 + \frac{1}{2}0.1625 + \frac{1}{2}0.1625$), which is relatively high. Likewise, a school located to the East of the power plant has a relatively low 'downwind' measure of only 0.16875 ($= 0.05 + \frac{1}{2}0.1625 + \frac{1}{2}0.075$).

Administrative Student Data: Our data consist of all public schools in North Carolina, the distance of each public school to the nearest power plant, and a measure of how often the wind blows from the relevant power plant toward the school.

[20] Usually, wind roses indicate the direction the wind blows *from*. We reverse this and show the direction wind is blowing *toward* to help with intuition.

We incorporate detailed administrative data from the North Carolina Education Research Center (NCERDC) and assign students to schools based on attendance.[21] Our student-level data include information on all public school students in the state for the 2000-01 to 2016-17 academic years. Importantly, the NCERDC data contain unique student identifiers, allowing us to track students over time.

The student-level data contain test scores for each student in mathematics and English for grades two through eight from standardized tests that are administered at the end of each school year in the state.[22] Test scores are reported on a developmental scale, which is designed such that each additional test score point represents the same knowledge gain, regardless of the student's grade or baseline ability. We standardize this scale at the student level to have a mean of zero and a variance of one for each grade-year. Student-level demographics include sex, race/ethnicity, socioeconomic status, English learner status, disability, and gifted status.

Summary statistics are reported in Table 1. Column (1) shows student characteristics for all students in the sample ($N = 2.5$ million, or 9.2 million student-year observations). North Carolina has a white student plurality and a substantial black minority population (28 percent), with Hispanic and Asian students making up a further eleven and three percent of the student body, respectively. Almost half are socioeconomically disadvantaged, ten percent report having a disability, and fifteen percent are gifted.

Columns (2)-(4) present summary statistics for the samples of our three sources of quasi-experimental variation. Column (2) restricts the data to students in schools located within $10km$ of a coal power plant, which corresponds to the sample used in the first and second quasi-experiments. While less than ten percent of the full sample attend schools within $10km$ of a coal plant, these students appear similar to the overall student body both in terms of test scores and demographics. Column (3) restricts the sample to students who attend a school within $10km$ of a natural gas

[21] Data citation: North Carolina Education Research Data Center (1995-2017).

[22] The exception is the second grade test, which is administered at the start of the school year for students in third grade. In addition, the second grade test was discontinued after 2007-08 and is not available in either 2005-06 for mathematics nor 2007-08 for English.

plant, while column (4) restricts the sample to students in schools within $10km$ of a coal plant that shuts down or converts to natural gas during the period of interest. The samples in both columns (3) and (4) consist of lower-performing students who come from more disadvantaged backgrounds. An important reason for these differences is that three of the coal plants that converted to natural gas are located near highly-disadvantaged neighborhoods.

4 Results

This section provides estimates of the effect of coal-fired power production on student performance using our three quasi-experiments, leveraging variation in production, wind, and plant closures. We also use the first two methodologies to measure the impact of natural gas power plants.[23]

Sorting: Before estimating the effects of coal-fired power production on student performance, we examine whether students sort endogenously into more/less polluted schools with changes in air pollution. Since our design leverages year-to-year variation in coal-fired power production, we note that any story whereby students sort in response to treatment must be driven by students sorting yearly in response to changes in power production, which seems unlikely. Our data allow us to test for sorting based on observables directly. Table 2 does so by regressing seven covariates on the production of the nearby power plant (Panel A) or production times downwind status (Panel B) along with school, grade, and year fixed effects. All of these point estimates are statistically insignificant, indicating that the composition of schools does not change systematically with year-to-year production variation – i.e., our 'treatment.'

Quasi-experiment 1: Production Variation. Our first quasi-experiment compares student performance in schools near power plants in years when power production is high to years when power production is low. Figure A.3 provides a rep-

[23]We do not have enough natural gas plant openings or closings to use the third source of variation.

resentation of the year-to-year production variation that we are using by displaying (logged) power production for the seven coal-fired power plants that are always open during our sample period. The figure makes clear that there is substantial plant-specific production variation over time that can be leveraged for identification. In addition, declines in power production are not strongly correlated across plants: during the Great Recession, for instance, there were pronounced production declines at Cliffside and GG Allen but production was relatively constant at the other five plants.

Figure 3 plots the nonparametric relationship between power production and student mathematics performance by distance to the plant. To do so, it groups schools into $2.5km$ bins and reports estimates of equation (1) for schools within those bins.[24] Figure 3(a) plots the relationship for schools near coal-fired power plants. Schools that are less than $10km$ from the coal-fired power plant experience approximately a 0.02σ decline in mathematics scores for each million Mwh increase in power production at the nearby power plant. After $10km$, the relationship between increased power production and test scores disappears, providing empirical support for our decision to restrict the effect of coal-fired power production to schools within $10km$ for our main analysis. Figure 3(b) plots the same relationship for natural gas fired power production. Visually, there is no relationship between natural gas power production and student performance.

Panel A of Table 3 reports the results of estimating equation (1) for schools within $10km$ of a coal-fired power plant for both mathematics and English. In line with the visual evidence, results indicate that a one million Mwh increase in coal-fired power production is associated with a 0.02σ decline in mathematics test scores, which is statistically significant at the one percent level. The impact on English scores is about half of the mathematics score effect and is statistically significant at the five percent level.

Panel B of Table 3 estimates the same equation, but for schools within $10km$ of

[24]We combine the 0-$2.5km$ and 2.5-$5km$ bins as there are few schools within $2.5km$ of a power plant.

a natural gas power plant. Given that natural gas-fired power production generates substantially lower levels of harmful emissions, we expect a smaller relationship between natural gas production and student performance. Indeed, our point estimates indicate no clear relationship between natural gas power production and student performance.

Quasi-experiment 2: Wind *and* Production Variation. Our second quasi-experiment incorporates wind variation into our first quasi-experiment. To start, Figure 4 plots the nonparametric relationship between coal-fired power production and student mathematics performance by distance to the plant *and* by whether the plant is 'downwind' or 'upwind' of the power plant. To do so, we delineate schools as 'downwind' or 'upwind' based on whether the school faces above or below median wind blowing from the power plant according our downwind measure (see equation (6)) and then we group schools into $2.5km$ bins and report the results of equation (1) for schools within those bins. For schools 'downwind' and within $10km$ of the power plant, students experience approximately a 0.03σ decline in mathematics scores for each one million Mwh increase in production at the nearby power plant. For 'downwind' schools further than $10km$, there is no clear relationship between test performance and power production. Schools 'upwind' of the power plant do not appear to be substantially affected by increased coal-fired power production.

Table 4 reports estimates of equation (1) for schools within $10km$ that are 'downwind' (Panel A) and 'upwind' (Panel B) of the coal-fired plant for both mathematics and English. As expected, we see that a one million Mwh increase in power production leads to large and statistically significant declines in mathematics scores of 0.03σ for schools that are 'downwind' of the coal-fired plant. As in the figure, for schools 'upwind' of the plant, there is no clear relationship between power production and student performance.

Panel C then interacts coal-fired power production with our continuous measure of how downwind a school is from the plant as described by equation (2). To help interpret these point estimates, we divide our downwind measure by 0.15 to make the scale of the estimates roughly comparable to the differences between Panels A and B given that our wind measure is about 0.15 units higher for schools 'downwind' relative to 'upwind' of the power plant. Our estimates yield near-identical effect sizes to our simple 'upwind' relative to 'downwind' comparison. Once again, estimates for English are about half of those for mathematics. Table A.3 repeats the quasi-experiment for natural gas plants and finds no clear relationship between natural gas power production and student performance in either 'downwind' or 'upwind' schools.

Quasi-experiment 3: Coal Plant Closures. Our third quasi-experiment leverages coal-fired plant closures and conversions to natural gas. Given that we find no evidence that power production using natural gas lowers student performance (see Tables 3 and A.3), we do not distinguish between a coal plant closure or conversion to natural gas. Our study therefore incorporates ten events: six coal plant closures and four conversions to natural gas. The closure/conversion date we use for each plant is given in the last column of Table A.1.

Figure 5 plots estimates from equation (4) for schools 'downwind' (Figure 5(a)) and 'upwind' (Figure 5(b)) of the plant, which show the evolution of student performance in the years leading up to and after closure. In the years leading up to the coal plant closure, we see that student performance is stable for both 'downwind' and 'upwind' schools. Given this, there appears to be no differential pre-trend among students in schools near plants about to close relative to those in schools near plants that will close in the future (or have closed).

After the plant closes at event time zero we see a jump in student performance among schools 'downwind' from the plant. Test scores continue to increase somewhat after plant closure, probably because some of our 'closed' plants continue to produce some coal-based electricity for a year or two afterward.[25] No such jump in test scores

[25] We consider a plant closed if it experiences greater than a seventy-five percent drop in its production. For most plants, this indicates their complete closure, but some plants keep a boiler in operation for a few ensuing years. For example, we consider Riverbend to have closed after 2010-11 when it retired four of its boilers, but the plant kept one boiler in operation for the subsequent two years.

after plant closure is observed for schools 'upwind' of the plant.

Table 5 reports results from the event study represented by equation (3). The point estimates indicate that a coal-fired power plant closure leads to a 0.05-0.06σ increase in both mathematics and English performance among schools within $10km$ of the power plant. Panels A and B delineate these results into schools that are 'downwind' (Panel A) and 'upwind' (Panel B) of the coal-fired power plant. As expected, the majority of the effect is driven by schools downwind of the power plant.

The average coal plant closure in the sample had coal power production of about one and a half million Mwh pre-closure, implying that our estimates in the event-study design are about 1.5-2 times those found in our first two quasi-experimental designs. While these differences are not statistically significant, one driver may be that the closed coal plants were older and less efficient (in terms of pollution) than the coal plants that remained open. Indeed, many of the closed plants did not have flue-gas desulfurization scrubbers and so had much higher levels of sulfur dioxide emissions.

4.1 Robustness

Placebo Test: One might be concerned that power plants jointly reduced production due to a statewide shock (e.g., the Great Recession) that differently affected students living downwind relative to upwind of power plants. While Figure A.3 reveals that production variation is, for the most part, plant-specific, we also conduct a placebo test where we estimate equation (2) but randomly assign the power production of a coal-fired plant to that of a different coal-fired plant. Figure A.4 reports the point estimates from these placebo tests and shows that they are centered around zero, only exceeding the estimate found in Panel C of Table 4 once (as one would expect by chance). The placebo test clearly indicates that our results are driven by production changes in nearby power plants (rather than statewide shocks).

Functional Form: Table A.4 recreates our main results from Table 4 using logged power production rather than production in millions of Mwh. This helps alleviate concerns that our results are solely driven by large power plants with substantial production variation in levels as the log specification captures the effect of a percentage change in power production instead. Results are qualitatively similar when the log specification is used.

Wind Measure: Our 'downwind' measure defined in equation (6) gives half weight to adjacent wind directions. Figure A.5 graphs the point estimate from our second quasi-experimental design that interacts power production with wind direction when different weights are applied to adjacent wind directions. The point estimate is remarkably stable when anything between zero through equal weights are given to adjacent wind directions.

5 Nationwide Results

Our results indicate that coal-fired power production in North Carolina causes a substantial and significant decline in test scores. Given that *nationwide* coal use declined by 45 percent from 2007 to 2018,[26] we would expect the performance of students throughout the country to have benefited from reduced coal use. This section takes our results from North Carolina and extrapolates them nationwide.

5.1 Repeating our Quasi-Experiment Using National Data

As a first step, we confirm our results in North Carolina using national performance data from the U.S. Department of Education.[27] While our power plant data are available nationwide and so are easily adapted, the performance data only contain school-level proficiency rates. These proficiency rates vary substantially across states due to different testing regimes and also within-state due to changes in testing over

[26] Total electric power generation from industrial sources using coal fell from 2.02 billion Mwh in 2007 to 1.15 billion Mwh in 2018 according to the EIA.

[27] Available from United States Department of Education (2019).

time. To deal with these issues, we construct school proficiency rankings within each state-year cell and use this measure as our outcome.[28] Appendix B.1 provides a detailed description of our data sources and the construction of our national data set.

With nationwide data in hand, we alter equation (1) and run the following regression:

$$y_{s(p)t} = \gamma_{s(p)} + \beta prod_{s(p)t} + \phi Z_{s(p)t} + \lambda_t + \epsilon_{s(p)t}\,, \quad for\ X_{s(p)t} \leq \bar{X}\,, \tag{7}$$

where $y_{s(p)t}$ is the statewide school proficiency ranking of school s that is nearest to power plant p in school year t, $prod_{s(p)t}$ represents the annual power production (in millions of Mwh) of plant p near school s during school year t, $Z_{s(p)t}$ includes time-varying school controls (e.g., student-teacher ratio, testing rate, enrollment, percent socioeconomically disadvantaged, percent female, and percent from different ethnicity groups), and $\gamma_{s(p)}$ and λ_t are school and year fixed effects, respectively. As before, $X_{s(p)}$ denotes the distance of school s to power plant p and $\bar{X}$ is the relevant radius, which we set to $10km$ as in our main specification. β is our coefficient of interest and represents the effect of a one million Mwh increase in power plant production on a school's statewide proficiency ranking.

Figure 6(a) plots the relationship between increased coal-fired power production and the school's statewide proficiency ranking in mathematics, grouping schools into $5km$ bins and reporting the results of equation (7) within those bins.[29] As in North Carolina, we observe a negative relationship between power production and performance among schools within $10km$ of the coal plant. Beyond $10km$, no such relationship is evident. As before, we find no statistically significant relationship

[28] Alternatively, we could use proficiency rates directly and include state-year fixed effects. We use within-state rankings instead for two reasons: (i) state-year fixed effects are collinear with production variation in states with one power plant, and (ii) proficiency thresholds vary substantially across states and so an identical (homogeneous) effect could lead to differential proficiency rate increases in different states. Results are qualitatively similar when we employ this alternative method: we find that a one million Mwh increase in coal-fired power production decreases the percent of students scoring proficient by 0.26 (statistically significant at the five percent level).

[29] We lack precision to graph results by $2.5km$ bins as done for North Carolina.

between power production and test scores for natural gas plants (Figure 6(b)).

Table B.6 reports the results from equation (7). We find that every one million Mwh increase in coal-fired power production decreases the statewide mathematics proficiency rank of schools within $10km$ by 0.3-0.4 percentile ranks. For comparison, if we run regression (7) using our North Carolina data, we find a point estimate of 0.42, indicating that our nationwide results are similar to those found in North Carolina. This gives us confidence that our North Carolina results can be extrapolated nationwide.

5.2 Calculating Impact of Reduced Coal Use

We calculate the impact of reduced coal usage on student performance by multiplying the estimated impact of one million Mwh of coal-fired power production on student performance obtained for North Carolina (-0.02σ) by the enrollment weighted change in exposure to coal-fired power generation within $10km$ of a school from 2006-07 to 2017-18. (See Appendix B.2 for details on the construction of our coal-fired exposure measure.)

For this calculation to be valid, several assumptions are required, although we believe these assumptions are conservative and so our estimated effect is a lower bound of the impact of reduced coal use on student cognition. First, we assume only students attending schools within $10km$ of a coal-fired power plant are affected, in line with results from Figure 3(a). Second, students are assumed to not be affected by coal-fired power generation in the summer months. Third, the impact of emissions from coal-fired power production is linear and homogeneous within $10km$. Fourth, the emissions intensity of coal-fired power plants has been constant over time.[30] Supporting this view, we see little difference in our point estimates after flue gas

[30]Figure A.1 indicates this assumption is conservative as coal power plants' emissions efficiency improved from 2004 to 2018, although these improvements are most notable for sulfur dioxide due to the installation of flue-gas desulfurization scrubbers. Health research generally considers other pollutants from coal plants such as particulate matter and nitrogen oxides to be more significant health threats than sulfur dioxide, indicating that improvements to sulfur dioxide efficiency may not affect test scores appreciably. For instance, the World Health Organization lists the most harmful air-pollutants for people's health in developed countries as being: particulate matter, ground level ozone, nitrogen oxides, and indoor radon (World Health Organization, 2017).

desulfurization scrubbers – which greatly improve sulfur dioxide emissions intensity – were installed in North Carolina.[31] Fifth, replacement sources of electricity do not influence student performance. As discussed previously, coal-fired power production was predominantly replaced by natural gas and we find no relationship between natural gas power production and student performance (see Figure 3(b)), providing support for this assumption.[32]

Impacts Throughout the U.S: The impacts from the decline in coal use feature a high degree of spatial variation. Figure 7(a) presents our estimate of the increase in test scores from 2006-07 to 2017-18 due to the decline in coal usage by state. On one hand, Western states had extremely low levels of coal usage in 2006-07 and so the decline in coal use had little effect there. Midwestern states, on the other hand, saw steep declines in coal usage. For example, students' average coal-fired power production exposure dropped almost ten-fold from 800,000 Mwh in 2007 to 90,000 Mwh in 2017-18 in the most affected state of Illinois.[33]

For the Midwest as a whole, the average student had 408,000 Mwh of coal-fired power production occur within $10km$ of their school during the school months of September-May in 2006-07. By 2017-18, this number had dropped to 128,000 Mwh. Our point estimates indicate that a *one year* decline of one million Mwh in coal-fired power production raises test scores by 0.02σ. Therefore, a student who experiences average 2017-18 coal use relative to 2007-08 coal use for *one year* would score 0.006σ (=0.02*(0.408-0.128)) higher. The decline in coal use, however, affects students every year they attend school. Converting these yearly estimates into a cumulative impact is difficult as the impact of educational interventions often "fade out." We make these conversions using the fade out estimates of teacher quality

[31] Specifically, among coal-fired plants that are open throughout our analysis period flue-gas desulfurization scrubbers were installed between 2006-2010. Our preferred point estimates (mimicking column (2) of Table 3) are -0.021 for both the time period preceding and following their installation among these always-open plants.

[32] Increased renewable energy production also supplanted coal-fired power plants; it is unlikely these sources would impact student performance given that they are emissions-free.

[33] One reason for the large change in coal exposure in Illinois is the closure of three coal-fired power plants in 2010 (Fisk, Crawford, and State Line) that where near or within the city limits of Chicago.

reported in Chetty et al. (2014), which are similar to those reported in the class size literature (Krueger and Whitmore, 2001). These estimates imply that the average Midwestern student experiencing 2017-18 coal levels would score a full 0.016σ higher than a student facing 2007-08 coal use.[34,35] The complete elimination of coal-fired power production would further raise Midwestern test scores by 0.007σ.

The potential gains from the complete elimination of coal also vary substantially across states, especially as there are relatively few states in the country that still produce substantial amounts of electricity from coal. Highlighting a few states that can still achieve substantial test scores gains from the elimination of coal-fired power production (as of 2018): West Virginia (0.044σ), Nebraska (0.023σ), Wisconsin (0.017σ), Wyoming (0.015σ), Colorado (0.012σ), and Missouri (0.011σ). The spatial variation identified here gives policymakers a set of schools to focus on to improve test scores via mitigation efforts (e.g., closing coal plants, moving schools, or providing air filters).

Inequality: The decline in coal use also had considerable impacts on inequality since underprivileged students are more likely to attend schools in polluted regions (Currie, Voorheis, and Walker, 2020). In the Midwest, we estimate that the decline in coal use from 2006-07 to 2017-18 reduced the black-white test score gap by 0.023σ and the socioeconomic test score gap by 0.005σ.[36] Figure 7(b) exhibits the decline in the black-white test score gap by state due to the decline in coal use over the past decade.

[34] Chetty et al. (2014) only report fade out up to four years after the intervention. We set the fade out rate for years five onward at the same level as four years after the intervention, although given that fourth year fade out is eighty percent our results are not terribly sensitive to this assumption. For example, our cumulative effect for students at the end fifth grade is: $0.021 * (0.408 - 0.128) * (1 + 0.53 + 0.36 + 0.26 + 0.22 + 0.22)$. We then average these cumulative effects for each grade (implicitly assuming enrollment is identical across grades).

[35] Nationwide, we calculate that coal exposure dropped from 195,000 Mwh in 2006-07 to 62,000 Mwh in 2017-18, implying that the average U.S. student experiencing 2017-18 coal levels would score a full 0.008σ higher than a student facing 2007-08 coal use.

[36] The respective nationwide numbers are a 0.006σ decline in the black-white test score gap and a 0.003σ decline in the socioeconomic test score gap. The socioeconomic test score gap is defined as the difference in test scores between students who are and who are not socioeconomically disadvantaged according to their income eligibility for child nutrition programs.

6 Conclusion

The last decade has seen a precipitous decline in coal-fired power production, causing dramatic improvements in air quality across the United States. This paper quantifies the impact of the shift from coal-fired power production to cleaner sources such as natural gas on student performance.

To do so, we leverage three sources of variation to estimate the impact of coal-fired power plant emissions on student performance: (i) year-to-year production, (ii) persistent wind patterns, and (iii) plant closures. Quasi-experiments leveraging each source of variation indicate that every one million Mwh of increased coal-fired power plant production lowers student performance in schools within $10km$ by around 0.02σ. Nearly the entirety of this effect is concentrated in schools downwind from the power plant.

We find no such relationship between natural gas power production and student performance, indicating that the switch from coal to natural gas has generated substantial improvements in student achievement. The boost to cognition exhibits a high degree of spatial variation, with Midwestern states receiving substantial test score boosts while western states experienced little change. In particular, we calculate that the drop in coal-fired power production from 2006-07 to 2017-18 in the Midwest led to a test score increase of 0.016σ and a reduction in the black-white test score gap of 0.023σ.

Our estimates suggest a new lever that policymakers can use to improve the parlous state of education in the United States. Even though coal use has declined, more than 40 percent of people in the country still live in areas with unhealthy levels of air pollution (American Lung Association, 2019), indicating that air quality improvements or mitigation can generate further educational gains (Gilraine, 2020). In addition, given that underprivileged children are more likely to live in polluted areas (Currie et al., 2020), policies that reduce exposure to airborne pollutants can also serve to reduce the pervasive and persistent test score gaps in education.

References

Alexander, Diane and Janet Currie (2017), "Is it who you are or where you live? Residential segregation and racial gaps in childhood asthma." *Journal of Health Economics*, 55, 186–200.

Almond, Douglas, Janet Currie, and Valentina Duque (2018), "Childhood circumstances and adult outcomes: Act II." *Journal of Economic Literature*, 56, 1360–1446.

American Lung Association (2019), "The state of the air 2019." Technical report, URL https://www.lung.org/our-initiatives/healthy-air/sota/key-findings/.

Amster, Eric and Clara Lew Levy (2019), "Impact of coal-fired power plant emissions on children's health: A systematic review of the epidemiological literature." *International Journal of Environmental Research and Public Health*, 16, 2008.

Anderson, Michael L (forthcoming), "As the wind blows: The effects of long-term exposure to air pollution on mortality." *Journal of the European Economic Association*.

Barreca, Alan I., Matthew Neidell, and Nicholas J. Sanders (2017), "Long-run pollution exposure and adult mortality: Evidence from the Acid Rain Program." Working Paper 23524, National Bureau of Economic Research, URL http://www.nber.org/papers/w23524.

Beach, Brian and W. Walker Hanlon (2018), "Coal smoke and mortality in an early industrial economy." *Economic Journal*, 128, 2652–2675.

Bennion, H.S. (1940), "Electric power in american industry." *Military Engineer*, 32, 393–396.

Bharadwaj, Prashant, Matthew Gibson, Joshua Graff Zivin, and Christopher Neilson (2017), "Gray matters: Fetal pollution exposure and human capital formation." *Journal of the Association of Environmental and Resource Economists*, 4, 505–542.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2007), "From the cradle to the labor market? The effect of birth weight on adult outcomes." *Quarterly Journal of Economics*, 122, 409–439.

Burtraw, Dallas, Karen Palmer, Anthony Paul, and Matt Woerman (2012), "Secular trends, environmental regulations, and electricity markets." *Electricity Journal*, 25, 35–47.

Cesur, Resul, Erdal Tekin, and Aydogan Ulker (2013), "Air pollution and infant mortality: Evidence from the expansion of natural gas infrastructure." Working Paper 18736, National Bureau of Economic Research, URL http://www.nber.org/papers/w18736.

Chan, H. Ron, B. Andrew Chupp, Maureen L. Cropper, and Nicholas Z. Muller (2018), "The impact of trading on the costs and benefits of the Acid Rain Program." *Journal of Environmental Economics and Management*, 88, 180–209.

Chay, Kenneth Y. and Michael Greenstone (2003a), "Air quality, infant mortality, and the Clean Air Act of 1970." Working Paper 10053, National Bureau of Economic Research, URL http://www.nber.org/papers/w10053.

Chay, Kenneth Y. and Michael Greenstone (2003b), "The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession." *Quarterly Journal of Economics*, 118, 1121–1167.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011), "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *Quarterly Journal of Economics*, 126, 1593–1660.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff (2014), "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*, 104, 2633–79.

Clay, Karen, Joshua Lewis, and Edson Severnini (2016), "Canary in a coal mine: Infant mortality, property values, and tradeoffs associated with mid-20th century air pollution." Working Paper 22155, National Bureau of Economic Research, URL http://www.nber.org/papers/w22155.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker (2015), "Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings." *American Economic Review*, 105, 678–709.

Currie, Janet, Eric A. Hanushek, E. Megan Kahn, Matthew Neidell, and Steven G. Rivkin (2009a), "Does pollution increase school absences?" *Review of Economics and Statistics*, 91, 682–694.

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

Currie, Janet, John Voorheis, and Reed Walker (2020), "What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality." Working Paper 26659, National Bureau of Economic Research, URL http://www.nber.org/papers/w26659.

Currie, Janet and Reed Walker (2011), "Traffic congestion and infant health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3, 65–90.

Currie, Janet and Reed Walker (2019), "What do economists have to say about the Clean Air Act 50 years after the establishment of the Environmental Protection Agency?" *Journal of Economic Perspectives*, 33, 3–26.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell (2014), "What do we know about short-and long-term effects of early-life exposure to pollution?" *Annual Review of Resource Economics*, 6, 217–247.

Dahl, Gordon B. and Lance Lochner (2012), "The impact of family income on child achievement: Evidence from the Earned Income Tax Credit." *American Economic Review*, 102, 1927–56.

De Gouw, Joost A., David D. Parrish, Gregory J. Frost, and Michael Trainer (2014), "Reduced emissions of CO2, NOx, and SO2 from us power plants owing to switch from coal to natural gas with combined cycle technology." *Earth's Future*, 2, 75–82.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif (2019), "The mortality and medical costs of air pollution: Evidence from changes in wind direction." *American Economic Review*, 109, 4178–4219.

Deryugina, Tatyana, Nolan Miller, David Molitor, and Julian Reif (2020), "Geographic and socioeconomic heterogeneity in the benefits of reducing air pollution in the United States." In *Environmental and Energy Policy and the Economy, Volume 2* (Matthew Kotchen, James H. Stock, and Catherine Wolfram, eds.), University of Chicago Press.

Deschenes, Olivier, Michael Greenstone, and Joseph S. Shapiro (2017), "Defensive investments and the demand for air quality: Evidence from the NOx budget program." *American Economic Review*, 107, 2958–89.

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

Ebenstein, Avraham, Victor Lavy, and Sefi Roth (2016), "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution." *American Economic Journal: Applied Economics*, 8, 36–65.

Fell, Harrison and Daniel T. Kaffine (2018), "The fall of coal: Joint impacts of fuel prices and renewables on generation and emissions." *American Economic*

Journal: Economic Policy, 10, 90–116.

Garces, Eliana, Duncan Thomas, and Janet Currie (2002), "Longer-term effects of Head Start." *American Economic Review*, 92, 999–1012.

Genc, Sermin, Zeynep Zadeoglulari, Stefan H. Fuss, and Kursad Genc (2012), "The adverse effects of air pollution on the nervous system." *Journal of Toxicology*, 2012.

Gilraine, Michael (2020), "Air filters, pollution and student achievement." *Ed Working Paper*, 20-188, URL https://www.edworkingpapers.com/ai20-188.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010), "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings." *Journal of Political Economy*, 118, 536–598.

Grönqvist, Hans, J. Peter Nilsson, and Per-Olof Robling (2017), "Early lead exposure and outcomes in adulthood." Working Paper 2017:4, Institute for Evaluation of Labour Market and Education Policy.

Heft-Neal, Sam, Jennifer Burney, Eran Bendavid, Kara Voss, and Marshall Burke (2019), "Air pollution and infant mortality: Evidence from Saharan dust." Working Paper 26107, National Bureau of Economic Research, URL http://www.nber.org/papers/w26107.

Heissel, Jennifer, Claudia Persico, and David Simon (forthcoming), "The impact of accumulated and acute exposure to traffic pollution on child academic outcomes." *Journal of Human Resources*.

Heyes, Anthony, Matthew Neidell, and Soodeh Saberian (2016), "The effect of air pollution on investor behavior: Evidence from the S&P 500." Working Paper 22753, National Bureau of Economic Research, URL http://www.nber.org/papers/w22753.

Holland, Stephen P., Erin T. Mansur, Nicholas Z. Muller, and Andrew J. Yates (forthcoming), "Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation." *American Economic Journal: Economic Policy*.

International Energy Agency (IEA) (2019), *Global Energy & CO2 Status Report 2019*. London: International Energy Agency.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker (2017), "Every breath you take – every dollar you'll make: The long-term consequences of the Clean Air Act of 1970." *Journal of Political Economy*, 125, 848–902.

Jha, Akshaya and Nicholas Z. Muller (2017), "Handle with care: The local air pollution costs of coal storage." Working Paper 23417, National Bureau of Economic Research, URL http://www.nber.org/papers/w23417.

Joskow, Paul L. (2013), "Natural gas: From shortages to abundance in the United States." *American Economic Review*, 103, 338–43.

Kleinman, Michael T. and Arezoo Campbell (2014), "Central nervous system effects of ambient particulate matter: The role of oxidative stress and inflammation." Technical report, California Air Resources Board, URL https://ww2.arb.ca.gov/sites/default/files/classic//research/apr/past/08-306.pdf.

Knittel, Christopher R., Konstantinos Metaxoglou, and Andre Trindade (2015), "Natural gas prices and coal displacement: Evidence from electricity markets." Working Paper 21627, National Bureau of Economic Research, URL http://www.nber.org/papers/w21627.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders (2016), "Caution, drivers! Children present: Traffic, pollution, and infant health." *Review of Economics and Statistics*, 98, 350–366.

Krueger, Alan B. and Diane M. Whitmore (2001), "The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR." *Economic Journal*, 111, 1–28.

Lleras-Muney, Adriana (2010), "The needs of the army using compulsory relocation in the military to estimate the effect of air pollutants on children's health." *Journal of Human Resources*, 45, 549–590.

Milligan, Kevin and Mark Stabile (2011), "Do child tax benefits affect the well-being of children? evidence from Canadian child benefit expansions." *American Economic Journal: Economic Policy*, 3, 175–205.

Muller, Nicholas Z. and Robert Mendelsohn (2009), "Efficient pollution regulation: Getting the prices right." *American Economic Review*, 99, 1714–39.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus (2011), "Environmental accounting for pollution in the United States economy." *American Economic Review*, 101, 1649–75.

Murphy, Joshua (2017), "The value of reducing power plant emissions: New evidence from the Clean Air Interstate Rule." Technical report, University of Toronto.

National Oceanic and Atmospheric Administration (2018), "Global hourly surface data." URL https://www7.ncdc.noaa.gov/CDO/cdopoemain.cmd?datasetabbv=DS3505&countryabbv=&georegionabbv=&resolution=40.

Nature (2009), "The shale revolution." *Nature*, 551–552.

Neidell, Matthew J. (2004), "Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma." *Journal of Health Economics*, 23, 1209–1236.

North Carolina Education Research Data Center (1995-2017), "Student, class and personnel files." URL http://childandfamilypolicy.duke.edu/research/hc-education-data-center/.

Persico, Claudia L. and Joanna Venator (2019), "The effects of local industrial pollution on students and schools." *Journal of Human Resources*, 0518–9511R2.

Pope, C. Arden and Douglas W. Dockery (2006), "Health effects of fine particulate air pollution: Lines that connect." *Journal of the Air & Waste Management Association*, 56, 709–742.

Ransom, Michael R. and C. Arden Pope (1995), "External health costs of a steel mill." *Contemporary Economic Policy*, 13, 86–97.

Roth, Sefi (2019), "The effect of indoor air pollution on cognitive performance: Evidence from the UK.", URL http://personal.lse.ac.uk/roths/JMP.pdf. Unpublished.

Sanders, Nicholas J. (2012), "What doesn't kill you makes you weaker: Prenatal pollution exposure and educational outcomes." *Journal of Human Resources*, 47, 826–850.

Schiavo, Joseph G. and Robert Mendelsohn (2019), "The effect of domestic air pollution mitigation and fracking on retirements of coal power plants." *Climate Change Economics*, 10, 1950008.

Schlenker, Wolfram and W. Reed Walker (2016), "Airports, air pollution, and contemporaneous health." *Review of Economic Studies*, 83, 768–809.

Schneider, Conrad G. and Jonathan M. Banks (2010), "The toll from coal: An updated assessment of death and disease from America's dirtiest energy source." Technical report, Clean Air Task Force, URL https://www.catf.us/resource/the-toll-from-coal/.

Simeonova, Emilia, Janet Currie, Peter Nilsson, and Reed Walker (forthcoming), "Congestion pricing, air pollution, and urban health." *Journal of Human Resources*.

United States Department of Education (2019), "Edfacts data files." URL https://www2.ed.gov/about/inits/ed/edfacts/data-files/index.html.

United States Energy Information Administration (EIA) (2020), “Form EIA-923 detailed data with previous form data (EIA-906/920).” URL https://www.eia.gov/electricity/data/eia923/.

US Environmental Protection Agency (2015), “Environmental protection agency. latest findings on national air quality: 2018 status and trends.” Technical report, URL https://19january2017snapshot.epa.gov/cleanpowerplan/clean-power-plan-final-rule-regulatory-impact-analysis_.html.

US Environmental Protection Agency (2018), “Latest findings on national air quality: 2018 status and trends.” Technical report, URL https://www.epa.gov/air-trends.

Venkatesh, Aranya, Paulina Jaramillo, W. Michael Griffin, and H. Scott Matthews (2012), “Implications of changing natural gas prices in the United States electricity sector for SO2, NOX and life cycle GHG emissions.” *Environmental Research Letters*, 7, 034018.

Ward, Courtney J. (2015), “It’s an ill wind: The effect of fine particulate air pollution on respiratory hospitalizations.” *Canadian Journal of Economics/Revue Canadienne D’économique*, 48, 1694–1732.

World Health Organization (2017), “Evolution of WHO air quality guidelines: Past, present and future.” Technical report.

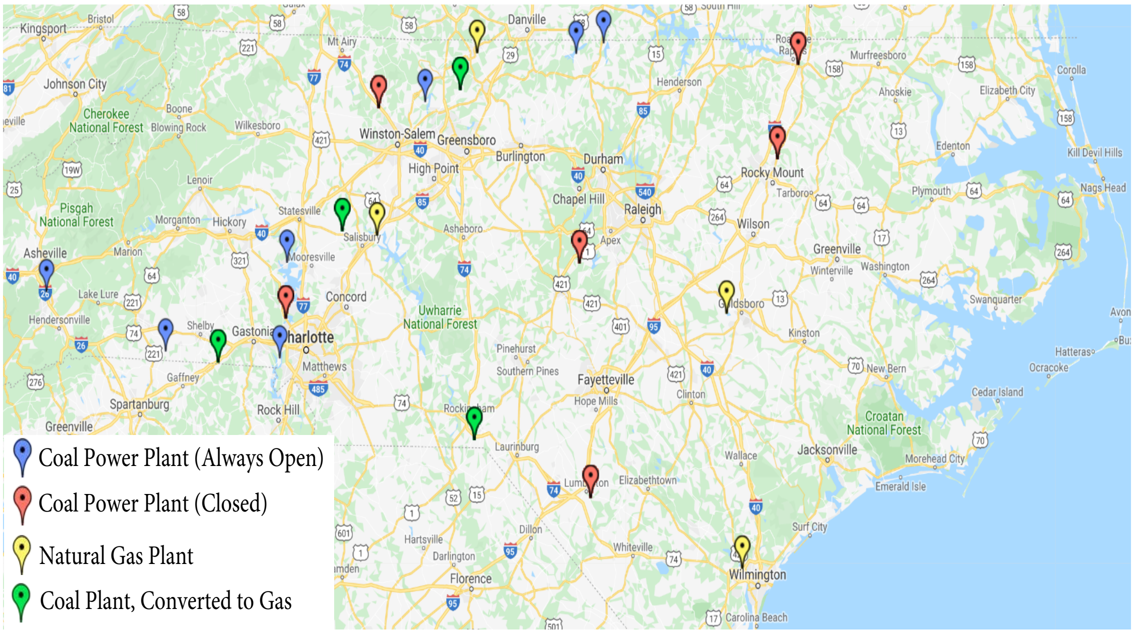
Yang, Muzhe and Shin-Yi Chou (2015), “Impacts of being downwind of a coal-fired power plant on infant health at birth: Evidence from the precedent-setting Portland rule.” Working Paper 21723, National Bureau of Economic Research, URL http://www.nber.org/papers/w21723.

Zhang, Qiang, Xujia Jiang, Dan Tong, Steven J. Davis, Hongyan Zhao, Guannan Geng, Tong Feng, Bo Zheng, Zifeng Lu, David G. Streets, et al. (2017), “Transboundary health impacts of transported global air pollution and international trade.” *Nature*, 543, 705–709.

Zhang, Yanxu, Daniel J. Jacob, Hannah M. Horowitz, Long Chen, Helen M Amos, David P. Krabbenhoft, Franz Slemr, Vincent L. St. Louis, and Elsie M. Sunderland (2016), “Observed decrease in atmospheric mercury explained by global decline in anthropogenic emissions.” *Proceedings of the National Academy of Sciences*, 113, 526–531.

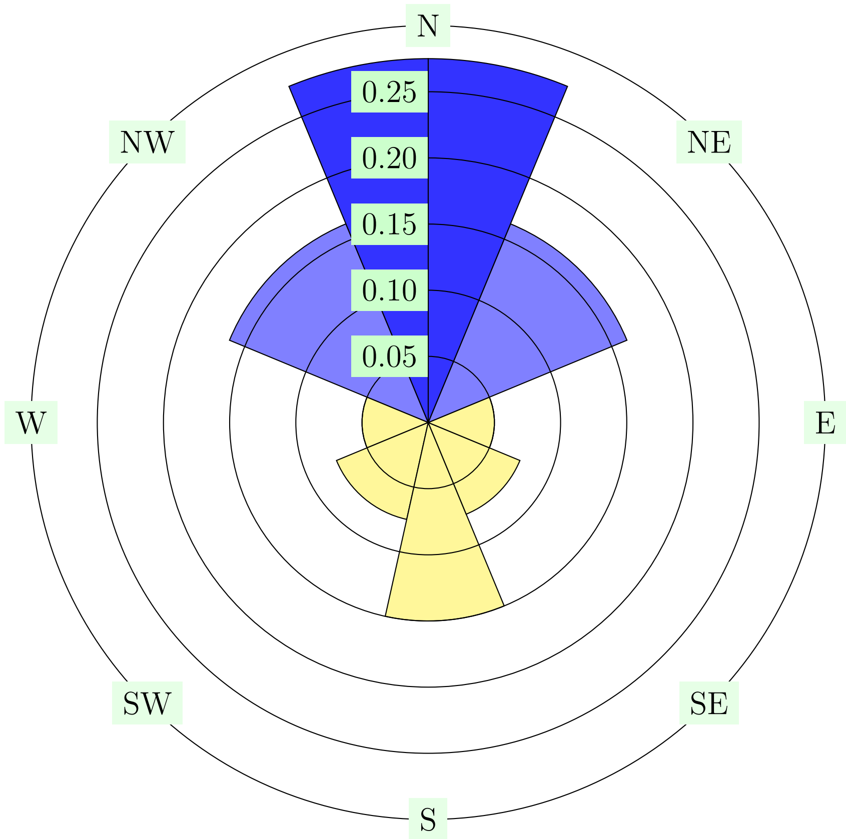
Zweig, Jacqueline S., John C. Ham, and Edward L. Avol (2009), “Air pollution and academic performance: Evidence from California schools.” *National Institute of Environmental Health Sciences*, 1–35.

Figure 1: Large Coal and Natural Gas Plants in North Carolina 2001-2017



Notes: This figure shows the geographic location of North Carolina's twenty-one large coal and natural gas plants. Out of seventeen coal plants in our sample, four converted to natural gas (denoted by green marker) and six closed outright (denoted by the red marker). We therefore have a total of seventeen coal plants (of which six closed and four converted to natural gas) and eight natural gas plants operating over our period of interest (2001-2017). Table A.1 reports the period of operation for these plants, along with distance to the nearest wind station, production statistics, and their closure or conversion date.

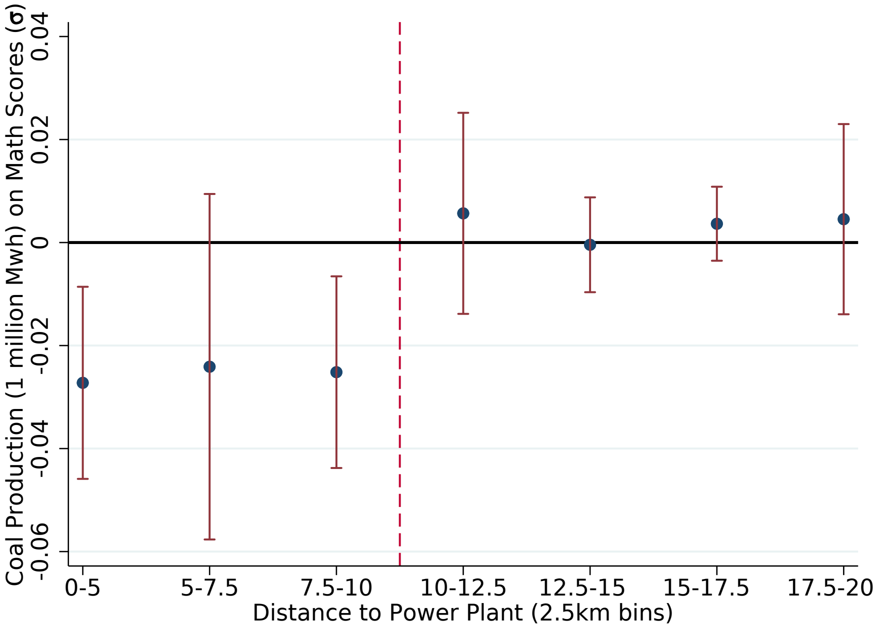
FIGURE 2: Constructing the 'Downwind' Measure: Example



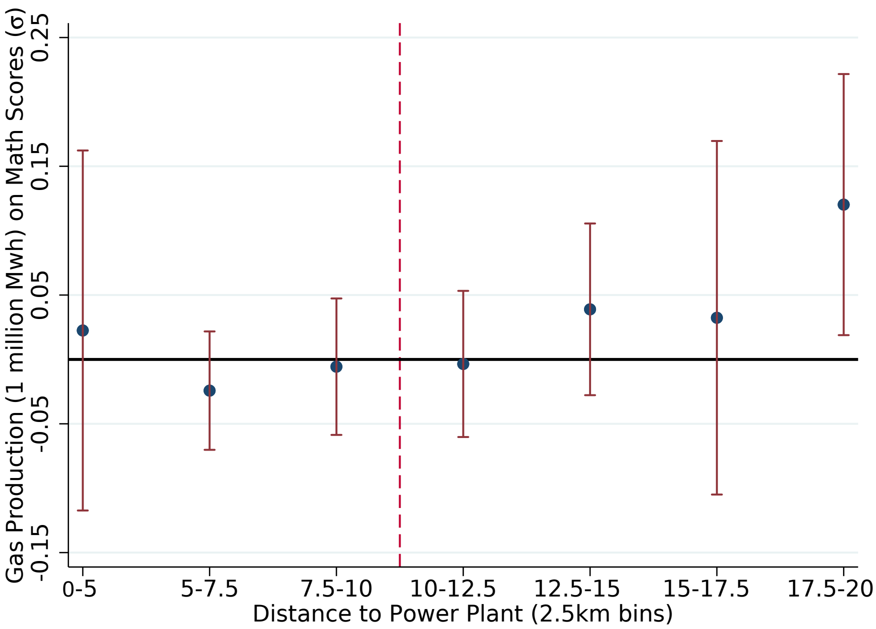
Notes: This figure provides an example of a typical wind rose in North Carolina. The circular format of the wind rose shows the direction the wind is blowing *toward* and the length of each circle segment around the circle shows the proportion of time the wind blows toward that particular direction during school hours. For instance, the dark blue shaded 'spoke' indicates that the wind blows towards the North 27.5 percent of the time. The 'downwind' measure for a school located due North of the power plant incorporates all three blue shaded spokes (N, NW, NE), with the lighter blue shaded spokes (NW, NE) being given half weight. The school's 'downwind' measure is thus 0.4375 ($= 0.275 + \frac{1}{2}0.1625 + \frac{1}{2}0.1625$), which is relatively high. Likewise, a school located due East of the power plant has a relatively low 'downwind' measure of only 0.16875 ($= 0.05 + \frac{1}{2}0.1625 + \frac{1}{2}0.075$).

Figure 3: Effect of Power Production on Student Mathematics Scores by Distance and Fuel Type

(a) Coal-Fired Power Production



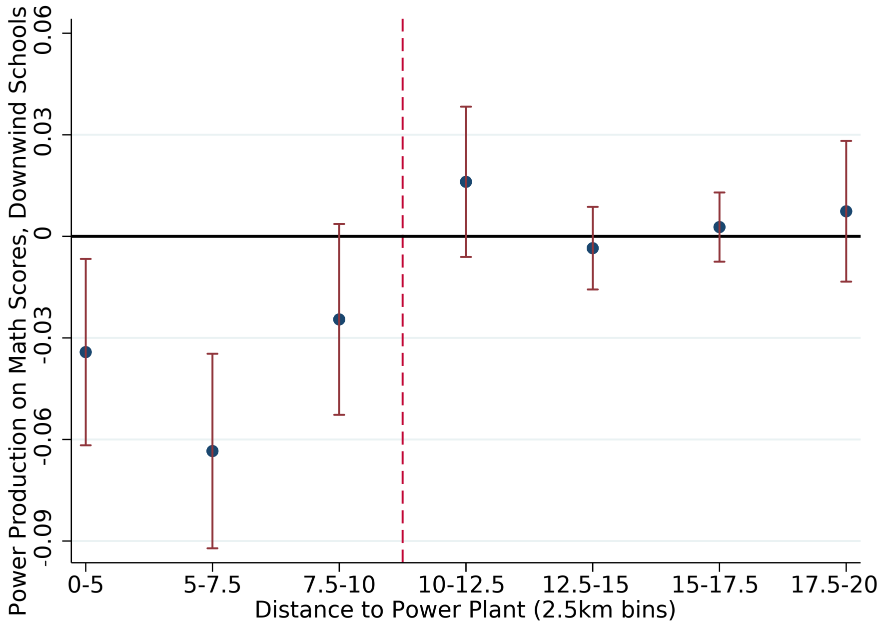
(b) Natural Gas Fired Power Production



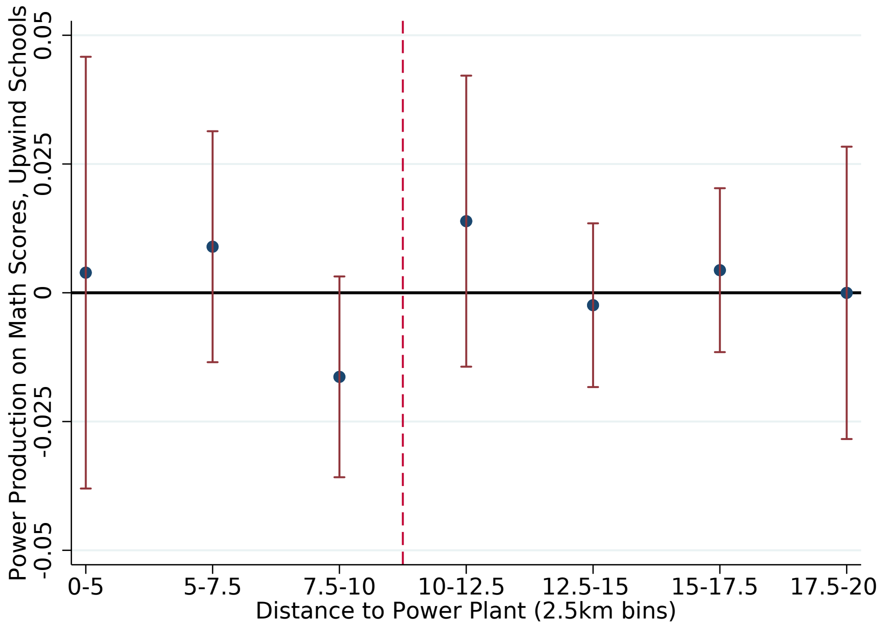
Notes: These figures use year-to-year power plant production variation to look at the effect of power production on test scores by estimating equation (1) for different distance bins and by fuel type. Effect sizes are in terms of standard deviations of the student test score distribution. Each point represents the result of a separate regression of equation (1) and includes controls for demographics and student, school, and grade-year fixed effects. Given these controls, the figures mimic the results from column (2) of Table 3. The two bins less than $5km$ are combined as there are few observations within 2.5km of a power plant. The solid horizontal line denotes a point estimate of zero while the vertical dashed line delineates observations closer to 10 kilometers from the power plant, which are the observations used in the main analysis. The whiskers represent 90 percent confidence intervals with standard errors clustered at the school level.

FIGURE 4: Effect of Coal-Fired Power Production on Student Mathematics Scores by Distance and whether School is 'Downwind' or 'Upwind' of Power Plant

(b) 'Downwind' of Coal-Fired Power Plant



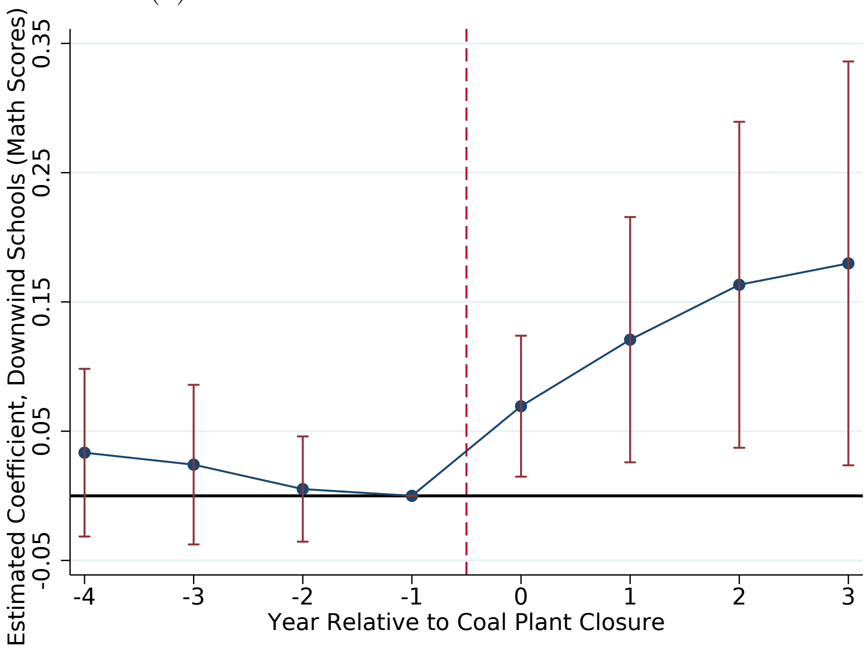
(b) 'Upwind' of Coal-Fired Power Plant



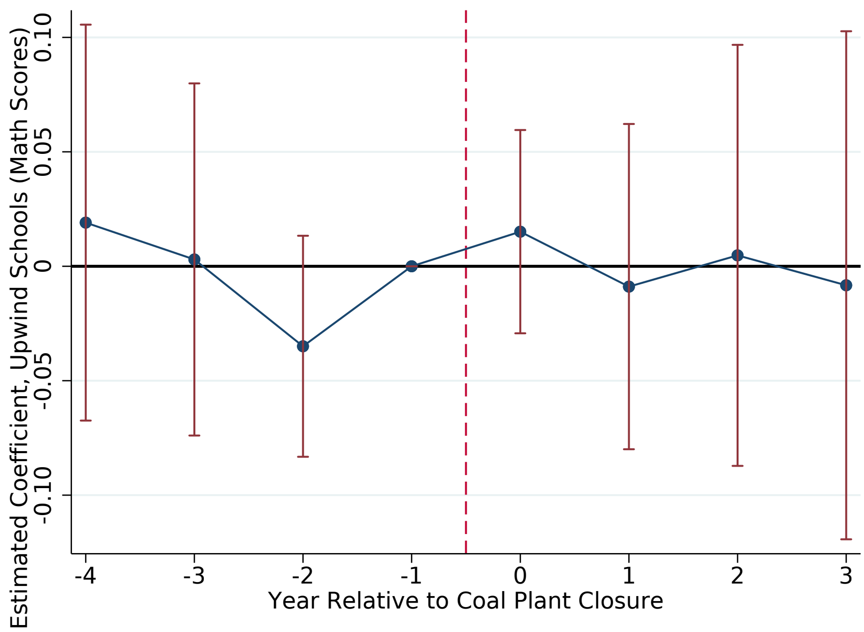
Notes: Figures estimate the effect of coal-fired power production on student performance by estimating equation (1) for different distance bins. Results are subdivided by whether the school is 'downwind' (Panel A) or 'upwind' (Panel B) of the coal-fired plant. 'Downwind' status is defined as facing above median wind, while 'upwind' schools face below median wind levels according to our downwind measure (see equation (6)) with the sample split occurring at the school level so that there are an equal number of schools in both panels for each distance bin. Effect sizes are in terms of standard deviations of the student test score distribution. Each point represents the result of a separate regression of equation (1) and includes controls for student demographics and student, school, and grade-year fixed effects. Given these controls, the figures mimic the results from column (2) of Table 4. The two bins less than $5km$ are combined as there are few observations within 2.5km of a power plant. The solid horizontal line denotes a point estimate of zero while the vertical dashed line delineates observations closer to 10 kilometers from the power plant, which are the observations used in the main analysis. The whiskers represent 90 percent confidence intervals with standard errors clustered at the school level.

Figure 5: Event-Study Estimates of the Effects of Coal Plant Closures on Students' Test Scores by Downwind and Upwind Status

(a) 'Downwind' of Coal-Fired Power Plant



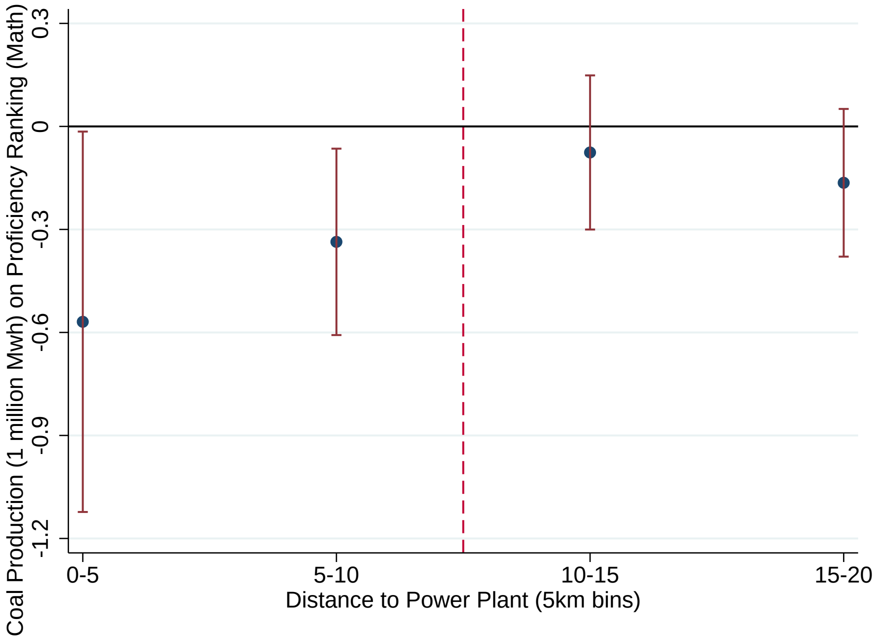
(b) 'Upwind' of Coal-Fired Power Plant



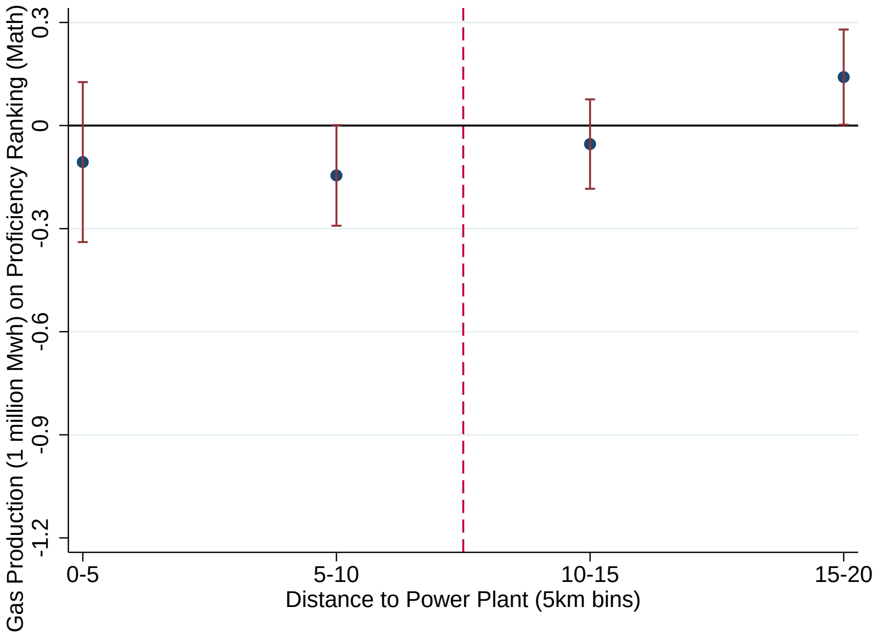
Notes: The above figures plot estimates from equation (4) and show the evolution of student performance in the years leading up to and after the plant closure. Results are subdivided by whether the school is 'downwind' (Panel A) or 'upwind' (Panel B) of the coal-fired plant. 'Downwind' status is defined as facing above median wind, while 'upwind' schools face below median wind levels according to our downwind measure (see equation (6)) with the sample split occurring at the school level so that there are an equal number of schools in both panels. Regressions include controls for student demographics and student, school, and grade-year fixed effects. Given these controls, the figures mimic the results from column (2) of Table 5. The solid horizontal line denotes a point estimate of zero while the vertical dashed line delineates the periods before and after the coal-fired power plant closed. The whiskers represent 90 percent confidence intervals with standard errors clustered at the school level.

Figure 6: Effect of Power Production on Mathematics Proficiency by Distance and Source: **National** Data

(a) Coal-Fired Power Production



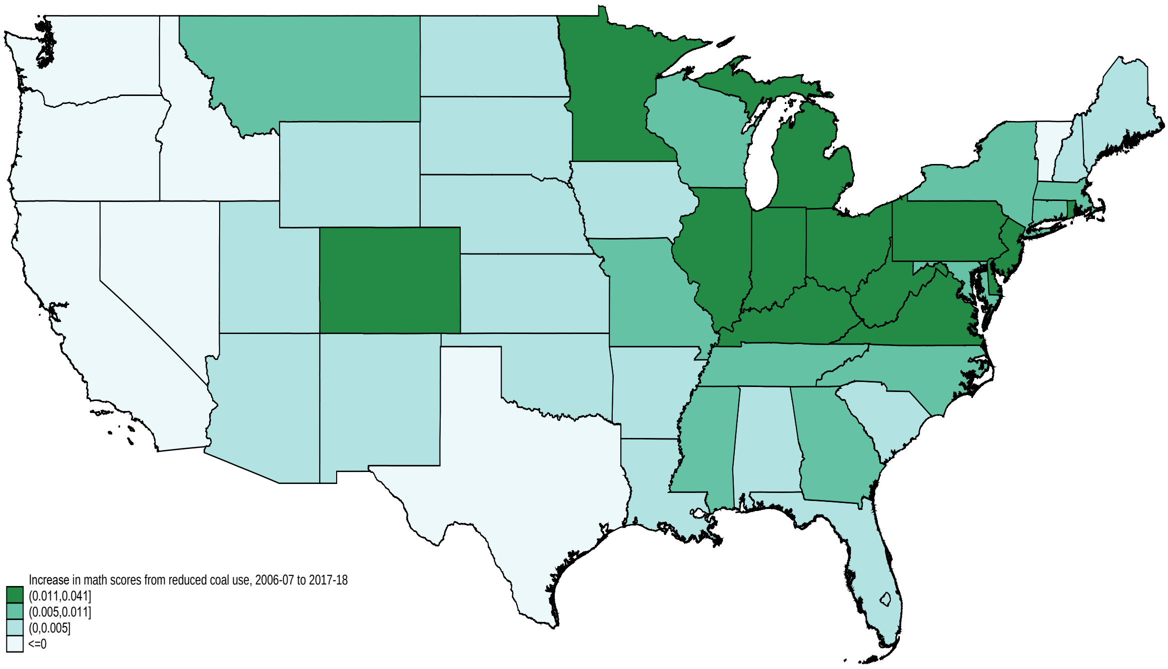
(b) Natural Gas Fired Power Production



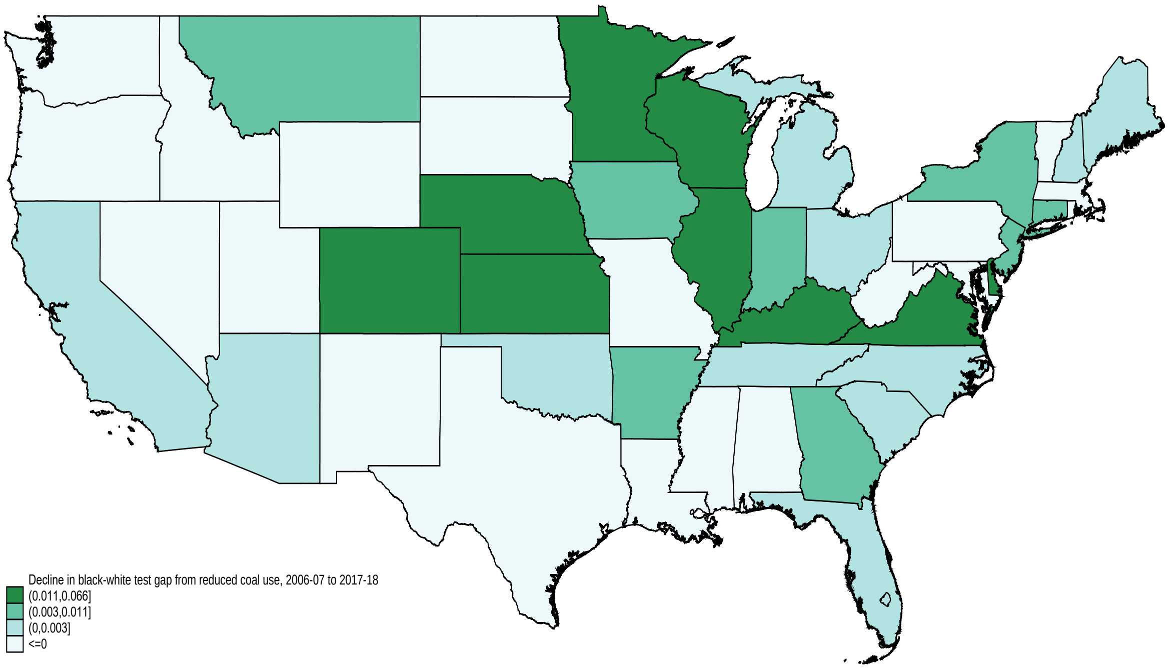
Notes: These figures use year-to-year power plant production variation to look at the effect of power production on test scores by estimating equation (7) which identifies the effect of a one Mwh increase in power production on student performance using the nationwide data (see Appendix B.1). The outcome variable is defined as the school's ranking based on their school wide proficiency rate in their state. Each point represents the result of a separate regression of equation (7) and includes controls for lag state proficiency rank, school demographics, and school and year fixed effects. Given these controls, the figures mimic the results from column (3) of Table B.6. The solid horizontal line denotes a point estimate of zero while the vertical dashed line delineates observations closer to 10 kilometers from the power plant, which are the observations used in the main analysis. The whiskers represent 90 percent confidence intervals with standard errors clustered at the school level. The scale of the y-axis is the same across both figures for comparability.

Figure 7: Spatial Variation in Impact of Reduced Coal Use from 2006-07 to 2017-18

(a) Aggregate Achievement (Mathematics)



(b) Black-White Test Score Gap (Mathematics)



Notes: These figures show spatial variation in the nationwide impact of reduced coal usage from 2006-07 through 2017-18. Figure 7(a) shows the aggregate improvement in student achievement by state, while Figure 7(b) shows the decline in the black-white test score gap by state. Effect sizes are in terms of standard deviations of the student test score distribution. We determine aggregate effects by calculating the average enrollment-weighted change in coal-fired power production within $10km$ of schools from 2006-07 to 2017-18 and multiply that number by our estimated effect of 0.02σ and then calculate cumulative impacts using fade out estimates from Chetty et al. (2014). For the black-white test score gap, we similarly calculate the average enrollment-weighted change in coal-fired power production within $10km$ of schools for black students relative to white students from 2006-07 to 2017-18 and multiply that number by our cumulative estimated effect. Alaska and Hawaii are excluded for visual clarity.

Table 1. Summary Statistics

	Full Sample[1] (1)	Within 10 km of Coal Plant (2)	Within 10 km of Gas Plant (3)	Event Study Sample (4)
Mean of Student Characteristics				
Mathematics Score (σ)	0.00	-0.01	-0.28	-0.22
Reading Score (σ)	0.00	-0.03	-0.25	-0.21
Lagged Mathematics Score (σ)[2]	0.02	0.01	-0.25	-0.19
Lagged Reading Score (σ)[2]	0.02	0.00	-0.23	-0.18
% White	54.6	55.7	49.3	44.7
% Black	27.6	30.4	34.0	39.8
% Hispanic	10.8	7.3	10.3	7.9
% Asian	2.5	2.3	0.8	1.5
% Socioeconomically Disadvantaged	47.9	46.8	64.0	59.5
% English Learners	4.4	3.2	3.2	3.7
% with Disability	9.7	7.0	12.4	9.5
% Gifted	15.4	13.5	13.7	9.6
% Repeating Grade	1.0	1.1	0.8	1.2
Distance to Power Plant (km)	36.8	6.8	6.2	6.7
Downwind Measure	0.24	0.24	0.25	0.24
# of Plants	21	17	8	10
# of Students	2,509,400	178,065	31,606	78,251
Observations (student-year)	9,247,841	469,095	72,335	191,965

[1] Data coverage: grades 3-8 from 2000-01 through 2016-17. Third grade lagged test scores are not available after 2008-09 (inclusive) due to the end of the grade 3 pre-test nor during 2005-06 (mathematics) and 2007-08 (English). Grade-years lacking lagged own subject test scores are dropped. Summary statistics are reported for the mathematics sample.

[2] Lagged test scores are generally missing for about ten percent of the sample.

Table 2. Checking for Sorting on Observables

	Lagged Mathematics (1)	Lagged English (2)	Percent White (3)	Percent Black (4)	Percent Hispanic (5)	Percent Disadvantaged (6)	Percent EL (7)
A. Coal-fired Power Plant Production							
Production	0.010	0.004	-0.19	0.08	-0.16	0.57	-0.05
(1 million Mwh)	(0.011)	(0.009)	(0.29)	(0.21)	(0.27)	(0.48)	(0.09)
B. Coal-fired Power Plant Production by Wind Measure							
Production	0.007	0.004	0.15	0.00	-0.30	0.17	-0.07
$*$(downwind$\div$0.15)	(0.007)	(0.005)	(0.19)	(0.12)	(0.18)	(0.26)	(0.06)
Observations	434,602	428,055	468,728	468,728	468,728	458,561	467,135
# of Students	162,923	161,665	177,990	177,990	177,990	175,292	177,740

Notes: This table sheds light on whether there is year-to-year sorting on observables based on nearby production of the coal-fired power plant. This ensures that the composition of schools does not change systematically with year-to-year production variation – i.e., our 'treatment.' Specifically, we regress a given covariate on the 'treatment' (i.e., production or production times downwind status) along with school, grade, and year fixed effects. Effectively, Panel A estimates equation (1) while Panel B reports results from equation (2) with our outcome being replaced by a covariate. EL is an acronym for English learners, while 'disadvantaged' refers to socioeconomically disadvantaged. We divide our continuous wind measure by 0.15 for consistency with the other tables. The number of observations changes in each column as covariates are sometimes missing. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

TABLE 3. Effect of Power Production Using Production Variation by Plant Type

	Mathematics Scores (σ)			*English Scores (σ)*		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Coal-fired Plants						
Production	-0.021***	-0.021***	-0.019**	-0.008*	-0.009**	-0.009**
(1 million Mwh)	(0.009)	(0.008)	(0.008)	(0.004)	(0.004)	(0.004)
Observations	469,095	469,095	434,602	466,947	466,947	433,020
# of Students	178,065	178,065	162,923	177,067	177,067	161,942
B. Natural Gas Plants						
Production	0.012	0.007	0.006	0.000	-0.008	-0.007
(1 million Mwh)	(0.016)	(0.015)	(0.019)	(0.008)	(0.008)	(0.006)
Observations	72,335	72,335	68,479	72,266	72,266	68,305
# of Students	31,606	31,606	29,694	31,584	31,584	29,654
Demographics Controls	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes

Notes: This table reports estimates from equation (1) which identifies the effect of a one Mwh increase in power production on student performance. Panel A reports estimates for schools nearby a coal-fired power plant, while Panel B does so for schools neighboring a natural gas power plant. The sample is restricted to schools within $10km$ of a power plant and to years that the power plant was in operation. Table A.1 reports the power plants used in the analysis by type and their period of operation. Effect sizes are in terms of standard deviations of the student test score distribution. Each cell represents the result of a separate regression of equation (1) and includes student, school, and grade-year fixed effects. 'Demographic controls' include student level demographics and mean demographics at the school-grade level and include: ethnicity, gender, socioeconomically disadvantaged status, disability status, gifted status, English learner status, and grade repeating status. Missing indicators are used for students with missing demographics. 'Lagged test scores' consist of a cubic polynomial in prior lagged mathematics and reading scores interacted with grade dummies along with average lagged school-grade test scores. Students with missing lagged own-subject scores are dropped while a missing indicator is used for those missing a lagged other subject score. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 4. Effect of **Coal** Power Plant Production Using Production and Wind Variation

	Mathematics Scores (σ)			*English Scores (σ)*		
	(1)	(2)	(3)	(4)	(5)	(6)
A. 'Downwind' from Coal-fired Plant						
Production	-0.033***	-0.033**	-0.033***	-0.009	-0.009	-0.011*
(1 million Mwh)	(0.010)	(0.010)	(0.010)	(0.007)	(0.007)	(0.007)
B. 'Upwind' from Coal-fired Plant						
Production	-0.004	-0.004	-0.002	-0.005	-0.006	-0.006
(1 million Mwh)	(0.012)	(0.012)	(0.012)	(0.005)	(0.005)	(0.005)
C. Continuous Wind Measure						
Production	-0.030**	-0.031**	-0.030**	-0.018**	-0.020**	-0.021**
$*$(downwind÷0.15)	(0.014)	(0.014)	(0.014)	(0.008)	(0.009)	(0.009)
Demographics Controls	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes
Observations	469,095	469,095	434,602	466,947	466,947	433,020
# of Students	178,065	178,065	162,923	177,067	177,067	161,942

Notes: Panels A and B contrast the effect of increased coal-fired power production for schools that are 'downwind' relative to 'upwind' of the coal-fired power plant according to our downwind measure (see equation (6)). Specifically, Panel A estimates equation (1) for 'downwind' schools facing above median wind, while Panel B does so for 'upwind' schools with below median wind levels. Splitting the sample into 'downwind' and 'upwind' schools is done at the school level so that there are an equal number of schools in both panels. Panel C reports results from equation (2) which uses year-to-year variation in coal-fired power production along with across school variation in our continuous wind measure. We divide our continuous wind measure by 0.15 to make our estimates comparable to the contrast between Panel A and B as, on average, our wind measure is about 0.15 units higher for 'downwind' schools relative to 'upwind' schools. Effect sizes in Panels A and B report the change in standardized test scores for a one Mwh increase in power production. Given the additional differencing layer, Panel C reports the change in standardized test scores for a one Mwh increase in power production *combined* with a 0.15-unit increase in our wind measure. The sample is restricted to schools within $10km$ of a power plant and to years that the power plant was in operation (as a coal plant). Table A.1 reports the power plants and their period of operation. Each cell represents the result of a separate regression and includes student, school, and grade-year fixed effects. 'Demographic controls' include student level demographics and mean demographics at the school-grade level and include: ethnicity, gender, socioeconomically disadvantaged status, disability status, gifted status, English learner status, and grade repeating status. Missing indicators are used for students with missing demographics. 'Lagged test scores' consist of a cubic polynomial in prior lagged mathematics and reading scores interacted with grade dummies along with average lagged school-grade test scores. Students with missing lagged own-subject scores are dropped while a missing indicator is used for those missing a lagged other subject score. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table 5. Event-Study of Coal-Fired Plant Closures

	Mathematics Scores (σ)			*English Scores (σ)*		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Event-Study						
post	0.065***	0.060***	0.048**	0.056***	0.055***	0.061***
	(0.022)	(0.021)	(0.021)	(0.016)	(0.016)	(0.015)
B. 'Downwind' from Coal-fired Plant						
post	0.093***	0.090**	0.103***	0.076**	0.069**	0.081**
	(0.035)	(0.035)	(0.036)	(0.033)	(0.034)	(0.032)
C. 'Upwind' from Coal-fired Plant						
post	0.031	0.031	0.016	0.038*	0.041**	0.052***
	(0.028)	(0.027)	(0.022)	(0.022)	(0.020)	(0.018)
Demographics Controls	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes
Observations	191,965	191,965	159,200	191,338	191,338	158,741
# of Students	78,251	78,251	66,879	78,202	78,202	66,648

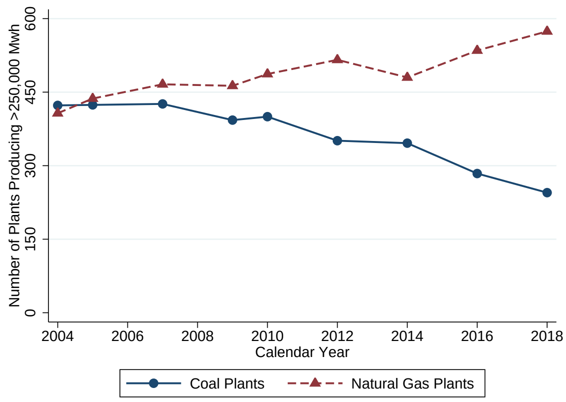
Notes: Panel A estimates results from the event study quasi-experiment defined by equation (3). The event-study utilizes ten coal plant closures and conversions; the names of these plants along with the date of the event are reported in Table A.1. Panels B and C estimate results from the event study quasi-experiment defined by equation (3) for schools that are 'downwind' relative to 'upwind' of the coal-fired power plant. Specifically, Panels B and C estimate equation (3) for 'downwind' schools facing above median wind and 'upwind' schools with below median wind levels according to our downwind measure (see equation 6). Splitting the sample into 'downwind' and 'upwind' is done at the school level so that there are an equal number of schools in both panels. The sample is restricted to schools within $10km$ of one of the ten power plants that closed or converted to natural gas during our period of study (2001-2017). Each cell represents the result of a separate regression of equation and includes student, school, and grade-year fixed effects. 'Demographic controls' include student level demographics and mean demographics at the school-grade level and include: ethnicity, gender, socioeconomically disadvantaged status, disability status, gifted status, English learner status, and grade repeating status. Missing indicators are used for students with missing demographics. 'Lagged test scores' consist of a cubic polynomial in prior lagged mathematics and reading scores interacted with grade dummies along with average lagged school-grade test scores. Students with missing lagged own-subject scores are dropped while a missing indicator is used for those missing a lagged other subject score. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Online Appendix

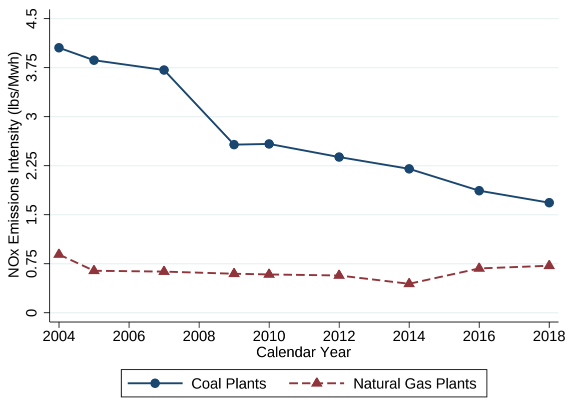
A Appendix Figures and Tables

FIGURE A.1: Power Plant Numbers and Emissions Intensity over Time by Fuel Type: United States

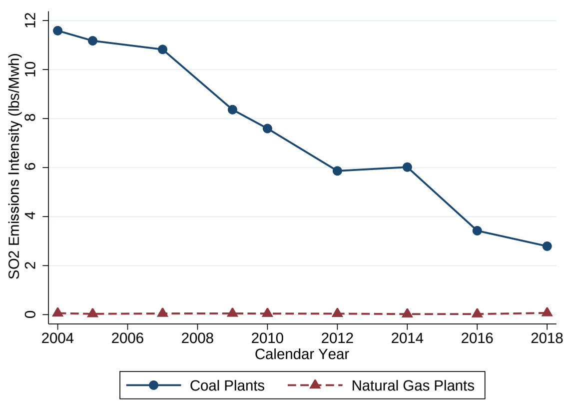
(a) Number of Power Plants by Fuel Type



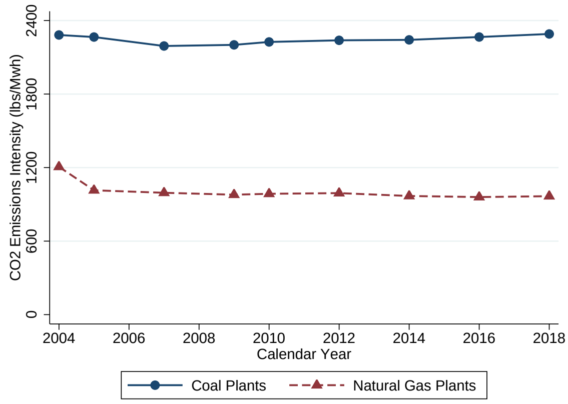
(b) Nitrogen Oxides (NO_x) Emissions Intensity by Fuel Type



(c) Sulfur Dioxide (SO_2) Emissions Intensity by Fuel Type



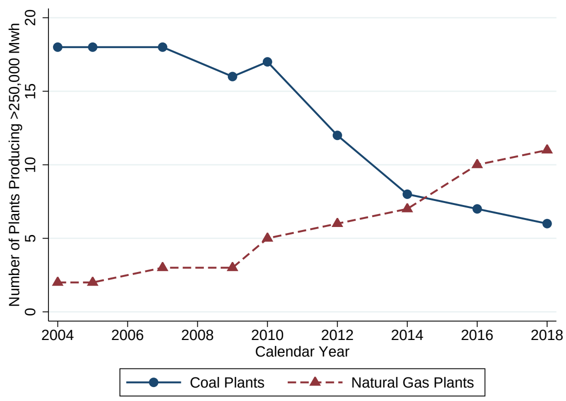
(d) Carbon Dioxide (CO_2) Emissions Intensity by Fuel Type



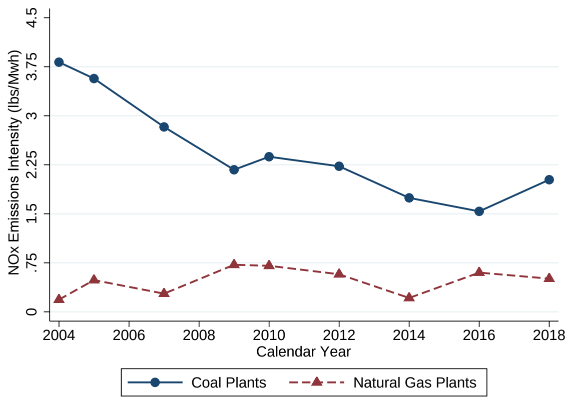
Notes: These figures plot number of large plants (i.e., those producing more than 250,000 Mwh of electricity in a given year) using coal or natural gas as their primary fuel source along with the emissions intensity of these plants in terms of NO_x, SO_2, and CO_2. Primary fuel source is coded as producing greater than 75% of electricity using that fuel type. About thirty plants per year that meet the electricity production criteria are omitted as they produce electricity using roughly equal amounts of gas and coal and so a primary fuel source cannot be coded. Data only cover roughly every second year (2004, 2005, 2007, 2009, 2010, 2012, 2014, 2016, 2018). They come from the Emissions & Generation Resource Integrated Database (eGRID) and are available at https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid. Emission figures are on the same scale as those in Figure A.2 so that nationwide emissions intensity over time can be compared to emissions intensity in North Carolina.

Figure A.2: Power Plant Numbers and Emissions Intensity over Time by Fuel Type: North Carolina

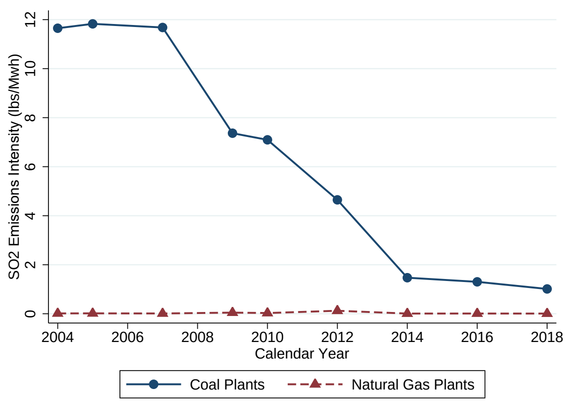
(a) Number of Power Plants by Fuel Type



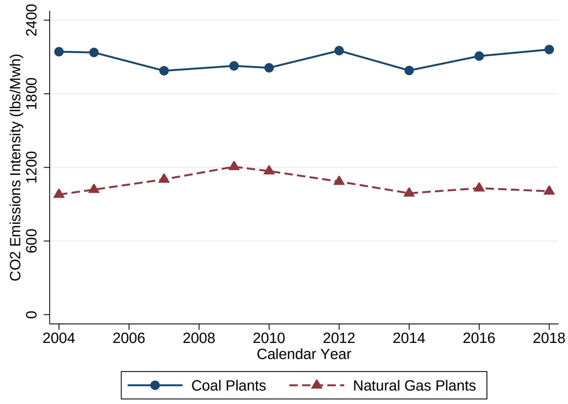
(b) Nitrogen Oxides (NO_x) Emissions Intensity by Fuel Type



(c) Sulfur Dioxide (SO_2) Emissions Intensity by Fuel Type

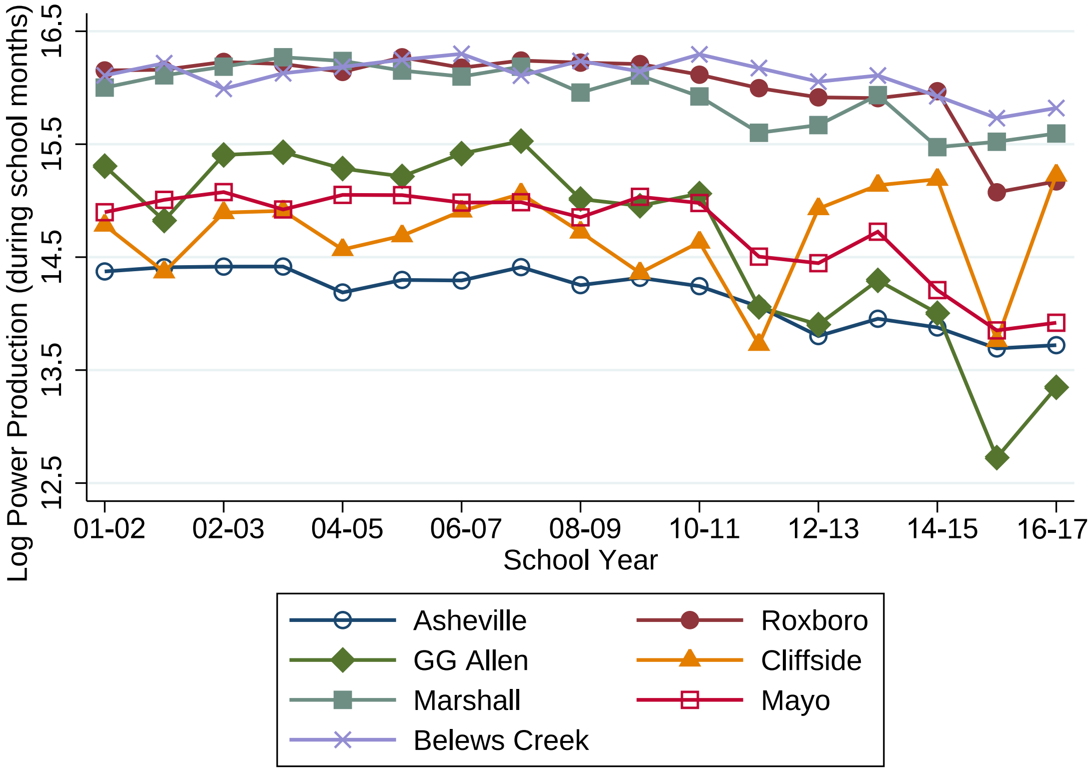


(d) Carbon Dioxide (CO_2) Emissions Intensity by Fuel Type



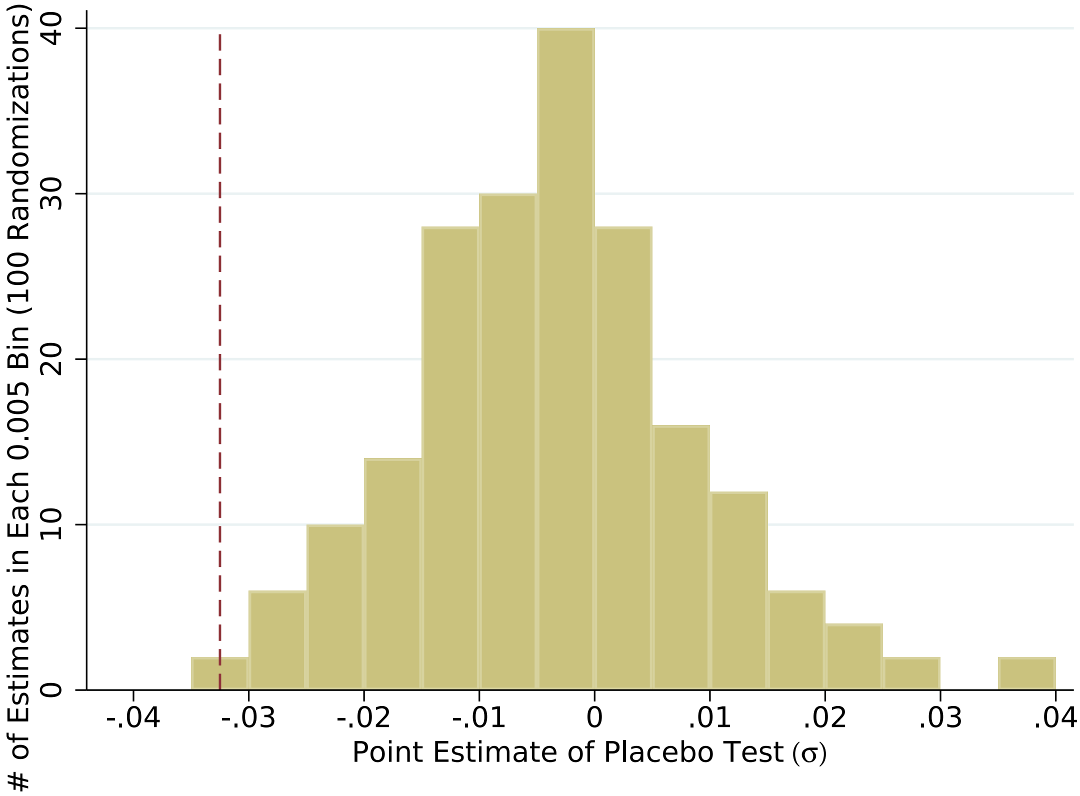
Notes: These figures plot number of large plants in North Carolina (i.e., those producing more than 250,000 Mwh of electricity in a given year) using coal or natural gas as their primary fuel source along with the emissions intensity of these plants in terms of NO_x, SO_2, and CO_2. Primary fuel source is coded as producing greater than 75% of electricity using that fuel type. Data only cover roughly every second year (2004, 2005, 2007, 2009, 2010, 2012, 2014, 2016, 2018) and come from the Emissions & Generation Resource Integrated Database (eGRID) and are available at https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid. Emission figures are on the same scale as those in Figure A.1 so that efficiency over time in North Carolina can be compared to nationwide efficiency.

FIGURE A.3: Power Production by Individual Coal Plants over Time



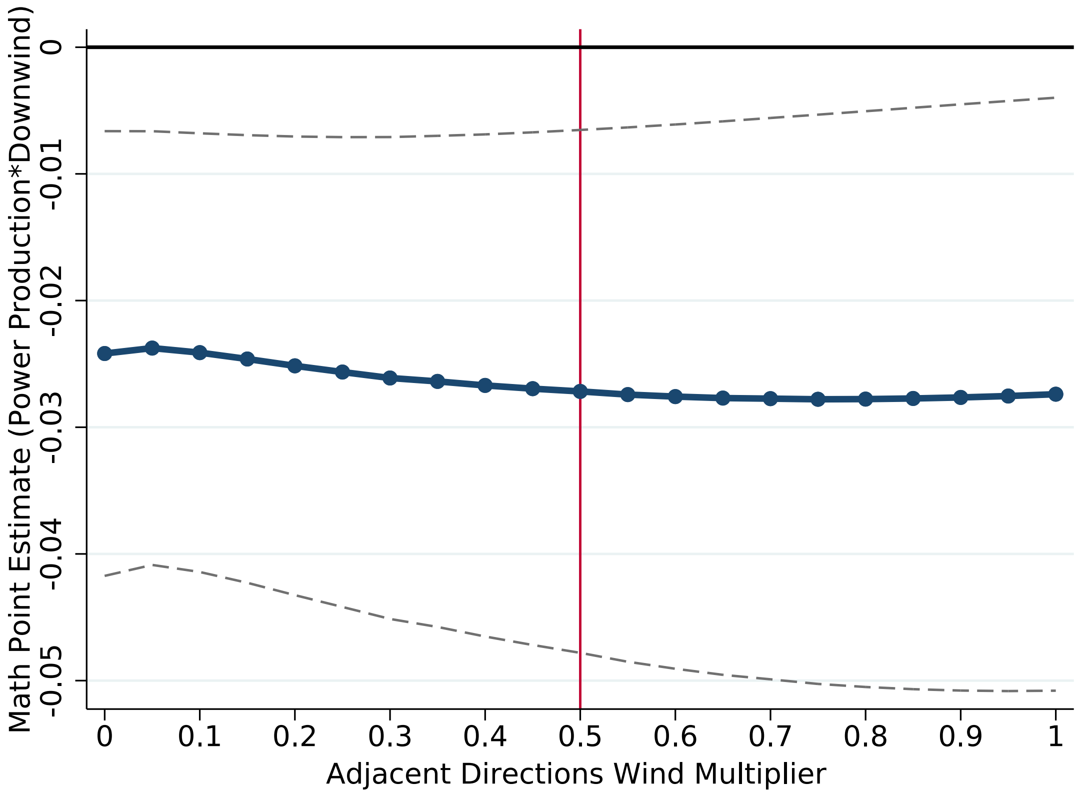
Notes: This figure captures the year-to-year production variation that we are using by displaying logged power production (during school months Sep-May) for the seven coal-fired power plants that are always open during our sample period (2000-01 through 2016-17). The figure is restricted to always-open coal plants for visual clarity, although year-to-year production variation among coal plants that closed is also used for quasi-experiments 1 and 2 (although the production variation caused by closure itself is not).

Figure A.4: Placebo Test



Notes: This figure reports the results of a placebo test where we run the second quasi-experimental design that captures the effects of power production and persistent wind patterns on mathematics scores (i.e., estimating equation (2)) but randomly assign the power production of a coal-fired plant to that of a different coal-fired plant. This is done to alleviate concerns that power plants may jointly reduce production due to a statewide shock (e.g., the Great Recession) that happens to affect students living downwind relative to upwind of power plants differentially. The sample is restricted to schools within $10km$ of a power plant and to years that the power plant was in operation (as a coal plant). Table A.1 reports the power plants and their period of operation. The placebo test is run 100 times, with the point estimates for each placebo placed in bins of 0.005σ. The x-axis denotes these 0.005σ bins while the y-axis indicates the number of placebo tests with a point estimate falling in that bin. Regressions include controls for student demographics and student, school, and grade-year fixed effects and our continuous wind measure is divided by 0.15 to make our estimates comparable to those reported in column (2) of Panel C in Table 4. The dashed vertical line represents the point estimate of -0.031σ found in column (2) of Panel C in Table 4.

FIGURE A.5: Robustness to the Wind Measure Used



Notes: This figure reports results from equation (2), which uses year-to-year variation in coal-fired power production along with across school variation in our continuous wind measure, when different weights are given to adjacent wind directions in the calculation of our downwind measure (see equation (6)). Specifically, the x-axis indicates the weight given to adjacent wind directions, with zero denoting no weight and one denoting equal weight. The continuous wind measure is divided by the difference between the downwind measure in 'downwind' schools relative to 'upwind' schools (split by above or below median according to the downwind measure) to keep point estimates consistent for the various 'downwind' definitions. The sample is restricted to schools within 10km of a power plant and to years that the power plant was in operation (as a coal plant). Table A.1 reports the power plants and their period of operation. Regressions include controls for student demographics and student, school, and grade-year fixed effects. The point estimate when adjacent wind directions are given half weight – denoted by the vertical line – effectively reports the same estimate as that reported in column (2) of panel C in Table 4, although it is slightly smaller as the difference between 'upwind' and 'downwind' schools is 0.13 rather than the 0.15-unit difference used to scale point estimates in Table 4.

Table A.1. Large Coal and Gas Power Plants in North Carolina

Plant Name (PlantID)	Fuel Type	Dates of Operation	(Lat, Lon)	Mean Production [Min-Max]	Km to Wind Station	Closure/Conversion Notes
Asheville Energy Plant (2706)	Coal	All years	(35.4731, -82.5417)	1,407,923 [883,586-1,824,702]	4.6	N/A
Cape Fear (2708)	Coal	2000-01 to 2010-11	(35.595, -79.0495)	1,271,130 [616,218-1,497,818]	4.9	Officially retired in October 2012, but retired 2 (of 4) coal-fired units in 2011 leading to a major decline in coal-fired energy production. Given 2011-12 production was less than one-quarter of that in 2010-11, 2011-12 is coded as first post-closure academic year.
H.F. Lee (2709)	Coal/Gas (Converted)	Coal: 2000-01 to 2011-12 Gas: 2012-13 to 2016-17	(35.3799, -78.0878)	Coal: 1,246,328 [702,694-1,566,722] Gas: 4,406,655 [2,736,551-5,100,208]	11.8	Retired all coal-fired units in September 2012 and so 2012-13 is coded as first post-closure academic year. Replaced by gas-fired combined cycle plant which started production in December 2012.
Roxboro (2712)	Coal	All years	(36.4833, -79.0731)	9,130,228 [3,520,377-11,632,098]	23.4	N/A
LV Sutton (2713)	Coal/Gas (Converted)	Coal: 2000-01 to 2012-13 Gas: 2013-14 to 2016-17	(34.2831, -77.9853)	Coal: 1,676,127 [754,695-2,368,378] Gas: 2,893,843 [2,006,652-3,311,531]	8.0	Coal-fired units were retired in November 2013 and so 2013-14 is coded as first post-closure academic year. Replaced by gas-fired combined-cycle plant that came online in late-2013.
W.H. Weather-spoon (2716)	Coal	2000-01 to 2010-11	(34.5875, -78.9755)	439,230 [218,852-648,041]	8.0	Plant retired in September 2011 and so 2011-12 is coded as first post-closure academic year.
GG Allen (2718)	Coal	All years	(35.1897, -81.0122)	2,946,652 [336,182-5,537,831]	6.4	N/A

Buck (2720)	Coal/Gas (Converted)	Coal: 2000-01 to 2010-11 Gas: 2011-12 to 2016-17	(35.7133, -80.3767)	Coal: 916,882 [383,752-1,418,080] Gas: 3,048,455 [1,717,465-3,711,169]	10.0	Officially retired April 2013, but 2 (of 4) coal-fired units were retired in mid-2011. Plant experienced a severe decline in production in 2009 (>75%) and its production levels never recovered. Given that, 2008-09 is coded as first post-closure academic year. Replaced by gas-fired combined-cycle plant that started production in 2011-12.
Cliffside (2721)	Coal	All years	(35.22, -81.7594)	2,546,192 [916,173-4,101,390]	14.9	N/A
Dan River (2723)	Coal/Gas (Converted)	Coal: 2000-01 to 2007-08 Gas: 2012-13 to 2016-17	(36.4862, -79.7208)	Coal: 460,957 [251,109-928,976] Gas: 2,873,443 [1,778,531-3,650,357]	12.9	Coal-fired plant officially retired on April 1, 2012, but major production effectively ceased after 2007-08 with production levels in 2008-09 being below 200,000Mwh, less than one-fifth 2007-08 production. 2008-09 is thus coded as first post-closure year. Natural gas plant opened in December 2012.
Marshall (2727)	Coal	All years	(35.5975, -80.9658)	8,429,976 [4,942,812-11,637,477]	18.6	N/A
Riverbend (2732)	Coal	2000-01 to 2010-11	(35.36, -80.9742)	1,074,186 [420,822-1,795,004]	15.2	Although officially retired on April 1, 2013, major power production ended on October 2012 with the retirement of four turbines in October 2012 with production in 2011-12 being below 40,000Mwh. Given this, 2011-12 is coded as the first post-closure year.
Mayo (6250)	Coal	All years	(36.5278, -78.8917)	2,599,171 [1,036,881-3,526,635]	28.2	N/A
Richmond/Smith Energy (7805)	Gas	2001-02 to 2016-17	(34.8392, -79.7406)	3,407,601 [287,921-8,753,174]	6.0	Began operation late-2001, and had a large expansion in late-2011.

Rowan (7826)	Gas	2003-04 to 2016-17	(35.7314, -80.6019)	1,245,976 [78,189-3,013,099]	12.0	Main energy production started in 2003-04 (although there had been a small amount of since 2001). Large expansion in 2008-09.
Belews Creek (8042)	Coal	All years	(36.2811, -80.0603)	9,715,447 [5,503,389-11,990,858]	21.9	N/A
Edgecombe Genco (10384)	Coal	2000-01 to 2010-11	(36.0373, -77.7533)	513,661 [339,506,693,125]	21.7	Continued producing some electricity until 2014-15, but production in 2011-12 fell by over two-thirds to below 140,000Mwh. 2011-12 is therefore coded as first post-closure year.
RJ Reynolds Tobaccoville (50221)	Coal	2000-01 to 2002-03	(36.2521, -80.3638)	201,242 [118,498-264,569]	18.3	Closed on March 1, 2004. 2003-04 is coded as first post-closure year as production for 2004-05 school year fell by over fifty percent (and production ceased in March of that year).
Roanoke Valley I (54035)	Coal	2000-01 to 2013-14	(36.4364, -77.6167)	1,122,406 [675,799-1,312,637]	11.9	Closed in 2017, but effectively ended major production in 2014, with 2014-15 production almost one-tenth of 2013-14 production. 2014-15 is therefore coded as first post-closure year.
Rockingham County CT (55116)	Gas	2011-12 to 2016-17	(36.3297, -79.8297)	385,307 [90,263-775,481]	12.1	Always open, but only generated significant energy production (>120k) after 2010-11.
Cleveland County Generating Facility (57029)	Gas	2013-14 to 2016-17	(35.1705, -81.4166)	458,470 [230,472-660,632]	19.3	Natural gas plant opened March 2013.

Table A.2. Summary Statistics: Upwind versus Downwind of Power Plant

	Within 10 km of Coal Plant		*Within 10 km of Gas Plant*	
	Downwind	Upwind	Downwind	Upwind
	(1)	(2)	(3)	(4)
Mean of Student Characteristics				
Mathematics Score (σ)	-0.05	0.03	-0.29	-0.27
Reading Score (σ)	-0.06	0.01	-0.30	-0.21
Lagged Math Score (σ)[2]	-0.03	0.04	-0.27	-0.23
Lagged Reading Score (σ)[2]	-0.03	0.02	-0.29	-0.19
% White	51.5	59.6	47.1	51.6
% Black	33.0	27.9	36.3	31.6
% Hispanic	9.1	5.6	9.5	11.1
% Asian	2.7	1.9	0.4	1.2
% Free or Reduced Price Lunch	45.2	48.2	65.8	62.2
% English Learners	3.8	2.7	2.8	3.5
% with Disability	7.2	6.9	11.5	13.3
% Gifted	11.5	15.3	11.8	15.5
% Repeating Grade	1.1	1.2	1.2	0.5
Distance to Power Plant (km)	6.9	6.7	6.7	5.8
Downwind Measure	0.31	0.17	0.31	0.18
# of Students	93,065	97,615	16,458	16,661
Observations (student-year)	226,363	242,732	36,216	36,119

Notes: This table breaks the summary statistics shown in columns (2) and (3) of Table 1 into 'upwind' and 'downwind' schools. 'Upwind' schools are defined as schools facing above median wind, while 'downwind' schools are those facing below median wind. Splitting the sample into 'downwind' and 'upwind' schools is done at the school level so that there are an equal number of schools in both columns (although the number of students does vary across columns). The sample is restricted to schools within 10km of a power plant and to years that the power plant was in operation (as a coal or gas plant). Table A.1 reports the power plants and their period of operation. Most covariates feature little missing data (<1% of sample), with the exception being lagged test scores which are missing for about ten percent of the sample.

Table A.3. Effect of **Natural Gas** Power Plant Production Using Production and Wind Variation

	Mathematics Scores (σ)			*English Scores (σ)*		
	(1)	(2)	(3)	(4)	(5)	(6)
A. 'Downwind' from Natural Gas-fired Plant						
Production	0.001	-0.005	-0.007	0.018	0.000	0.001
(1 million Mwh)	(0.014)	(0.020)	(0.022)	(0.019)	(0.017)	(0.016)
B. 'Upwind' from Natural Gas-fired Plant						
Production	0.048	0.035	0.027	0.030	0.012	0.010
(1 million Mwh)	(0.061)	(0.052)	(0.053)	(0.047)	(0.026)	(0.027)
C. Continuous Wind Measure						
Production	0.024	0.017	0.019	0.019	0.018	0.009
$*$downwind÷0.15	(0.033)	(0.032)	(0.032)	(0.019)	(0.018)	(0.017)
Distance Controls	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes
Observations	72,335	72,335	68,479	72,266	72,266	68,305
# of Students	178,065	178,065	29,694	31,584	31,584	29,654

Notes: Panels A and B contrast the effect of increased natural gas-fired power production for schools that are 'downwind' relative to 'upwind' of the natural gas-fired power plant according to our downwind measure (see equation (6)). Specifically, Panel A estimates equation (1) for 'downwind' schools facing above median wind, while Panel B does so for 'upwind' schools with below median wind levels. Splitting the sample into 'downwind' and 'upwind' schools is done at the school level so that there are an equal number of schools in both panels. Panel C reports results from equation (2) which uses year-to-year variation in coal-fired power production along with across school variation in our continuous wind measure. We divide our continuous wind measure by 0.15 to make our estimates comparable to the contrast between Panel A and B as, on average, our wind measure is about 0.15 units higher for 'downwind' schools relative to 'upwind' schools. Effect sizes in Panels A and B report the change in standardized test scores for a one Mwh increase in power production. Given the additional differencing layer, Panel C reports the change in standardized test scores for a one Mwh increase in power production *combined* with a 0.15-unit increase in our wind measure. The sample is restricted to schools within 10km of a power plant and to years that the power plant was in operation (as a coal plant). Table A.1 reports the power plants and their period of operation. Each cell represents the result of a separate regression of equation and includes student, school, and grade-year fixed effects. 'Demographic controls' include student level demographics and mean demographics at the school-grade level and include: ethnicity, gender, socioeconomically disadvantaged status, disability status, gifted status, English learner status, and grade repeating status. Missing indicators are used for students with missing demographics. 'Lagged test scores' consist of a cubic polynomial in prior lagged mathematics and reading scores interacted with grade dummies along with average lagged school-grade test scores. Students with missing lagged own-subject scores are dropped while a missing indicator is used for those missing a lagged other subject score. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Table A.4. Effect of **Coal** Power Plant Production Using **Log** Production and Wind Variation

	Mathematics Scores (σ)			*English Scores (σ)*		
	(1)	(2)	(3)	(4)	(5)	(6)
A. 'Downwind' from Coal-fired Plant						
log(production)	-0.081***	-0.080***	-0.079***	-0.052***	-0.052***	-0.052***
	(0.023)	(0.023)	(0.022)	(0.013)	(0.014)	(0.014)
B. 'Upwind' from Coal-fired Plant						
log(production)	-0.002	-0.002	0.002	0.005	0.004	0.006
	(0.024)	(0.024)	(0.025)	(0.011)	(0.011)	(0.012)
C. Continuous Wind Measure						
log(production)	-0.056***	-0.057***	-0.055***	-0.040***	-0.043***	-0.043***
$*$downwind÷0.15	(0.020)	(0.020)	(0.019)	(0.012)	(0.012)	(0.011)
Demographics Controls	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes
Observations	469,095	469,095	434,602	466,947	466,947	433,020
# of Students	178,065	178,065	162,923	177,067	177,067	161,942

Notes: This table recreates Table 4 using log(production) rather than production in millions of Mwh. Panels A and B contrast the effect of increased coal-fired power production for schools that are 'downwind' relative to 'upwind' of the coal-fired power plant according to our downwind measure (see equation (6)). Specifically, Panel A estimates equation (1) for 'downwind' schools facing above median wind, while Panel B does so for 'upwind' schools with below median wind levels. Splitting the sample into 'downwind' and 'upwind' schools is done at the school level so that there are an equal number of schools in both panels. Panel C reports results from equation (2) which uses year-to-year variation in coal-fired power production along with across school variation in our continuous wind measure. We divide our continuous wind measure by 0.15 to make our estimates comparable to the contrast between Panel A and B as, on average, our wind measure is about 0.15 units higher for 'downwind' schools relative to 'upwind' schools. Effect sizes in Panels A and B report the change in standardized test scores for a one hundred percent increase in power production. Given the additional differencing layer, Panel C reports the change in standardized test scores for a one hundred percent increase in power production *combined* with a 0.15-unit increase in our wind measure. The sample is restricted to schools within $10km$ of a power plant and to years that the power plant was in operation (as a coal plant). Table A.1 reports the power plants and their period of operation. Each cell represents the result of a separate regression and includes student, school, and grade-year fixed effects. 'Demographic controls' include student level demographics and mean demographics at the school-grade level and include: ethnicity, gender, socioeconomically disadvantaged status, disability status, gifted status, English learner status, and grade repeating status. Missing indicators are used for students with missing demographics. 'Lagged test scores' consist of a cubic polynomial in prior lagged mathematics and reading scores interacted with grade dummies along with average lagged school-grade test scores. Students with missing lagged own-subject scores are dropped while a missing indicator is used for those missing a lagged other subject score. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

B Nationwide Analysis

Our nationwide analysis supplements our North Carolina study by extending our design to utilize nationwide performance data on school proficiency rates. Here, we detail the data construction for our nationwide sample.

B.1 Data

School Performance Data: For performance data with nationwide coverage we draw on school level proficiency data reported by the U.S. Department of Education.[37] These data report mathematics and English proficiency rates on the state assessment by school from 2009-10 through 2017-18 and are (to the authors' knowledge) the only school level performance data with national coverage.[38] Proficiency rates vary substantially across states with mathematics proficiency rates ranging from thirty percent in New Mexico to almost eighty percent in Iowa. Given that each state uses a different test, these differential proficiency rates for the most part capture differences in the difficulty of each state's standardized test rather than differences in student performance across states. Furthermore, state proficiency rates vary substantially within states across time due to changes in statewide testing regimes. Notably, many states adopted Common Core standards during this time period, often causing large declines in statewide proficiency rates. While school fixed effects (which subsume state fixed effects) account for across state differences in test difficulty, we deal with within-state changes in testing regimes by constructing school proficiency rankings within each state-year cell. State proficiency rankings have the benefit of being relatively consistent under different testing regimes, allowing us to utilize within state-year variation even for states with one power plant (which is not the case if we include state-year fixed effects).

Data report school level proficiency rates, although for privacy reasons do not

[37] Available at https://www2.ed.gov/about/inits/ed/edfacts/data-files/index.html.

[38] While the Stanford Education Data Archive contain school level performance data, these data are not disaggregated by year which is required to use the time variation inherent in our empirical strategy.

report proficiency rates for schools with less than five students and 'blur' the proficiency rates for schools with up to three-hundred students by reporting achievement ranges (e.g., 5-10 percent proficient). When achievement ranges are reported, the midpoint of the achievement range is assigned as the school level proficiency rate (e.g., 5-10 percent proficient is assigned as 7.5 percent proficient). The data also separates out the number of test takers by grade, with high school state standardized tests reported as a separate category. High school testing differs substantially from testing in lower grades due to the fact that students can select different high school exams (e.g., 'Foundations of Mathematics' rather than 'Algebra II') whereas all students generally must take the same exam in third through eighth grade. We therefore create a separate statewide proficiency ranking for high schools and omit schools that serve both elementary and high school students (e.g., K-12 schools). Our results (available upon request) are similar if high schools are dropped or K-12 schools are included.[39]

School Demographic Data: School demographic data from 2009-10 through 2017-18 is collected from the National Center for Education Statistics (NCES).[40] Data is collected on: enrollment, ethnicity, gender, socioeconomically disadvantaged status, and the ratio of students to full time equivalent teachers. These data allow us to construct the following school year level controls: enrollment, percent of students that are male, percent belonging to a given ethnicity group (e.g., African-American, Asian, Hispanic, and White), percent socioeconomically disadvantaged, and the student-teacher ratio. We also use the number of test takers from the performance data to calculate school level test-taking rates. This data is then merged to the school performance data using the NCES school id, with 98.3 percent of schools with performance data being matched to the student demographics data.

Power Plant Data: As in our North Carolina analysis, data on national power plant locations and production are drawn from the EIA Monthly Generation and

[39]Three-quarters of the schools in our data are elementary and middle schools that have no high school tests conducted. Fifteen percent of our schools are high schools and the remaining ten percent are schools that have both middle school and high school grades (e.g., K-12 schools).

[40]Available at https://nces.ed.gov/ccd/.

Fuel Consumption Time Series File (EIA-923) for 2009 through 2018 (as performance data are only available 2009 onwards). We mimic the restrictions we used for North Carolina whereby we only include plants with average power production levels using fossil fuels of over 250,000 Mwh during the academic year over their period of operation. Once again, we match the monthly power production to the academic year by summing power production generated from September to May for each plant by fuel type.

The national data requires some additional data cleaning relative to our North Carolina analysis. First, unique plants in the EIA data are sometimes on the same industrial site as another plant. We combine these plants together by manually investigating all power plants within $3km$ of each other and assigning them an identical plant identifier if there are no residential areas between them. Second, the national data includes power plants with significant power production from oil or biomass fuels, which we code as a third plant type in addition to our coal and natural gas plants. Third, several power plants produce electricity using both coal and natural gas. In our main results, we assign the plant type using the predominant fossil fuel, although we also drop these combined fuel source plants in column (4) of Table B.6. Fourth, power plants can sometimes be in close proximity to each other making it difficult to attribute the effect of increased power production to a single plant. We drop these cases in column (5) of Table B.6. Schools are then matched to the nearest power plant according to physical distance,[41] which creates our main analysis sample of 88,626 schools matched to 922 power plants.

Summary Statistics: Table B.5 reports summary statistics from our nationwide sample. Column (1) shows the summary statistics for schools across the nation, with column (2) restricting the data to schools within $10km$ of a coal power plant. Compared to the nation at large, schools near coal power plants perform worse and have students that come from a more disadvantaged background. Column (3)

[41]We calculate the distance to five closest power plants so that we can make the appropriate sample restrictions to exclude schools that are near two power plants that have substantial fossil fuel power production in the same school year.

reports the data for schools within 10km of a natural gas plant; these schools are somewhat higher performing and feature substantially more Hispanic students relative to the coal plant sample. Differences between the coal and gas plant samples are likely driven by the geographic location of plants: coal plants dominate in the Midwest, for instance, where there are relatively few Hispanic students.

B.2 Calculating Effects for Each State

To calculate the impact of reduced coal usage on student performance for each state we multiply the estimated impact of one million Mwh of coal-fired power production on student performance (-0.02σ) by the enrollment weighted change in exposure to coal-fired power generation within 10km of each school in the state from 2006-07 to 2017-18. We construct the change in coal-fired production exposure by once again drawing on national power plant production from the EIA Monthly Generation and Fuel Consumption Time Series File (EIA-923 for the 2017-18 data, EIA-906/920 for the 2006-2007 data). This time, we do not restrict attention only to plants with production above $250,000$ Mwh during the school year. Instead, we consider all coal producing plants in the United States, summing September to May power production by plant to obtain overall academic year production. Plant location data is then drawn from the EIA-860 generator information series. We match each plant operating in the 2017-2018 academic year to the 2018 location data, and match each plant operating in the 2006-07 academic year to the 2012 location data, which is the earliest available. For plants operating in 2006-07 which are not included in the 2012 location data, we obtain their locations from the United States Environmental Protection Agency (EPA) 2007 Emissions and Generation Resource Integrated Database (eGRID).[42] The final sample contains 665 plants in 2006-07 and 390 plants in 2017-18. There are 363 plants which are operational in both time periods.

We then combine our national plant-level coal production data with school-

[42] Available at https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid.

level data from the National Center for Education Statistics (NCES) Elementary and Secondary Information System (ELSI)[43] for the 2006-07 and 2017-18 academic years, which itself comes from the Department of Education's Common Core of Data survey. For each public school in the United States, our data contains a school name, state, latitude and longitude, as well as total enrollments broken down by race and socioeconomically disadvantaged status. The final sample contains $94,164$ schools in 2006-07 and $95,256$ schools in 2017-18. There are $80,328$ schools which had positive enrolments in both 2006-07 and 2017-18. Each school is matched to all coal plants within a 10 kilometer radius separately in each academic year. This allows us to sum the total coal production of all plants within these radii, identifying the total Mwh generated in the vicinity of each school, which we code in increments of one million Mwh. We then calculate state averages of coal production within the vicinity of all schools', weighted by total or subgroup-specific enrollment. Subtracting the 2017-18 values from those obtained in 2006-07 gives us the enrollment weighted average change in exposure to coal-fired power generation within $10km$ of all schools in the state from 2006-07 to 2017-18.

[43]Available at https://nces.ed.gov/ccd/elsi/tableGenerator.aspx.

Table B.5. Summary Statistics: **National** Data

	All Schools[1] (1)	Within 10 km of Coal Plant[2] (2)	Within 10 km of Gas Plant[3] (3)
Mean of School Characteristics			
State Proficiency Rank (Math)	48.5	39.2	42.4
State Proficiency Rank (English)	48.3	38.7	41.7
Enrollment	587.2	569.4	660.6
% White	53.7	47.5	31.8
% Black	15.8	24.8	21.5
% Hispanic	22.0	20.9	36.4
% Asian	4.1	2.8	6.5
% Free/Reduced Price Lunch	52.6	59.6	61.4
Student-Teacher Ratio	17.0	16.5	17.5
Distance to Power Plant (km)	33.2	6.3	6.1
# of Plants[4]	922	351	523
# of Schools	88,626	5,607	14,582
Observations (school year)	665,511	42,824	106,911

[1] Data coverage: 2009-10 through 2017-18. Includes all schools, regardless of distance to nearby power plant.
[2] Includes all schools within 10km of a coal-fired power plant that produced more than 250,000 Mwh of electricity from coal during at least one school year from 2009-10 through 2017-18.
[3] Includes all schools within 10km of a natural gas fired power plant that produced more than 250,000 Mwh of electricity from natural gas during at least one school year from 2009-10 through 2017-18.
[4] Includes any power plant that produced more than 250,000 Mwh of electricity from coal, natural gas, oil, or biomass during at least one school year from 2009-10 through 2017-18.

Table B.6. Effect of Power Production on Mathematics Proficiency Rank Using Production and Wind Variation by Plant Type: **National** Data

	All Plants (1)	All Plants (2)	All Plants (3)	Primary Fuel Only Plants (4)	No Other Nearby Plants (5)
A. Coal-fired Plants					
Production	-0.39**	-0.32**	-0.39***	-0.46***	-0.32**
(1 million Mwh)	(0.16)	(0.16)	(0.15)	(0.17)	(0.16)
Observations	42,824	39,507	33,937	28,211	29,752
# of Schools	5,607	5,525	5,319	4,866	4,873
B. Natural Gas Plants					
Production	-0.05	-0.11	-0.12	-0.11	-0.08
(1 million Mwh)	(0.09)	(0.09)	(0.008)	(0.09)	(0.09)
Observations	106,911	96,209	82,218	75,131	62,679
# of Students	14,582	14,223	13,602	13,197	10,914
Demographics Controls	No	Yes	Yes	Yes	Yes
Lag State Prof. Rank	No	No	Yes	Yes	Yes

Notes: This table reports estimates from equation (7) which identifies the effect of a one Mwh increase in power production on student performance using the nationwide data (see Appendix B.1). Panel A reports estimates for schools nearby a coal-fired power plant, while Panel B does so for schools neighboring a natural gas power plant. The sample is restricted to schools within 10km of a power plant the produced at least 250,000 Mwh over school months for one school year using that fossil fuel (coal for Panel A and natural gas for Panel B). Column (4) restricts the sample to power plants where no more than 250,000 Mwh of power is produced by an alternative fossil fuel, while column (5) restricts the sample to schools with only one power plant within 10km of them. The outcome variable is defined as the school's ranking based on their school wide proficiency rate in their state. Each cell represents the result of a separate regression and includes school and year fixed effects. 'Demographic controls' include school level controls for: ethnicity, gender, socioeconomically disadvantaged status, class size, and enrollment. 'Lag State Prof. Rank' consist of a cubic polynomial in prior lagged statewide proficiency rank in mathematics and reading scores interacted with state dummies. Standard errors are clustered at school level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.