Three Essays on Energy Economics

by

Louis Demetri Preonas

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Meredith Fowlie, Co-chair
Professor Severin Borenstein, Co-chair
Professor Maximilian Auffhammer
Assistant Professor W. Reed Walker

Spring 2018

Three Essays on Energy Economics

Abstract

Three Essays on Energy Economics

by

Louis Demetri Preonas

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Associate Professor Meredith Fowlie, Co-chair

Professor Severin Borenstein, Co-chair

Electricity powers the modern economy, and the electricity supply chain is notoriously complex. Power plants must develop stable relationships for fuel procurement, as their long-run profitability hinges on securing a cheap, reliable fuel supply. Electric utilities may also lack the incentive to provide a reliable power supply to all potential customers, which could hamper economic productivity. The physical properties of electricity transmission create inherent challenges in providing power to all regions of the grid, while simultaneously incentivizing economically efficient production decisions. In this dissertation, I study three potential market failures in electricity supply: (i) market power in U.S. coal transportation; (ii) under-electrification of India's rural poor; and (iii) short-run allocative inefficiencies in Indian electricity dispatch. In each case, my findings are of substantial economic importance due the scale of the electric power industry, which is essential to virtually all economic activity. Climate change only raises the stakes, and alleviating electricity market failures has the potential to increase carbon dioxide emissions and further harm the planet.

In the first chapter, I investigate how market power in the transportation of coal might impact U.S. climate policies. Economists have widely endorsed pricing CO_2 emissions to internalize climate change-related externalities. Doing so would significantly affect coal, which is the most carbon-intensive major energy source. However, U.S. coal markets exhibit an additional distortion, as the railroads that transport coal to power plants can exert market power. This upstream distortion can mute the price signal of a corrective tax, due to changes in markups or incomplete tax pass-through. I provide the first empirical estimates of how coal-by-rail markups respond to changes in coal demand. I find that rail carriers reduce coal markups when downstream power plant demand changes, due to a decrease in the price of natural gas (a competing fuel). I estimate markup changes that vary substantially across coal plants, resulting from a combination of heterogeneous transportation market structure and plant-specific demand shocks. Since low natural gas prices and a CO_2 emissions tax similarly disadvantage coal, observed decreases in coal markups imply that pass-through of a federal carbon tax to coal power plants may be

heterogeneous and incomplete. This could substantially erode the environmental benefits of a price-based climate policy. My results suggest that decreases in coal markups have increased recent climate damages by $2.3 billion, compared to a counterfactual where markups do not change.

In the second chapter, coauthored with Fiona Burlig, we study the impacts of energy access in the developing world. Over 1 billion people still lack electricity access. Developing countries are investing billions of dollars in rural electrification, targeting economic growth and poverty reduction, despite limited empirical evidence. We estimate the effects of rural electrification on economic development in the context of India's national electrification program, which reached over 400,000 villages. We use a regression discontinuity design and high-resolution geospatial data to identify medium-run economic impacts of electrification. We find a substantial increase in electricity use, but reject effects larger than 0.26 standard deviations across numerous measures of economic development, suggesting that rural electrification may be less beneficial than previously thought.

In the third chapter, coauthored with Fiona Burlig and Akshaya Jha, we examine short-run allocative inefficiencies in Indian electricity supply. Electricity consumption is highly correlated with economic development. Understanding and resolving the drivers of economic inefficiencies in electricity markets is critical to supporting economic growth. We quantify the costs of short-run misallocation in Indian electricity supply. We assemble a novel dataset on daily production from each utility-scale power plant in the country and administrative measures of plant-specific marginal operating costs, and calculate the total variable costs of electricity generation in India to be approximately $29 billion per year. We next construct the "least-cost" counterfactual where we dispatch power plants in order of lowest-to-highest marginal cost. We find that this least-cost dispatch results in total annual operating costs that are roughly $4.7 billion lower than observed dispatch. Once we account for transmission constraints, we find a remaining misallocation wedge of $3.2 billion per year. We find evidence that this wedge results from market design and political economy considerations, but find little evidence of market power.

Contents

List of Figures

List of Tables

Acknowledgements

I owe an immense debt of gratitude to my dissertation advisors, Meredith Fowlie and Severin Borenstein. Meredith convinced me to come to Berkeley sight unseen, and I was *her* student from day one. She was always my strongest, most enthusiastic advocate. When she sensed I was struggling, she proactively initiated weekly one-on-one meetings devoted to keeping me on track. These meetings were *the* highlight of graduate school, and as we navigated the weeds of my research (amid countless detours and wrong turns), I grew by leaps and bounds—simply by watching Meredith's creative problem solving as she fired on all (fuel-efficient) cylinders. Above all, Meredith taught me to research both forward *and* backward, making choices that balance the information I currently have against a desired-yet-feasible endpoint.

I became Severin's student after a few years in Severin's orbit, where the gravitational pull is unusually strong. This reflects both his uncanny ability to intuit deep economic insights, and his genuine, holistic dedication to seeing his colleagues succeed. He fielded my distressed phone calls between job interviews, and freely dispensed mental health therapy during my recurring bouts of imposter syndrome. Without fail, I would feel wiser, more confident, and more put-together upon leaving Severin's office than I had when I entered. Through Severin, I learned the value of selecting research projects that are exciting to me *and* important to society. I am profoundly humbled to be joining his long, distinguished list of advisees.

In addition to my two primary advisors, I enjoyed the rare privilege of having five unofficial sub-advisors within the Energy Institute. Max Auffhammer expertly shepherded me through seemingly every phase of my PhD, from introductory econometrics, to second-year paper, to orals, to the academic job market. I could summarize six years of academic progress using six pivotal conservations in Max's office, each of which kinked my trajectory upwards. I will never forget which regressions friends don't let friends run. Reed Walker taught me how to properly conceptualize a research project, and consistently identified weaknesses in my research that others had overlooked. He encouraged me to aim high, and his constructive feedback (as much as anyone's) propelled me toward success on the job market. Catherine Wolfram was always exceedingly generous with her time and attention, dating back to our very first conversation—when she called me in Costa Rica to give a persuasive 45-minute sales pitch for why I should come to Berkeley ARE, despite being a professor in a different department! Her razor-sharp insights helped me learn to navigate the awkward transition from initial results to full-fledged paper. Lucas Davis taught me to clarify, distill, and express my ideas. His critique of one of my presentations—"There were just too many ideas!"—yielded a higher marginal benefit on my future career than any single comment I received in grad school. Jim Sallee showed me what deep, big-picture economic thinking looked like. I aspire to one day think half as big.

The Energy Institute has been my home base for the past five years, and I could not have asked for a more supportive group of coworkers. Casey Hennig initially brought me into the fold, and always made sure my computer had enough RAM. Cristina Bentley provided invaluable technical, logistical, and moral support as I careened toward the PhD finish line. Paula Pedro always offered sound advice on research, music, and life. Ela é uma ótima colega, e uma amiga ainda melhor. Andy Campbell helped to maintain an open exchange of ideas between faculty, students, and EI staff, and he even trusted me to man the `@energyathass` Twitter account. The incomparable, indispensable Karen Notsund helped me navigate the ins and outs of Berkeley student fees, external grant applications, and mystifying data protocols. She gave me pep talks when I needed them, made sure I always received a paycheck, brought me alfajores after each visit to Chile, and bought me a pie on each of my last five birthdays.

I am also incredibly grateful for the exceptional training, mentoring, and support I received from the Department of Agricultural and Resource Economics. ARE faculty were continuously invested in my success and genuinely cared about my progress, even if I wasn't their student. Leo Simon showed me why an affine function is not linear, and taught me to visualize Cramer's Rule despite my inability to comprehend 3D graphics. Betty Sadoulet taught me how to skin, gut, and butcher a regression table, as only she could. Jeremy Magruder repeatedly managed to see the forest for the trees, and helped me put the pieces together on each of my research projects. Sofia Villas-Boas bent over backwards to get me a job, and helped to foster the warm, supportive research culture that allowed me to thrive. Peter Berck regaled me with valuable advice on research, the job hunt, and life in the economics profession. Carmen Karahalios wore many hats with aplomb—she was my ultimate confidant, fixer, and cheerleader. Diana Lazo put up with my assorted reimbursement-related shenanigans, and helped me submit 151 job applications. A Carmen y Diana les debo montones de agradecimiento.

I would be remiss without thanking my brilliant, supportive classmates. For the past six years, they consistently challenged me to hone my skills, sharpen my ideas, and exceed my own expectations. Fiona Burlig was the ideal research partner in every way. She was always ready to brainstorm new project ideas, to fight me on LaTeX formatting minutiae, and to call me out when I was being ridiculous. I look forward to many more years of friendship, street food, coauthorship, and mind-meld. Josh Blonz graciously paved the way for my entire grad school experience, from RFF to Berkeley ARE, to the Energy Institute, to the job market. He helped me evaluate countless identification strategies, while coaching me on exciting new ways to optimize my life. Kenny Bell urged me fight my verbose tendencies and get to the punchline, while also inspiring me to hone my chicken-frying skills. I still cannot believe we saw Paul McCartney at Candlestick *and* Radiohead at the Greek. Andrew Stevens taught me the benefits of engaging with the broader campus community. He was always the responsible adult in the room, despite being younger when he finished his PhD than I was when I started mine! Matt Woerman showed me the value of a meticulous whiteboard derivation, and helped me stay grounded during some mentally/emotionally trying stretches. Ceren Baysan drove me on hikes, and fed me snacks to keep my awake through countless afternoon lectures and seminars. Jimmy Gillan helped me learn to stay honest and stick to my guns, especially when I felt the urge to short-cut a problem set. Becca Taylor encouraged me to run longer distances, and was always willing to sit down and hash out research strategies. Daniel Tregeagle indulged my longwinded rants, and enlightened me on the history of economic thought. Patrick Baylis, Susanna Berkouwer, Judd Boomhower, Aluma Dembo, Karl Dunkle Werner, Walter Graf, Sylvan Herskowitz, Akshaya Jha, Erin Kelley, John Loeser, Eva Lyubich, Dave McLaughlin, Erica Myers, Ed Rubin, and Derek Wolfson individually contributed to my success to a greater degree than each of them realizes.

Writing this dissertation was a never-ending journey, and I appreciate every little bit of help I received along the way. Jason Trout exorcised and reanimated my computer at least four times. Liz Vadas shipped me a Zingerman's build-your-own Reuben kit on dry ice (the Rolls Royce of edible mail-order gifts), with enough protein and grease to sustain me through my most frantic dissertation writing push. The music of Caetano Veloso shocked me out of a seemingly insurmountable writer's block, while also helping to restore my depleted mental and emotional health. My inordinately accommodating roommates Daniel Tregeagle, Andrea Wooding, Harrison Mantas, Emma Connolly, Becca Taylor, and Cristine Tenant *never* complained when I converted our kitchen table into my personal research command center, often for weeks at a time. My chiropractor Korosh Shakeri was instrumental in helping me to overcome injuries from a cycling accident—including broken ribs, whiplash, migraines, and light sensitivity. I cannot fathom my recovery without his care.

Completing a PhD is expensive, and I benefited from an embarrassment of riches. I am grateful to Energy Institute at Haas for employing me for four years, and for providing extensive financial and non-financial research support. I am also grateful to the Department of Agricultural and Resource Economics, the Haas School of Business, and the National Bureau of Economic Research, each of which covered a portion of my salary, tuition, and benefits. Grant funding from the College of Natural Resources, the Development Impact Lab, the Berkeley Institute for Transparency in the Social Sciences, and the Giannini Foundation afforded me the time and resources to get research projects off the ground. Resources for the Future's Joseph L. Fisher Dissertation Fellowship allowed me to complete my dissertation comfortably, while my life was consumed by the academic job market. Throughout my time as a graduate student, Chris McAlpine deftly managed my personal finances so that I didn't have to.

I owe much of my success to my incredibly supportive community of Bay Area family and friends. Pam and Dave Fullerton welcomed me into their home, and into their family life. They cooked me countless Sunday dinners, and even took the opportunity to introduce me to a few famous economists. If every PhD student had their own Thea Fighter nearby, PhD completion rates would skyrocket. John Fullerton invited me to chase after dogs in Montclair, and to chase him down off-trail moguls in Little Cottonwood Canyon.

Anastasia Fullerton took me boating in Lake Tahoe (not one of my best days), and was always the life of the party. George Fullerton helped me access a few vital datasets, and he even gave me a reason to attend my only Cal sporting event (though, admittedly, I felt obliged to root for his Stanford squad). India Fullerton was the epitome of a good dog. Stephen Preonas introduced me to the Bay Area nightlife, and nights out with Stephen were as unpredictable as they were memorable. Bryan Klofas and Michaela Parnell took me on backpacking trips where I learned to track satellites in the night sky. Meredith Stott drove me on many hikes, and one outrageously fun kayaking trip. Dave McLaughlin and Sarah Jo Szambelan fed me ribs, hosted numerous game nights, and always made me feel at home. Having my dear friend Adam Stern nearby, even for such a short time, was the ultimate treat. And what can I possibly say about Leslie Briggs? For 12 years running, she's been the best friend a guy could ask for. My Bay Area experience was in large part the Leslie Briggs experience, and I mean that in the best possible way.

Outside of California, several of my long-distance friendships helped to sustain me through six years of grad school—most notably, Richard Xu, Megan Jacobs Swain, Matt Swain, Jesse Zarley, Katie Schultz, and Luke Bostian. I almost certainly would not be where I am today without four of my mentors. Connie Kelly, my eighth-grade math teacher, opened my eyes to the intrinsic joys of solving math problems. Jan Gerson, my undergraduate advisor, convinced me to pursue a degree in economics. Steve Salant, my undergraduate mentor, first got me excited about pursuing a career in economic research. Carolyn Fischer, my first research supervisor, threw me into the deep end and made sure I figured out how to swim. Both before and during grad school, my parents and brother always gave me their unconditional love and support. Every edible care package meant the world to me, and every late night phone call helped set me straight and move me one step closer to the finish line. (I am quite excited to be done with this three-hour time difference.) My extended family of aunts, uncles, and cousins also made for one hell of a cheering section. I lost two grandparents during my time as a graduate student. I still miss them terribly. They always encouraged me to work hard and play hard, and I know they would thrilled to see all of my hard work finally pay off—especially because I had plenty of fun along the way.

A PhD demands a tremendous amount of resourcefulness, creativity, and self-discipline. I dedicate my dissertation to the memory of three women who helped to instill these qualities in me. Evelyn Wheeler raised me to always save my leftovers, down to the last ladle full of soup. She made sure I would know how to open a watermelon at a picnic if no one remembered to bring a knife. Her indelible catchphrase—"Wellllll. Ah *jest don't know!*"—still echoes every time I face a problem that demands a creative solution. Molly Macauley exemplified a life devoted to policy-driven research, and she (unknowingly) served as one of my earliest academic role models. Above all, Molly taught me to listen before speaking. I still marvel at her ability to facilitate a meeting with the utmost efficiency, effectiveness, and grace. Finally, Louise Angermeier, my piano teacher throughout middle school and high school, imparted countless lessons that apply as well to economic research as they do to music. Louise valued precision, technique, and mechanics. She taught me to digest Rachmaninoff one measure at a time, while adhering to a painfully slow metronome. I use this same process each time I debug a piece of code, run a new regression, prepare for a presentation, or confront a challenging mathematical proof. Thanks to Louise, piano playing remains my one truly irreplaceable creative outlet and mental health salve. I could not have survived grad school without it.

Chapter 1

Market Power in Coal Shipping and Implications for U.S. Climate Policy[1]

1.1 Introduction

Economists have widely advocated policies that price carbon dioxide emissions to reflect their marginal external cost (Nordhaus (1993)). While such policies are efficient under perfect competition (Pigou (1932)), additional distortions such as market power reduce the efficiency of a Pigouvian tax (Buchanan (1969); Barnett (1980)). Economists have long understood that firms with market power may adjust prices in response to taxation (Cournot (1838)). However, there exists surprisingly little empirical research on how market power impacts the pass-through of an environmental tax, or the transmission of the desired price signal to market participants.

This paper investigates market power in the transportation of coal, and analyzes its potential impacts on the efficacy and efficiency of U.S. climate policy. Coal is likely the most environmentally damaging and carbon-intensive industry in the U.S. economy (Muller, Mendelsohn, and Nordhaus (2011)). Many geographically concentrated mines supply coal to many geographically dispersed power plants, and the railroads that transport coal from mines to plants can exercise market power (Busse and Keohane (2007)). If a carbon tax causes these oligopolist railroads to reduce coal markups, this could mute the carbon price signal received by power plants and erode the environmental benefits of the tax. While previous research has studied environmental and economic outcomes

[1]I thank Meredith Fowlie, Severin Borenstein, Maximilian Auffhammer, Lint Barrage, Cyndi Berck, Peter Berck, Susanna Berkouwer, Joshua Blonz, Fiona Burlig, Lucas Davis, Karl Dunkle Werner, Don Fullerton, James Gillan, Stephen Jarvis, Akshaya Jha, Jeremy Magruder, Dave McLaughlin, Elisabeth Sadoulet, James Sallee, Leo Simon, James Stock, Sofia Villas-Boas, Reed Walker, Matt Woerman, Catherine Wolfram, and Derek Wolfson for their invaluable feedback on this dissertation chapter. I also thank seminar participants at the Energy Institute at Haas, UC Berkeley, the Heartland ERE Workshop at University of Illinois Urbana-Champaign, University of Nevada Reno, University of Maryland, Cornell University, UC Santa Barbara, University fo Utah, Georgia State University, and University of Pennsylvania. All remaining errors are mine.

under a carbon tax, I provide the first estimates of how upstream market power in coal supply might impact climate policy outcomes.

I begin by estimating the size of the market power distortion in coal transportation, or the average markup levels faced by coal power plants. Then, I estimate how markups change due to changes in the demand for coal. Theory suggests that a shift in coal demand should cause a profit-maximizing railroad to reoptimize coal markups. Recent decreases in the price of natural gas, coal's primary competitor in electricity markets, represent a negative shock to power plants' coal demand. Since a carbon tax would induce a similar shift in coal demand, observed changes in coal markups due to decreases in the gas price may predict how railroads would reoptimize markups under a carbon tax.[2]

To identify markup levels, I exploit predetermined cross-sectional heterogeneity in market power. Some coal plants are "captive" to a single rail carrier and face an effective transportation monopoly; other plants may purchase coal from multiple railroads, and these "non-captive" plants face more competitive coal shipping. I implement a nearest-neighbor matching strategy to compare the price of coal delivered to captive vs. non-captive plants, which takes advantage of plants' inability to arbitrage spatial price differences. By flexibly controlling for coal commodity value and railroad freight costs, I recover the average differential markup faced by captive plants.

I use a difference-in-differences design to identify changes in coal markups caused by changes in the price of natural gas. This leverages two sources of cross-sectional heterogeneity: (i) geographic variation in transport market power, and (ii) variation in coal plants' sensitivity to gas price changes, which I predict from microdata on U.S. electricity generation. Using a simple oligopoly model as a guide, I combine these two sources of variation into a single cross-sectional predictor of markup changes, and interact this variable with the time series of gas prices. Regressing the delivered price of coal on this interaction in a panel fixed effects framework, I estimate the extent to which gas price changes cause differential changes in coal markups. Given that natural gas is the primary substitute for coal in electricity supply, negative shocks to the gas price disadvantage coal generation in a manner similar to a tax on CO_2 emissions (Cullen and Mansur (2017)). Hence, observed gas price shocks mimic the variation of a carbon tax, and my estimates of markup changes can help predict the pass-through of such a tax.

I find that coal plants facing the most market power in transportation pay \$2–5 per ton higher average markups, compared to plants facing the least market power. This translates to an average markup of 4–14 percent of delivered coal prices, explaining 13–41 percent of the average spatial gap between mines' sales prices and plants' delivered prices. I also find robust and statistically significant *changes* in markups for approximately 43 percent of plants—the subset of plants that face the most market power *and* are sensitive to competition from gas-fired generation. For these "markup-sensitive" plants, a \$1/MMBTU drop in gas price causes coal markups to fall by \$1 per ton. I find no evidence that markups change for the remaining 57 percent of coal plants, which are less sensitive to gas-fired competition *or* face relatively little market power.

[2] I use the terms "natural gas" and "gas" interchangeably. My analysis does not relate to gasoline.

I demonstrate that rail carriers reoptimize markups to effectively insulate some coal plants against shocks to their competitiveness. As decreasing gas prices reduce the marginal cost of gas-fired generation, "markup-sensitive" plants see their coal prices decrease, thereby reducing these plants' *own* marginal costs. Such offsetting changes in markups help this subset of coal plants to remain competitive with their gas-fired rivals. By contrast, over half of coal plants experience the full gas price shock, as their markups do not adjust. These heterogeneous impacts across plants are qualitatively consistent with the predictions of the static oligopoly model that I develop, implying that rail carriers indeed reoptimize markups heterogeneously to maximize profits in coal shipping.

Falling gas prices have given gas-fired plants a competitive advantage over coal-fired plants. This is similar to what might occur under a carbon tax, which would penalize coal as the more carbon-intensive fuel. Therefore, I can convert my estimated markup changes into the pass-through rates of an implicit carbon tax, or the rates at which rail carriers would have passed a mine-mouth carbon tax on to delivered coal prices.[3] For the subset of plants whose markups do not change, this translates to full pass-through, or implied pass-through rates statistically indistinguishable from 1. By contrast, this translates to incomplete pass-through for "markup-sensitive" plants, with plant-specific pass-through rates ranging from 0.98 to 0.42. This suggests that market imperfections in coal shipping are likely to distort the price signal of a federal carbon tax, such that certain coal plants may experience as little as 42 percent of the desired cost increase.

This paper contributes to four different literatures. First, my results contribute to the literature on market power in intermediaries. Atkin and Donaldson (2015) develop techniques to identify markups separately from transportation costs, and several recent papers estimate how oligopolistic intermediaries influence both upstream and downstream outcomes (e.g., Startz (2016); Ganapati (2017)). While these studies typically focus on differentiated product markets, coal is a globally traded commodity that is relatively homogeneous. I leverage a unique feature of coal markets—limited spatial arbitrage between power plants—to credibly identify transport markups while invoking relatively few structural assumptions on coal demand. My results have important implications for many commodities with high geographic specificity and high transportation costs, including crude oil, cement, and metals.

My analysis also contributes to the literature on coal intermediaries, which has largely focused on the railroads' interactions with upstream mines (e.g., Kolstad and Wolak (1983); Wolak and Kolstad (1988)), rather than downstream power plants. A notable exception is Busse and Keohane (2007), who provide the first evidence of price discrimination due to geographic variation in coal shipping during the 1990s. My results demonstrate that heterogeneous coal markups have persisted through recent years, and I am the

[3]The physical location of the tax along the coal supply chain should not change the economic interpretation of pass-through, in the absence of additional market distortions beyond rail market power (Weyl and Fabinger (2013)). "Forward" pass-through of a mine-mouth tax (i.e. a cost shock to rail shipping) follows the standard formulation of a cost shock passed through to final goods prices. However, in practice, carbon taxes are typically levied on electricity sales.

first to show that markup changes have led to economically meaningful impacts on CO_2 emissions.

Second, my results contribute to a growing empirical literature on environmental policy in the presence of market power. Given widespread evidence of market power in major polluting industries (e.g., Bushnell, Mansur, and Saravia (2008) on electricity markets; Hastings (2004) on gasoline markets), surprisingly few studies have empirically estimated the theoretically ambiguous interactions of these two market failures. Mansur (2007) finds that market power in electricity markets can increase pollution abatement under environmental regulation. On the other hand, Ryan (2012) and Fowlie, Reguant, and Ryan (2016) find that emissions regulation exacerbates market power distortions in the cement industry. I find that changes in coal markups may significantly erode the environmental benefits of a carbon tax, as incomplete pass-through would mute the price signal felt by a subset of coal plants. My results suggest that incomplete pass-through increased CO_2 emissions damages during my sample period by roughly \$2.3 billion, compared to a full pass-through counterfactual. Hence, the magnitude of this effect would likely be economically meaningful, despite the fact that incomplete pass-through would only impact a fraction of coal plants.

Incomplete pass-through of a carbon tax could increase or decrease welfare, depending on the size of the tax relative to marginal external costs. If the tax were equal to the social cost of carbon, then the presence of markups would restrict coal consumption past the socially optimal quantity. In this case, incomplete pass-through would reduce the markup distortion and likely increase welfare.[4] However, real-world carbon prices are typically much smaller than the estimated social cost of carbon (Carl and Fedor (2016); Revesz et al. (2017)). Under a suboptimally low carbon tax, coal markups would increase welfare by raising coal prices closer to their marginal social cost. In this case, incomplete pass-through would lower coal markups and reduce welfare.

Third, my analysis contributes to the literature on estimating tax pass-through in energy markets. Previous studies have often found heterogeneous pass-through of energy taxes, due to variation in market structure both across and within industries (e.g., Ganapati, Shapiro, and Walker (2016) in manufacturing; Pouliot, Smith, and Stock (2017) in transport fuels). Muehlegger and Sweeney (2017) find that pass-through in petroleum refining also varies by whether cost shocks are firm-specific or common across all firms. I find heterogeneous pass-through due to a combination of these two effects: spatial variation in the competitiveness of coal shipping, and variation in coal plants' sensitivity to relative fuel price shocks. To my knowledge, this is the first evidence that pass-through of a carbon tax in the U.S. electricity sector may be heterogeneous and incomplete.

Heterogeneous pass-through implies that the economic incidence of a carbon tax would likely vary across coal plants. I apply the theoretical tools of Weyl and Fabinger (2013) to translate pass-through to incidence, which reveals substantial heterogeneity in the share

[4] Coal also emits harmful local air pollutants such as SO_2, NO_x, and particulate matter. A tax greater than the social cost of carbon may partially internalize damages from these other pollutants. Hence, incomplete pass-through could reduce welfare even under a tax equal to marginal CO_2 damages.

of the implied tax burden borne by plants. While most plants bear the full decline in profits from a gas price drop, a subset of plants bear less than half, with the remainder coming out of railroad oligopoly rents. This finding contributes to the literature on environmental tax incidence, which has shown that imperfect competition and heterogeneous pass-through can shift the tax burden towards producers and make climate policy less regressive (Ganapati, Shapiro, and Walker (2016); Stolper (2017)). In my setting, shifting a share of the tax burden upstream from coal power plants may also benefit electricity consumers, potentially reducing the regressivity of a carbon tax.

Finally, my results contribute to the literature on fuel-switching between coal and natural gas. Recent decreases in the gas price have crowded out coal-fired generation, thereby reducing CO_2 emissions from the U.S. electricity sector. While several previous studies have estimated the magnitude of these environmental benefits (e.g., Holladay and LaRiviere (2017); Fell and Kaffine (forthcoming)), I show that decreasing coal markups have likely attenuated this shift away from coal. A simple counterfactual exercise suggests that short-run fuel substitution could have yielded 8 percent greater CO_2 abatement, if coal markups had not changed. This suggests that previous retrospective analyses may have understated the potential environmental benefits of a carbon tax, if the tax is large enough to drive coal markups close to zero and eliminate the countervailing effect of incomplete pass-through.

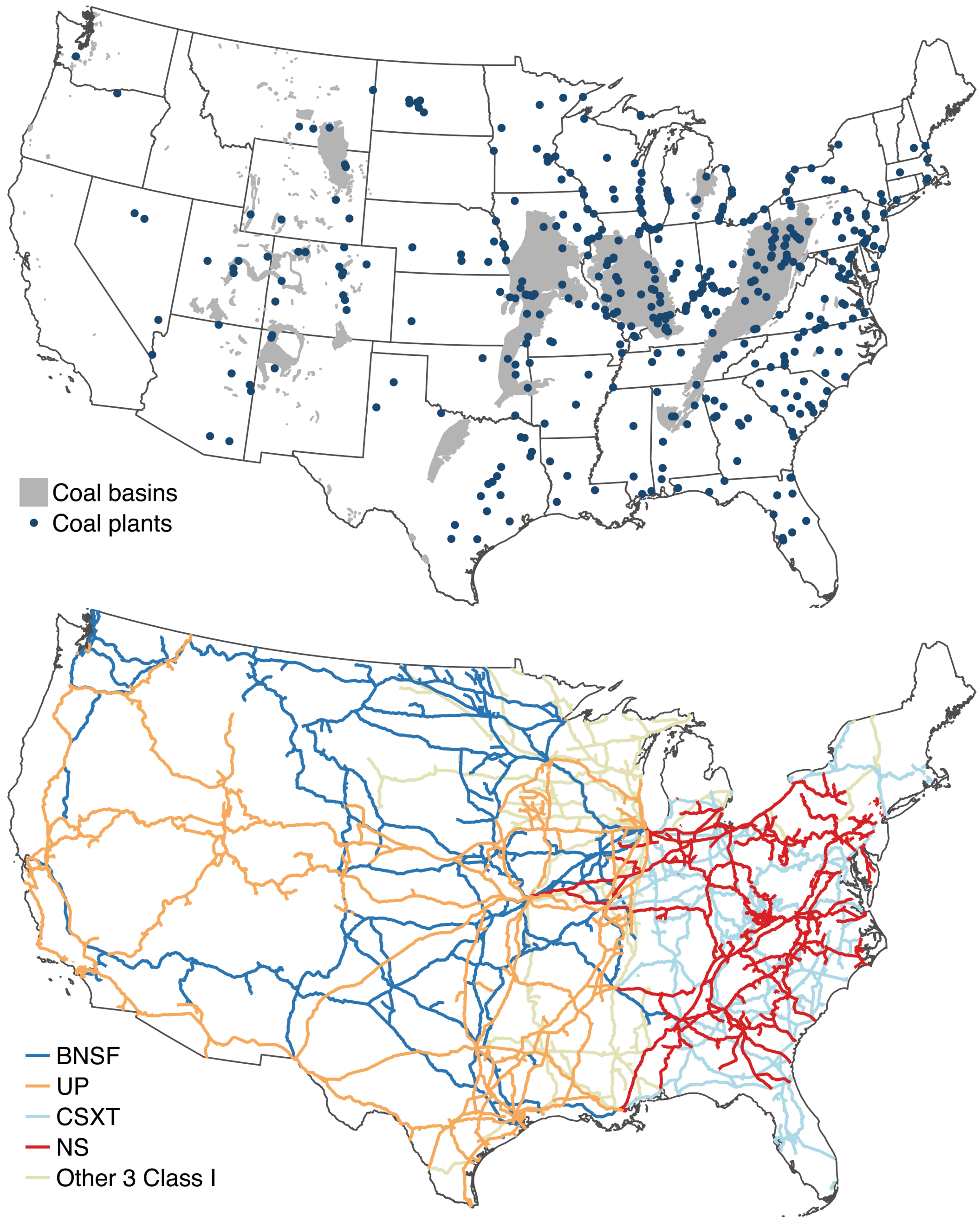
This paper proceeds as follows. Section 1.2 describes the institutions of U.S. coal markets and the recent boom in natural gas production. Section 1.3 develops a static oligopoly model to predict how railroads reoptimize markups as gas prices change. Section 1.4 outlines the data I use to implement my empirical strategy, detailed in Section 1.5. Section 1.6 reports results from estimating levels and changes of coal markups. Section 1.7 analyzes the implications of these results for climate policy. Section 1.8 concludes.

1.2 U.S. Coal Markets

U.S. coal markets feature three primary types of agents: mining firms, power plants, and transport intermediaries.[5] Mines are spatially concentrated in regions with productive coal deposits, most notably the Powder River Basin in northeastern Wyoming, and the Appalachian Basin in West Virginia and Kentucky. By contrast, coal power plants are spatially dispersed across the country, due to the regionally fragmented nature of electricity markets. Coal is heavy relative to its commodity value, and plants located far from mines incur substantial coal shipping costs (Joskow (1985)). Railroads are the dominant transportation mode, and a few large rail carriers deliver over 70 percent of coal shipments. Figure 1.2.1 maps the geographic configuration of coal producing regions, coal power plants, and major rail lines. Plants located on navigable waterways may also re-

[5]Over 90 percent of U.S. coal consumption occurs in the electric power sector. My analysis does not include other industrial consumers of coal, such as steel, cement, and paper manufacturers. I also ignore coal imports (less than 2 percent of U.S. consumption) and exports (roughly 3 percent of U.S. production).

Figure 1.2.1: U.S. Coal Geography



Notes: This top panel maps all productive deposits of power plant grade coal (i.e. bituminous and sub-bituminous) in the contiguous U.S. The vast majority of coal production occurs in three regions: the Powder River Basin in northeastern Wyoming; the Appalachian Basin in West Virginia and eastern Kentucky; and the Illinois Basin in southern Illinois and western Kentucky. Dots denote all 430 large coal-fired electric power plants that operated between 2002–2015. The bottom panel maps major rail lines owned and operated by the seven Class I rail carriers. Two rail carriers each dominate the West (BNSF, Union Pacific) and East (CSX Transportation, Norfolk Southern). I combine the remaining three carriers (Canadian National, Canadian Pacific, Kansas City Southern) into a single color, as these rail carriers cover smaller territories far from most major coal deposits.

ceive coal shipments via barge, a more competitive outside option with low barriers to entry. Barges contribute roughly 17 percent of coal deliveries.[6]

Four firms control most of the coal shipping industry, with two large rail carriers dominating both the western and eastern U.S. (see Figure 1.2.1). Ever since the Staggers Act of 1980 substantially weakened rail price regulations, railroads have been able to set freight shipping rates with limited government oversight (MacDonald (1989)).[7] In cases where a single rail carrier exhibits "market dominance" along a given route, regulators may intervene to prevent rail revenues from exceeding 180 percent of total variable costs.[8] This means that rail oligopolists have significant leeway to exercise market power and negotiate complex long-term contracts with power plants (Joskow (1988)). By allowing carriers to extract oligopoly rents and exploit economies of scale, the Staggers Act also spurred a series of railroad mergers; the 33 "Class I" railroads of 1980 have consolidated into the 7 Class I railroads of today (Schmidt (2001); Prater, Sparger, and O'Neil (2014)).[9]

Three factors have led to substantial spatial dispersion in coal-by-rail markups. First, unlike most commodities, coal consumption must occur in precise geographic locations with potentially limited access to transportation networks. While some power plants have the option to purchase coal from multiple rail carriers or via barge (by virtue of their locations), other plants must rely on a single rail carrier for all coal deliveries. Second, many plants are constrained to buy a particular type of coal, produced in only one mining region (Joskow (1987)). This further restricts plants' shipping options, as mines may also have limited access to rail and water networks.[10] Third, the resale of coal is cost-prohibitive, because infrastructure is built to carry coal *to* (not *away from*) plants (Busse and Keohane (2007); Jha (2015)). Hence, plants are unable to arbitrage away spatial price differences, allowing railroads to charge higher markups to plants with fewer shipping options.[11]

[6]Trucks also transport a small share of coal deliveries. However, trucking is relatively costly and likely cannot compete directly with rail and water (Busse and Keohane (2007)).

[7]MacDonald (2013) offers a comprehensive history of U.S. railroad regulation.

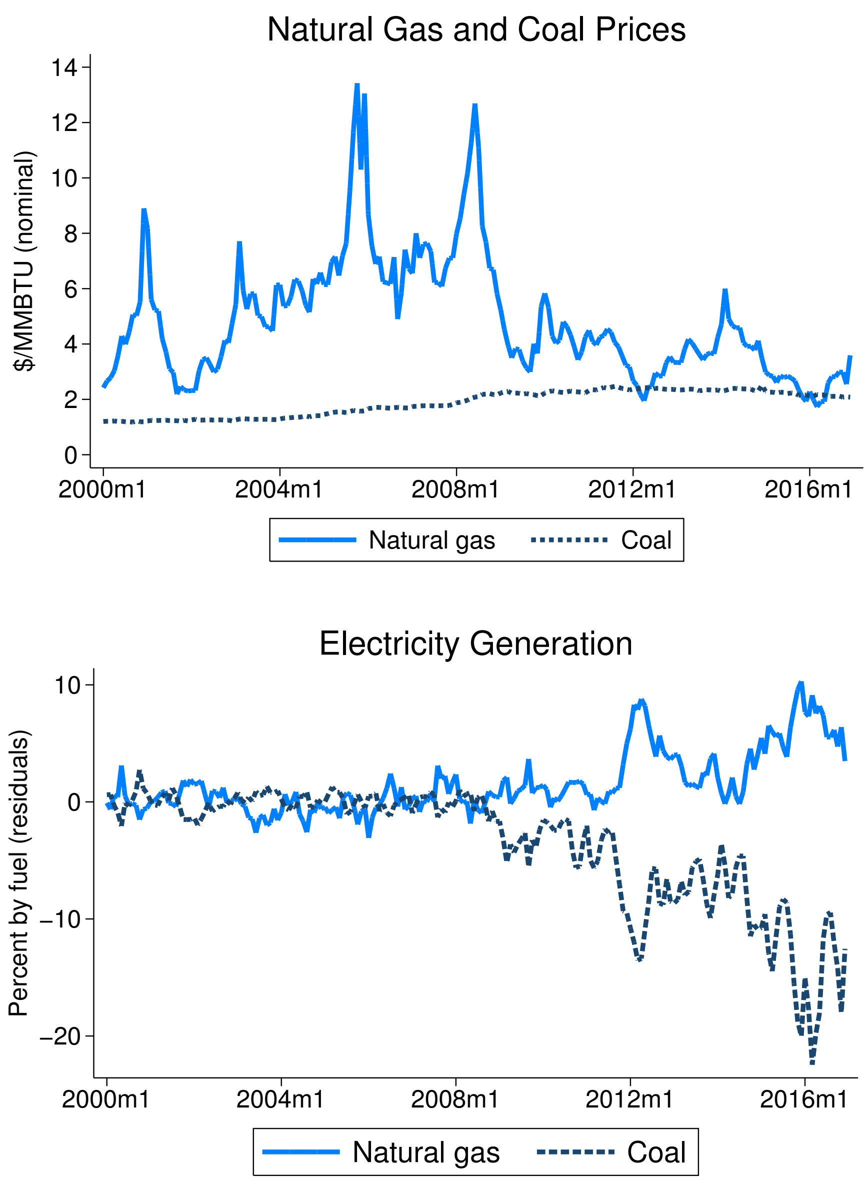
[8]In practice, regulators loosely interpret this threshold such that railroads may earn an adequate return on investment (Wilson (1996)). While the Surface Transportation Board reviews only 1–7 rate challenges each year, rate cases for coal shipping occur more frequently than for all other commodities combined (`https://www.stb.gov/stb/industry/Rate_Cases.htm`).

[9]The Class I designation includes carriers with annual operating revenues exceeding $453 million. These seven firms account for approximately 69 percent of rail mileage and 94 percent of rail freight revenues.

[10]Coal's physical characteristics vary across coal regions, and even across mines within a region. Plants typically value coal with high energy content (or BTUs per ton), and with low sulfur and ash content (which create local air pollution). Plants self-calibrate to a pre-specified mix of coal attributes, and deviations can reduce the efficiency of boilers and pollution-control devices (Kerkvliet and Shogren (1992)). Also, many plants comply with SO_2 regulations by burning low-sulfur coal from Wyoming's Powder River Basin (Schmalensee and Stavins (2013)). If such a plant has access to two rail carriers and the Ohio River, but only one rail carrier connects to the Powder River Basin, then it has effectively one shipping option.

[11]Power plants may purchase coal directly from rail carriers; alternatively, plants may purchase freight services from railroads and separately purchase coal from upstream mines. This distinction does not

Figure 1.2.2: U.S. Fuel Prices and Electricity Generation



Notes: The top panel reports the monthly average Henry Hub spot prices for natural gas, and the monthly average cost of coal delivered to electric generating plants. This average cost of coal does not account for heterogeneous coal attributes, including BTU content by weight, sulfur content, or ash content. The bottom panel plots U.S. monthly electricity generation by fuel as a percent of total monthly generation, controlling for month fixed effects and a 2000–2008 time trend for each fuel.

U.S. coal consumption has declined over the past decade, largely due to decreases in the price of natural gas. Technological advances in hydraulic fracturing ("fracking") have led to a boom in natural gas extraction, causing a historic drop in U.S. gas prices.[12] Because coal plants compete directly with natural gas generation in electricity markets, low gas prices have crowded out coal-fired electricity generation. The top panel of Figure 1.2.2 shows how the fracking boom has led to a decline in U.S. gas prices since 2008, and the bottom panel illustrates how the electricity sector has shifted towards gas and away from coal. The corresponding decrease in coal demand has likely caused rail oligopolists to reoptimize coal markups. These observed changes in markups can predict what might occur under a carbon tax, which would similarly disadvantage coal relative to low-carbon natural gas (Cullen and Mansur (2017)). If coal markups decrease (increase), this would dampen (magnify) the carbon tax price signal as it passes along the coal supply chain. This effect would likely be heterogeneous across plants, due to variation in pre-existing markups *and* variation in plants' exposure to gas-fired competition.[13]

1.3 Theoretical Framework

In this section, I develop a simple Cournot oligopoly model of railroad intermediaries who sell coal to power plants. This allows me to predict how markups should respond to gas price changes, which depends on three key dimensions of plant heterogeneity: (1) number of potential rail carriers; (2) availability of water transport as a more competitive outside option; and (3) price elasticity of demand for coal as an input to electricity production.

1.3.1 Symmetric Rail Oligopoly

Consider power plant j that is a price-taker in the market for coal. This plant consumes a specific type of coal from origin o, which is produced at constant marginal cost by a perfectly competitive mining sector.[14] Plant j is fully captive to N_{oj} identical rail

---

affect the economic interpretation of delivered coal markups. My analysis treats rail intermediaries as both the owners of the commodity and the providers of freight services.

[12] Two separate technological innovations have facilitated the "fracking boom": horizontal drilling and hydraulic fracturing. Fitzgerald (2013) provides a comprehensive overview of these technological advances and their effect on the costs of gas extraction. The physical properties of natural gas make it expensive to export, which is why a domestic supply glut has depressed U.S. gas prices.

[13] Low gas prices have not impacted all coal plants equally. For a coal plant located in an electricity market with many gas-fired competitors, a negative gas price shock will likely cause its coal demand to decrease. For a coal plant in a market without any gas-fired competitors, the same gas price shock may have no effect on its coal demand. Coal plants also vary in their productive efficiency, and low gas prices should disproportionately hurt relatively inefficient plants.

[14] This assumption greatly simplifies my theoretical framework. In reality, coal supply may be upward-sloping, and mining need not be perfectly competitive, especially in Wyoming's Powder River Basin where a few large firms dominate mining operations (Atkinson and Kerkvliet (1986)). Appendix A.1.3 discusses the welfare implications of alternative market structures.

carriers for its coal deliveries from origin o, and each rail carrier i chooses the best-response quantity of coal q_{ioj} that maximizes its profits on route oj. In equilibrium, plant j consumes $N_{oj}q_{ioj} = Q_{oj}$ units of coal at price P_{oj}. Plant j cannot resell its purchased coal, meaning that P_{oj} is not restricted by a binding arbitrage condition and rail carriers may effectively treat each plant as its own isolated coal market.

Rail carrier i's profits from selling coal from origin o to plant j are:

$$\pi_{ioj}(q_{ioj}) = q_{ioj}\Big[P_{oj}(Q_{oj};\mathbf{Z}_{oj}) - C_o - S(\mathbf{T}_{oj})\Big] - F_{oj} \tag{1.1}$$

where P_{oj} is the plant j's inverse demand for coal shipped from origin o, as a function of Q_{oj} and a vector of parameters $\mathbf{Z}_{oj}$. C_o is the exogenous mine-mouth coal price in origin o. The function $S(\mathbf{T}_{oj})$ denotes the average per-unit cost of shipping coal on route oj, where $\mathbf{T}_{oj}$ is a vector of transportation cost parameters, including rail mileage, diesel costs, and the opportunity cost of a rail car. Finally, the oligopolist incurs F_{oj}, a fixed cost of entry on shipping route oj.[15] In reality, carrier i maybe also be subject to regulatory oversight, but I abstract from rail regulation in this simple model.

Taking rail carrier i's first-order condition, and rearranging in terms of a price-cost markup μ_{oj}:

$$\mu_{oj} \equiv P_{oj} - C_o - S(\mathbf{T}_{oj}) = -\left(\frac{\theta_{oj}}{N_{oj}}\right)\frac{\partial P_{oj}}{\partial Q_{oj}}Q_{oj} \tag{1.2}$$

where the "conduct parameter" θ_{oj} equals 1 under a pure Cournot oligopoly and 0 under perfect competition.[16] Plant j's markup depends on both its coal transportation options and its demand for type-o coal. If plant j is captive to a single rail carrier (i.e., $N_{oj} = 1$), it should face higher markups than if multiple carriers were competing on route oj. At the same time, if plant j is located on a navigable waterway and can receive type-o coal via barge, this should limit railroads' ability to set high markups. Since waterways have less restricted usage rights and lower barriers to entry, I treat barge shipments as a competitive outside option (i.e., $\theta_{oj} = 0$). Finally, if plant j's inverse demand for coal is relatively inelastic, it should face relatively higher markups, all else equal.

1.3.2 Comparative Statics for Coal Markups

Coal demand depends on the price of natural gas, because the two fuels compete in electricity dispatch. If the gas price decreases (increases), a coal plant may supply less

[15] I use a symmetric oligopoly model for tractability. In reality, each firm's shipping routes are constrained by track ownership and trackage rights, implying non-identical costs $S(\mathbf{T}_{oj})$ and F_{oj}. For simplicity, I assume that quantity q_{ioj} does not enter into $S(\mathbf{T}_{oj})$, which ignores rail capacity constraints or increasing returns to scale in shipping. My empirical analysis relaxes this assumption.

[16] $\theta_{oj} = \frac{\partial Q_{oj}}{\partial q_{ioj}}$ is identical for all i, by symmetry. I use this "conduct parameter" formulation for notational convenience (following Atkin and Donaldson (2015); Bergquist (2017)), and I treat θ_{oj} only as a continuous heuristic for distance from perfect competition. Calibrating θ_{oj} as a structural parameter can be problematic, as it only takes on a well-defined interpretation at a few values (Corts (1999)).

(more) electricity at a given coal price. The gas price also influences the elasticity of coal demand, by determining the range of coal prices over which a coal plant is marginal in electricity supply. A marginal plant has (locally) elastic coal demand, because its coal consumption responds to small changes in coal price. At lower coal prices, a coal plant will be inframarginal and its strict capacity constraint will bind. This translates to (locally) inelastic coal demand, as small changes in coal price will not change its coal consumption.

Figure 1.3.3 presents a stylized electricity market to illustrate how a negative gas price shock impacts both the level and slope of coal demand. There is a single coal plant with constant marginal cost, and an upward-sloping supply of gas-fired generation. Each technology's marginal costs scale with its respective fuel price, and the aggregate electricity supply curve depends on *both* fuel prices. The top panels show four supply curves, for four combinations of coal price (low, high) and gas price (high, low). In reality, electricity demand is stochastic and extremely inelastic; for simplicity, this stylized example assumes electricity demand is deterministic and perfectly inelastic.

At a given gas price, the plant's coal demand is the 1-to-1 mapping between coal price and coal consumption. Under a high coal price and high gas price (i.e., the solid supply curve in the top-right panel), the coal plant is marginal in the electricity market and generates at 70 percent capacity. Hence, it demands 70 percent of its throughput capacity for coal, or Q^* in the bottom-left panel. Comparing the bottom two panels, the gas price governs the range of coal prices for which the plant is marginal, and coal demand is not vertical. A negative gas price shock causes inverse coal demand to shift down *and* become less steep.[17]

Using my symmetric oligopoly model, I can derive how rail carriers should reoptimize coal markups in response to gas price changes. Let Z_j denote the gas price of coal plant j's competitors, which enters plant j's inverse demand function as an element of the parameter vector $\mathbf{Z}_{oj}$. The change in markup μ_{oj} that results from a small change in gas price Z_j is:[18]

$$\frac{d\mu_{oj}}{dZ_j} = \frac{\dfrac{\partial P_{oj}}{\partial Z_j}(2 + E_{D_{oj}} - N_{oj}) - \dfrac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}\theta_{oj}}{2 + E_{D_{oj}}} \tag{1.3}$$

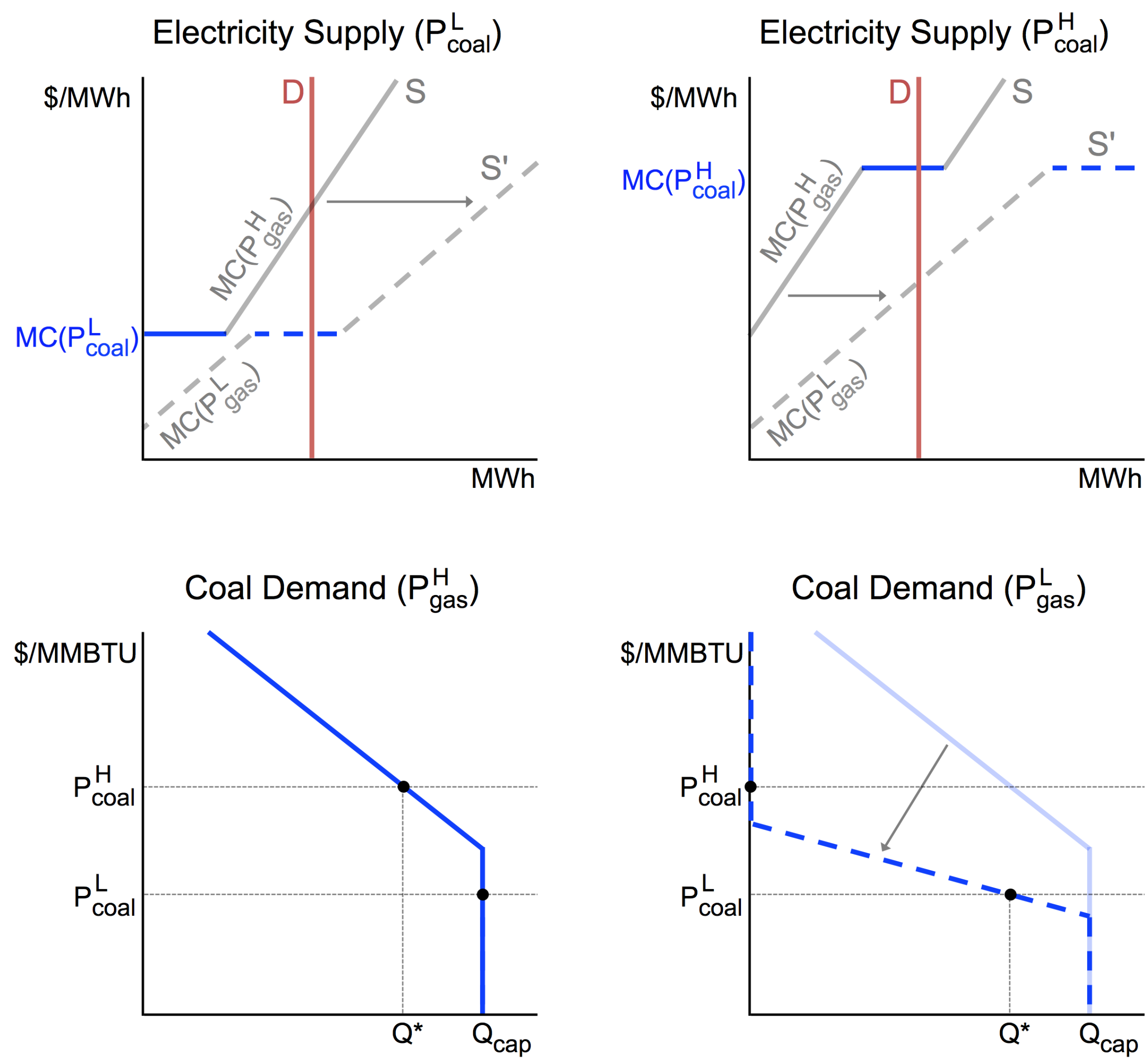
where $E_{D_{oj}}$ is the elasticity of the slope of inverse demand scaled by the degree of competitiveness $\frac{\theta_{oj}}{N_{oj}}$:

$$E_{D_{oj}} \equiv \left(\frac{\partial^2 P_{oj}}{\partial Q_{oj}^2}\right) \left(\frac{\partial P_{oj}}{\partial Q_{oj}}\right)^{-1} Q_{oj} \left(\frac{\theta_{oj}}{N_{oj}}\right) \tag{1.4}$$

[17] In reality, electricity dispatch may not order plants from lowest-to-highest cost, and plants may not maximize short-run profits. Demand realizations come from a continuous probability distribution, and electricity is not storable. Coal storage enables to plants to hedge against uncertainty in electricity markets, which must clear instantaneously. Hence, coal markets clear on a longer timescale, and coal demand should not have sharp kinks.

[18] Appendix A.1.1 provides a full derivation of this comparative static.

Figure 1.3.3: Coal Demand and Natural Gas Prices



Notes: This figure presents a stylized electricity market to illustrate the relationship between gas prices and coal demand. There is one coal generator with a fixed capacity, and constant marginal cost at a given coal price ($MC(P_{\text{coal}})$, in blue). There is also an upward-sloping supply of many small natural gas generators, with marginal costs that scale multiplicatively with the gas price ($MC(P_{\text{gas}})$, in grey). Electricity demand (D) is perfectly inelastic, and deterministic (for simplicity). The top panels show four electricity supply curves, each for a given combination of coal price (low in the left panel, high in the right panel) and gas price (high for solid lines, low for dashed lines). The bottom panels translate the coal plant's electricity production into its corresponding demand for coal (MWh out as a function of MMBTU of coal in, given the plant's fixed production technology). Under a high gas price (P^H_{gas}), the coal plant consumes at full capacity (Q_{cap}) given a low coal price (P^L_{coal}) and at Q^* given a high coal price (P^H_{coal}). If the gas price decreases to P^L_{gas}, the coal plant is now marginal given P^L_{coal} (where it had been inframarginal) and above the margin given P^H_{coal} (where it had been marginal). The decrease in gas price has caused inverse coal demand at Q^* to shift down and become less steep.

Equation (1.3) depends on the level, slope, and curvature of plant j's inverse demand. $\frac{\partial P_{oj}}{\partial Z_j}$ captures how gas price affects the *level* of inverse coal demand. If a negative gas price shock (i.e. $dZ_j < 0$) causes plant j's inverse coal demand to shift down as in Figure 1.3.3, then $\frac{\partial P_{oj}}{\partial Z_j} > 0$. The cross-partial $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j}$ captures how gas price affects the *slope* of inverse coal demand. If lower gas prices make inverse coal demand less steep (i.e. if $dZ_j < 0$ causes $\frac{\partial P_{oj}}{\partial Q_{oj}}$ to become less negative), then $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} < 0$. Finally, the change in markup depends on the degree to which inverse demand is concave ($E_{D_{oj}} > 0$) or convex ($E_{D_{oj}} < 0$). More concave demand will tend to increase $\frac{d\mu_{oj}}{dZ_j}$, while more convex demand will tend to decrease $\frac{d\mu_{oj}}{dZ_j}$.[19]

These three features of coal demand interact with route oj's rail market size (N_{oj}) and structure (θ_{oj}) to jointly determine how railroads should reoptimize markups when the gas price changes. The sign of $\frac{d\mu_{oj}}{dZ_j}$ is theoretically ambiguous and depends on the relative sizes of $\frac{\partial P_{oj}}{\partial Z_j}$, $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j}$, and $E_{D_{oj}}$, which may vary considerably across heterogeneous coal plants. Rail carrier behavior may also depart from the predictions of this simple model, especially if regulatory constraints bind or if markups are not truly independent across plants.[20] Below, I econometrically estimate the degree to which observed gas price changes have caused changes in coal markups. I directly estimate plant-specific demand parameters $\frac{\partial P_{oj}}{\partial Z_j}$, $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j}$, and $E_{D_{oj}}$, which I use to construct an empirical approximation of $\frac{d\mu_{oj}}{dZ_j}$. This allows me to test the theoretical predictions and empirical magnitudes of Equation (1.3).

1.4 Data

My analysis combines several publicly available datasets, published by U.S. government agencies. This section highlights the core datasets for my empirical analysis, including data on coal shipments, power plants, and the U.S. rail network.[21] It also describes how I use GIS data to construct a measure of plants' rail captiveness.

1.4.1 Data Sources

The Energy Information Administration's (EIA) Form 923 collects detailed data on coal deliveries to power plants, for all large U.S. coal plants. These data are at the month-

[19]This is a standard finding from the pass-through literature on imperfectly competitive product markets, where the pass-through rate is closely related to the curvature of demand (Weyl and Fabinger (2013)).

[20]In reality, rate regulation prevents carriers from extracting full (unconstrained) oligopoly rents. This simple model does not account for multiple-market negotiations between rail carriers or dynamic interactions between carriers and plants.

[21]Appendix A.2 describes each of these datasets in further detail, while also describing how I merge datasets and construct key variables.

shipment level, where "shipments" aggregate deliveries received on a single purchase order or contract, in a given month, with the same supplier, county of origin, and coal rank (i.e. bituminous vs. sub-bituminous). For each observation, EIA reports the total tons of coal delivered; their average BTU, sulfur, and ash content by weight; and the primary modes of transportation (e.g., rail, barge, truck). EIA also classifies each shipment as either a long-term contract or a spot market transaction.

Coal plants must report the average prices for each observation, inclusive of commodity costs, transportation costs, and markups. EIA redacts price data for independent power producers, and I only observe prices for utility-owned plants. My empirical analysis focuses on this subset of plants, which represent 77 percent of coal deliveries since 2002.[22] Because coal is a heterogeneous commodity without a uniform price index, I control for average mine-mouth prices at the county-year level, published in EIA's Annual Coal Report.

I merge coal shipment data with several EIA datasets on power plant characteristics (Form 860), operations (Forms 906 and 923), and pollution abatement (Forms 767, 860, and 923). The EPA eGRID database reports each plant's power control area (PCA), or its region on the electricity transmission grid. To estimate plant-specific coal demand parameters, I leverage data from the EPA's Continuous Emissions Monitoring System (CEMS), which report hourly generation and emissions for all large fossil fuel generating units.[23] This allows me to estimate each coal unit's probability of generating in a given hour, conditional on the relative prices of coal and natural gas.

I use detailed GIS data on the U.S. rail network published by the Bureau of Transportation Statistics (BTS). I apply a graph algorithm to find the shortest path along the rail network connecting each coal-producing county to each power plant destination.[24] Then, I calculate the proportion of each shortest route owned or operated by each of the 7 Class I rail carriers, and assign a "dominant" (modal) carrier to each route. BTS also reports the average traffic density of rail lines, which I integrate over the full length of each route as a proxy for rail network congestion. To control for time-series variation in shipping costs, I use the Association of American Railroads' (AAR) monthly fuel price index, which compiles survey data on actual diesel prices paid by railroad operators.[25] I also calculate each plant's proximity to a navigable river, Great Lake, or coastline. This allows me to identify the subset of plants with the option to receive waterborne coal deliveries.

[22] Beginning in the late 1990s, electricity market restructuring forced many vertically integrated utilities to sell their coal plants. Most of these divestments were in just four states (Pennsylvania, Illinois, Ohio, and New York), and the vast majority occurred before the start of the fracking boom. Previous research has focused directly on the effects of coal plant divestment (Cicala (2015); Chan et al. (2017)), and these studies have obtained non-disclosure agreements with EIA to unmask prices for non-utility plants.

[23] Most coal plants comprise multiple generating units (or boilers), each with different operating constraints and variable costs.

[24] Hughes (2011) applies a similar algorithm to calculate the shortest rail distance, and finds that GIS-derived shortest distances closely approximate (yet slightly understate) actual rail shipping distances. Appendix A.3 describes this shortest-distance algorithm in detail.

[25] Diesel purchases represent roughly half of railroads' total variable transportation costs.

1.4.2 Defining Rail Captiveness

I treat each power plant's location on the rail network as pre-determined. Plant geographic locations are obviously fixed, and I exclude the few plants constructed during my 2002–2015 sample period. More importantly, the rail network was largely static throughout this period, with very few changes in the ownership or trackage rights of individual rail lines.[26] This means that each coal plant has faced the same set of potential rail carriers.

I partition plants into two time-invariant groups, "captive" and "non-captive", based on their locations on the rail network and the counties from which they purchase coal. I define the "captive" group as plants that either (i) become unconnected from the rail network after removing any single Class I carrier, or (ii) become unconnected from *all* observed trading partners after removing the dominant carrier along each origin-destination route. For example, consider a plant that only purchases coal from two counties in Wyoming. I classify this plant as captive if a single Class I carrier controls all terminal nodes within a 7-mile radius. This plant is also captive if, after removing the dominant carrier on its shortest route to *each* Wyoming county, the new shortest routes *both* increase by over 300 miles.[27]

1.5 Empirical Strategy

The goal of my empirical analysis is to estimate how coal-by-rail markups vary across heterogeneous power plants. First, I estimate how markup *levels* vary across captive vs. non-captive plants, using a nearest-neighbor matching strategy. Then, I estimate how gas price shocks cause markups to *change* differentially across two sources of heterogeneity: (i) variation in market power, and (ii) variation in plants' sensitivity to competition from natural gas generation. These difference-in-differences estimates combine data on plants' transportation options with plant-specific coal demand estimates, using the comparative static $\frac{d\mu}{dZ}$ from my oligopoly model as a guide.

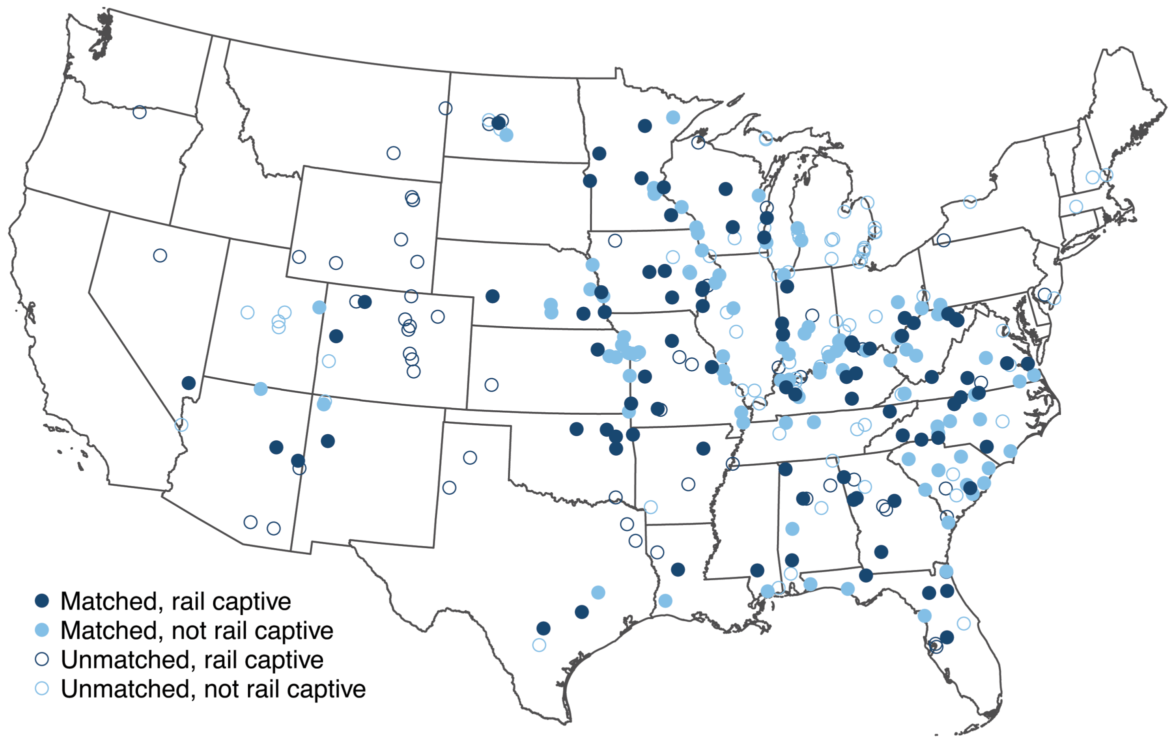
1.5.1 Matching Captive vs. Non-Captive Plants

Rail captiveness is not randomly assigned, and we might expect captive and non-captive plants to differ systematically. Because captiveness depends on geography, plants of each group might be spatially concentrated and burn similar types of coal, have similar operating characteristics, or face similar market conditions. Any observed or unobserved differences that are correlated with rail markups would lead to biased estimates of the markup differential between captive and non-captive groups.

[26] The last merger between Class I carriers occurred in 1999. 99.3 percent of Class I track mileage maintained constant ownership since 2006, the earliest year of available BTS data.

[27] Each of these thresholds is quite conservative. 7 miles is the 95th percentile of plants' distance to the closest rail node. A 300-mile increase in distance implies a 20 percent increase over the median delivered coal price. Appendix A.3 discusses my choice of thresholds, while Appendix A.5 reports sensitivity analysis.

Figure 1.5.4: Nearest Neighbor Matching



Notes: This map plots the location of all 324 utility-owned coal power plants in the full 2002–2015 sample. Rail captive plants are in navy and non-captive plants are in light blue; plants in the matched sample are filled, while unmatched plants are hollow. Match criteria: up to k nearest neighbors ($k = 3$), with a maximum distance of 200 miles; exact matches on coal rank; and removing non-utility and non-rail plants.

To address this identification challenge, I apply a nearest-neighbor matching strategy in the tradition of Heckman, Ichimura, and Todd (1997). I match each plant in the "captive" group to k plants in the "non-captive" group with the closest geographic proximity, with a maximum distance of 200 miles between matched plants. I also force exact matches on plants' preferred coal type from the pre-fracking period (2002–2006). I omit plants that rely exclusively on non-rail shipping modes (i.e. barges and trucks), and plants that are not utility-owned (for which I do not observe coal price data). I assign nearest-neighbor weights as the inverse of the number of matches. For example, if a non-captive plant is one of 3 matches for captive plant A and one of 2 matches for captive plant B, it receives a weight of $\frac{1}{3} + \frac{1}{2} = \frac{5}{6}$. Matched captive plants receive a weight of 1, while all unmatched plants receive a weight of 0. This ensures that weights sum to twice the number of matched captive plants.[28]

[28]My approach closely resembles that of Cicala (2015), who also imposes a maximum match distance of 200 miles. He and Lee (2016) similarly match coal plants that sell gypsum byproduct vs. non-gypsum plants.

Figure 1.5.4 maps the full sample of captive and non-captive plants. This reveals broad geographic overlap, except for certain regions where plants tend to be either only captive (i.e. the western Great Plains) or only non-captive (i.e. Michigan). This map also shows the outcome of my matching algorithm with $k = 3$ nearest neighbors. Matched plants tend to be located near multiple plants of the opposite group. While I allow matches up to 200 miles apart, most matched plants have a nearest neighbor within 100 miles.[29]

Table 1.5.1 presents summary statistics for both groups of plants from 2002–2006, including plant and coal characteristics. In the full sample, captive plants systematically purchase more low sulfur, sub-bituminous coal, and are more likely to participate in wholesale electricity markets. However, nearest-neighbor weights adjust the distributions of both captive and non-captive plants such that they are no longer statistically different. While geographic proximity alone does not ensure that matched non-captive plants can serve as plausible controls, distance-based matching yields covariate balance across a wide range of observables.

1.5.2 Markup Levels

I begin by estimating differences in markups between captive and non-captive plants. I estimate the following OLS regression, which is analogous to the markup expression I derive in Equation (1.2):

$$P_{ojms} = \tau D_j + \beta_C \mathbf{C}_{ojms} + S(\mathbf{T}_{ojms}\,;\beta_T) + \beta_X \mathbf{X}_{jm} + \eta_o + \delta_m + \varepsilon_{ojms} \tag{1.5}$$

P_{ojms} is the average delivered price of coal, for shipment s from county o to plant j in month m. D_j is an indicator for rail captiveness, and the coefficient τ estimates the average differential markup faced by captive plants, relative to non-captive plants. Since I do not directly observe coal markups, I use price as an outcome variable and control for shipment-level variation in both commodity value and shipping costs (i.e. $C_o + S(\mathbf{T}_{oj})$ in Equation (1.2)). I also include nearest-neighbor weights and plant-specific controls ($\mathbf{X}_{jm}$), in the style of a doubly robust estimator (Wooldridge (2007)). After controlling for both county fixed effects (η_o) and month-of-sample fixed effects (δ_m), the remaining variation in P_{ojms} is close to the variation I would use in the ideal experiment: comparing the price of two identical coal shipments to two otherwise identical coal plants, where only one plant is rail captive.

The matrix $\mathbf{C}_{ojms}$ controls for determinants of commodity value, including average heat, sulfur, and ash content; coal rank; and the average annual mine-mouth price for coal produced in county o. $\mathbf{C}_{ojms}$ also includes dummies for spot market transactions and contracts expiring within 2 years, since plants pay higher prices for (longer) contracts that minimize the risk of supply disruptions.[30] The matrix $\mathbf{T}_{ojms}$ controls for the two

[29]Appendix A.5.1 reports sensitivity analysis on the number of matches (k) and maximum match distance.

[30]Wolak (1996) finds that coal plants simultaneously purchase on long-term contracts and the spot market, even as contract purchases tend to have higher delivered prices. Jha (2017) estimates that the

Table 1.5.1: Summary Statistics – Captive vs. Non-Captive Coal Plants (2002–2006)

	Full sample			Matched sample ($k = 3$)		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	908.19	865.73	42.46	900.89	940.81	−39.92
	(780.72)	(742.77)	[0.57]	(69.32)	(92.86)	[0.73]
Coal-fired capacity (MW)	806.13	760.84	45.29	815.64	822.25	−6.61
	(738.72)	(703.91)	[0.52]	(66.58)	(89.73)	[0.95]
Number of coal units	2.36	2.62	−0.26	2.59	2.60	−0.01
	(1.32)	(1.64)	[0.08]*	(0.15)	(0.15)	[0.96]
Coal unit vintage (year)	1968.85	1962.88	5.97	1966.25	1962.46	3.79
	(13.90)	(13.34)	[0.00]***	(1.39)	(1.41)	[0.06]*
Annual capacity factor	0.63	0.60	0.03	0.63	0.63	−0.00
	(0.17)	(0.17)	[0.04]**	(0.01)	(0.01)	[0.93]
Heat rate (MMBTU/MWh)	11.09	11.06	0.03	10.97	10.73	0.24
	(1.40)	(1.52)	[0.86]	(0.14)	(0.13)	[0.20]
Scrubber installed	0.36	0.29	0.07	0.28	0.28	−0.00
	(0.48)	(0.45)	[0.12]	(0.05)	(0.06)	[0.95]
Market participant	0.49	0.71	−0.22	0.46	0.50	−0.04
	(0.50)	(0.46)	[0.00]***	(0.05)	(0.06)	[0.62]
	Full sample			Matched sample ($k = 3$)		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	48.82	44.00	4.82	47.44	44.96	2.48
	(47.90)	(43.70)	[0.29]	(4.10)	(5.19)	[0.71]
Sulfur content (lbs/MMBTU)	0.87	1.02	−0.15	0.79	0.85	−0.05
	(0.61)	(0.79)	[0.03]**	(0.06)	(0.07)	[0.56]
Ash content (lbs/MMBTU)	8.46	8.96	−0.50	8.08	7.94	0.14
	(4.21)	(8.24)	[0.46]	(0.37)	(0.42)	[0.80]
Share spot market	0.19	0.19	−0.00	0.18	0.16	0.02
	(0.29)	(0.25)	[0.87]	(0.03)	(0.02)	[0.65]
Share contracts expiring $\leq$ 2 years	0.22	0.24	−0.01	0.19	0.19	−0.00
	(0.25)	(0.26)	[0.61]	(0.02)	(0.02)	[0.97]
Share sub-bituminous	0.41	0.31	0.10	0.43	0.41	0.02
	(0.47)	(0.42)	[0.03]**	(0.05)	(0.06)	[0.80]
Average rail distance (miles)	554.91	620.34	−65.43	565.67	583.36	−17.69
	(385.90)	(417.90)	[0.12]	(40.37)	(42.48)	[0.76]
	Full sample			Matched sample ($k = 3$)		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	94	149	243	49	59	108
Preferred coal rank: sub-bituminous	77	76	153	36	36	72
Non-rail plants	17	14	31	0	0	0
Utility plants	148	176	324	87	96	183
Total plants	190	240	430	87	96	183

Notes: This table compares coal plants captive to a single rail carrier to non-captive plants. The left three columns include all CEMS electric power plants with at least 1 coal generating unit in 2002–2015, and with at least 1 reported coal delivery in both 2002–2006 and 2007–2015. The right three columns are weighted by nearest-neighbor matches, with unmatched plants receiving weight 0, matched captive plants receiving weight 1, and matched non-captive plants weighted by the inverse of the number of matches. Matching criteria: up to k nearest neighbors ($k = 3$), with a maximum distance of 200 miles; exact matches on preferred coal rank; and removing non-utility and non-rail plants. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

primary factors affecting the cost of transporting coal: the shortest rail distance between coal county o and plant j, and the average diesel price paid by rail carriers in month m. $\mathbf{T}_{ojms}$ also includes the log of shipment size, as marginal freight costs are likely decreasing in tons shipped. Finally, $\mathbf{T}_{ojms}$ includes the proportion of the shortest oj route on rail lines with high traffic density, to allow for higher costs along more congested routes. The function $S(\,\cdot\,)$ flexibly models shipping costs as the four-way interaction of the components of $\mathbf{T}_{ojms}$.[31] The matrix $\mathbf{X}_{jm}$ controls for both predetermined and time-varying plant characteristics, including each of the variables in panel A of Table 1.5.1. I cluster standard errors at the plant level, which allows for arbitrary within-plant serial correlation. In order to adjust the distribution of captive vs. non-captive plants, while also inflating each observation by the quantity of coal transacted, I weight by the product of the nearest-neighbor weights and shipment size.[32]

To interpret $\hat{\tau}$ as causal, D_j must be uncorrelated with plant unobservables, after nearest-neighbor matching and conditioning on observable characteristics in $\mathbf{X}_{jm}$. Captiveness is geographically predetermined, and 85 percent of matched plants predate the 1980 Staggers Act, which effectively legalized rail price discrimination. These plants likely could not have strategically influenced their degree of captiveness. It is also unlikely that coal plant unobservables impacted railroads' decisions to consolidate and increase market power, given that individual coal plants are small relative to the rail network. Spurious correlations could violate this identifying assumption if, for example, captive plants tended to have less sophisticated managers. As such a violation is unlikely, I interpret $\hat{\tau}$ as the causal effect of rail captiveness on markups. My subsequent results do not hinge on this interpretation.

1.5.3 Coal Demand Estimation

My theoretical framework illustrates how changes in markups likely depend on coal plants' sensitivity to the natural gas price. In order to account for this additional source of variation, I estimate plant-specific coal demand curves. In most settings with detailed data on production functions, this demand would follow from applying the Envelope Theorem to the firm's profit function at different factor prices, and inverting its production function to solve for the corresponding input quantity. However, four features of electricity markets make this approach infeasible for deriving coal demand.

First, regulated coal plants do not behave as short-run profit-maximizers; their profits are calibrated to a fixed rate-of-return, and they are not the residual claimants on

average regulated coal plant is willing to trade a \$1.66 increase in expected delivered coal price for a \$1.00 decrease in the standard deviation of delivered coal price.

[31] Appendix A.5 demonstrates that Equation (1.5) is not sensitive to the components of $\mathbf{C}_{ojms}$ and $\mathbf{T}_{ojms}$, which supports my interpretation of $\hat{\tau}$ as an average differential *markup*. Misspecification of these cost controls would mean that $\hat{\tau}$ might confound markups and cost differences.

[32] Observations in EIA's coal delivery data vary substantially by size, and this enables me to estimate the differential markup for the average ton of coal shipped, giving each ton of coal equal weight. Equation (1.5) also includes the log of shipment size as a linear control, as part of $\mathbf{T}_{ojms}$.

marginal costs of coal purchases or marginal revenues from electricity production. Second, for plants that do not participate in wholesale electricity markets, production decisions depend on complex engineering algorithms designed to balance the electricity grid; such algorithms may not define a marginal electricity price. Third, it would be extremely difficult to characterize the spatial and temporal constraints of the grid, where supply must respond to instantaneous changes in electricity demand, subject to available transmission capacity. Fourth, coal plants face their own dynamic operating constraints; delayed startup/shutdown decisions, ramping constraints, and maintenance outages would imply a unique state-dependent objective function for each plant.[33]

For these reasons, I estimate coal demand using a semi-parametric policy function approach, following Davis and Hausman (2016).[34] I predict electricity generation conditional on market conditions and fuel prices, allowing me to infer plant-specific coal demand curves (as in Figure 1.3.3). For each coal generating unit, I estimate the following time series regression, where CF_{uh} is unit u's capacity factor (i.e. generation divided by capacity) in hour h:

$$CF_{uh} = \sum_{b} \alpha_{ub} \mathbf{1}[G_{uh} \in b] + \sum_{b} \gamma_{ub} \mathbf{1}[G_{uh} \in b] \cdot CR_{ud} + \zeta_u CR_{ud} + \xi_u \mathbf{G}_{uh} + \omega_{uh} \tag{1.6}$$

Each unit-specific regression predicts unit u's generation conditional on *aggregate* fossil electricity generation in u's market region (G_{uh}, in discrete bins b), the daily ratio of u's marginal costs relative to the marginal costs of gas generators (CR_{ud}), and a matrix of controls ($\mathbf{G}_{uh}$).[35]

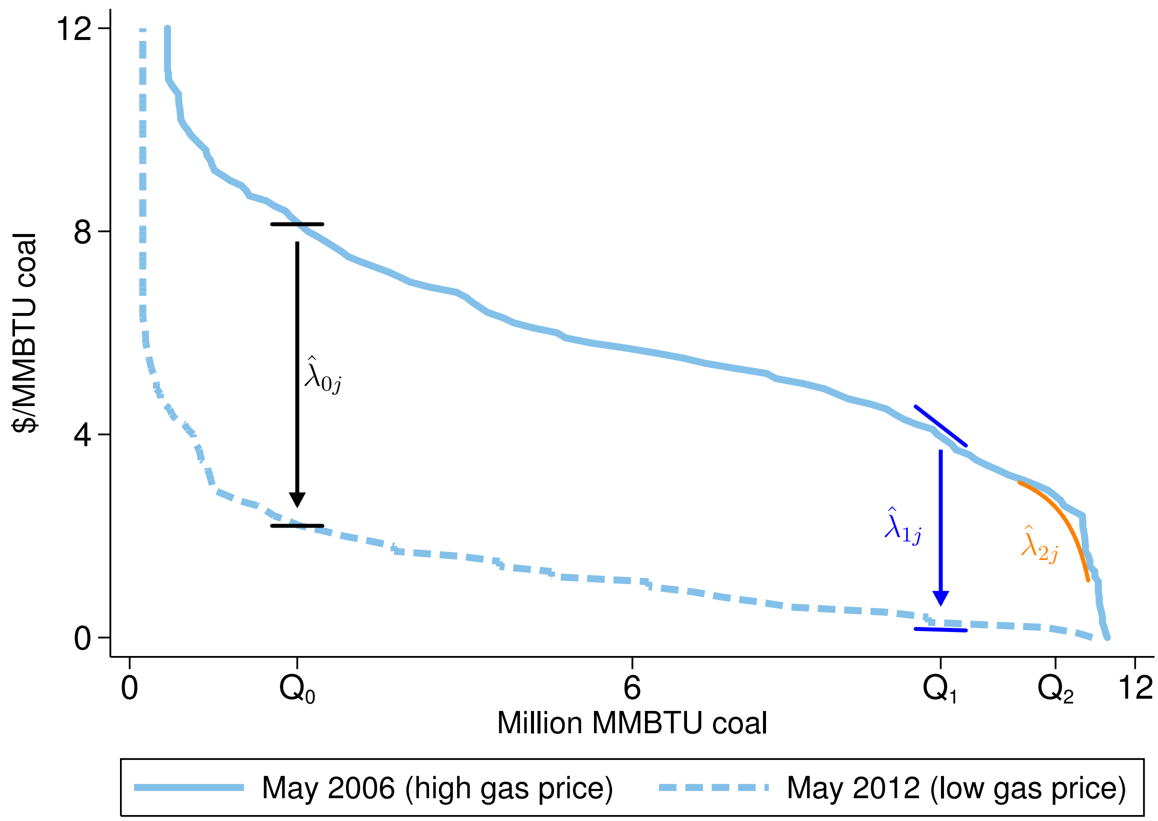
After estimating Equation (1.6) for each unit, I solve each fitted model in terms of the coal price (a component of CR_{ud}), and derive the counterfactual prices for which $\widehat{CF}_{uh} = 0.5$. These are the predicted coal prices at which unit u would have been exactly marginal in electricity supply, in each hour h. I integrate the distribution of counterfactual coal prices across all hours in each month and across each of plant j's units, transforming

[33] A large body of research addresses each of these issues. Fowlie (2010) finds that rate-of-return regulation distorts coal plants' incentives to adopt least-cost pollution abatement strategies, while Cicala (2015) finds that regulated plants do not minimize coal purchase costs. Cicala (2017) demonstrates that the lack of a marginal electricity price signal leads to large allocative inefficiencies under non-market dispatch. Borenstein, Bushnell, and Stoft (2000); and Davis and Hausman (2016) demonstrate how transmission constraints directly impact electricity market outcomes. Mansur (2008) and Reguant (2014) focus on power plants' dynamic production constraints and non-convex operating costs.

[34] I use the term "policy function" because I treat coal demand estimation as a prediction problem, rather than estimating an optimized function of price or quantity.

[35] $CF_{uh} \in [0, 1]$ by construction, and $CF_{uh} = 0$ if unit u does not operate in hour h. G_{uh} sums hourly CEMS generation across all units in u's market region. The cost ratio CR_{ud} divides unit u's marginal cost (including coal price and marginal environmental costs) by the average marginal cost of gas units in the same PCA. Generation bins b allow for a flexible relationship between CF_{uh}, G_{uh}, and CR_{ud}. $\mathbf{G}_{uh}$ includes the daily sum, maximum, minimum, and standard deviation of G_{uh}; the daily maximum temperature; hour-of-day, quarter-of-year, and year-of-sample fixed effects; and year dummies interacted with the daily sum of G_{uh}. Appendix A.4 describes my demand estimation procedure in further detail, and conducts sensitivity analysis on Equation (1.6).

Figure 1.5.5: Coal Demand Estimation Example



Notes: This figure plots two estimated coal demand curves for a representative plant j. The solid curve shows plant j's monthly coal demand for May 2006, when the gas price was high; the dashed curve shows its demand for May 2012, when the gas price was low. I estimate how a change in gas price affects the level ($\hat{\lambda}_{0j}$) and slope ($\hat{\lambda}_{1j}$) of plant j's coal demand. In this simplified example, a decrease in gas price decreases the level of inverse demand (i.e., $\hat{\lambda}_{0j} > 0$ at Q_0) and makes its slope less negative (i.e., $\hat{\lambda}_{1j} < 0$ at Q_1). The third parameter $\hat{\lambda}_{2j}$ estimates the average (local) curvature of coal demand, which is concave (i.e., $\hat{\lambda}_{2j} > 0$) at Q_2. My algorithm estimates all three parameters at plant j's observed coal quantities, using estimated demand curves and gas prices for all months. Appendix A.4.1 describes this algorithm in further detail.

unit-specific capacity factors into their corresponding coal quantities. This yields price-quantity mappings, or coal demand curves for each plant-month. Finally, I use monthly gas price variation to estimate three parameters that map directly to my comparative static in Equation (1.3):

$$(1.7) \qquad \hat{\lambda}_{0j} \sim \frac{\partial P_{oj}}{\partial Z_j} \qquad \hat{\lambda}_{1j} \sim \frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_j \qquad \hat{\lambda}_{2j} \frac{\theta_{oj}}{N_{oj}} \sim E_{D_{oj}}$$

$\hat{\lambda}_{0j}$ and $\hat{\lambda}_{1j}$ estimate how gas price affects the *level* and *slope* of plant j's inverse demand, respectively. $\hat{\lambda}_{2j}$ estimates the average *curvature* of plant j's inverse demand, where $E_{D_{oj}}$ is the elasticity of the slope of inverse demand (as defined in Equation (1.4)). Figure 1.5.5 describes these parameters graphically, using two estimated demand curves from one plant.

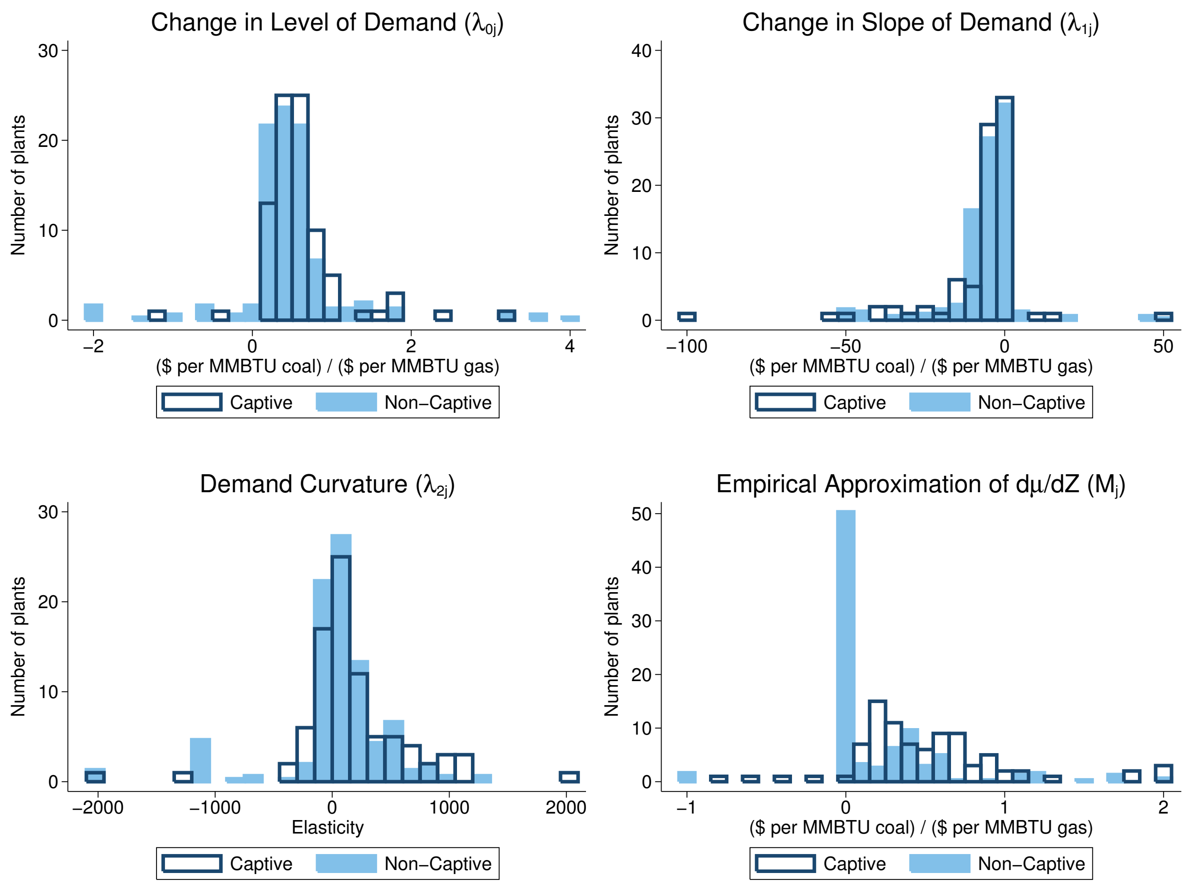
Figure 1.5.6 presents the estimated distributions of these parameters, separately for captive and non-captive plants. The distribution of $\hat{\lambda}_{0j}$ has a median of 0.46, which implies that for a \$1 per MMBTU decrease in gas prices, coal prices would need to fall by \$0.46 per MMBTU to maintain the median plant's baseline coal consumption.[36] These estimates reveal considerable heterogeneity across coal plants, with $\hat{\lambda}_{0j} \in [0, 1]$ for 86 percent of plants, implying reasonable elasticities of substitution between coal and gas. $\hat{\lambda}_{1j} < 0$ for 88 percent of plants, implying that low gas prices have caused most plants' inverse coal demand to become less steep. As gas prices have fallen, most coal plants have become more marginal in electricity markets, which has made them more coal-price-elastic. Finally, $\hat{\lambda}_{2j} > 0$ for 67 percent of plants, suggesting that coal demand tends to be (locally) concave.

Importantly, $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, and $\hat{\lambda}_{2j}$ are the outcome of a linear prediction algorithm that imposes no assumptions on plant j's objective function. I do assume that plants with multiple generating units operate these units independently, that Equation (1.6) is not misspecified, and that counterfactual coal consumption in each hour is either zero or at maximum capacity. Counterfactual coal prices also hold the rest of the electricity market constant, including the coal prices faced by other plants.[37] This means that my demand estimates could not predict the effects of a common shock to coal commodity prices. However, they *can* predict variation in plant j's idiosyncratic opportunity cost of coal — the very type of price changes that occur when a rail carrier reoptimizes plant-specific markups.

[36] Unlike my regressions on coal shipments where the rail carrier's relevant price is in dollars per ton, these demand parameters use dollars per MMBTU, in order to denominate coal in terms of its energy content (i.e. its value to power plants as a fuel input). BTU content varies substantially across coal shipments, with a mean (standard deviation) of 19.7 (3.4) MMBTU/ton in my estimation sample.

[37] In reality, rail carriers may jointly reoptimize markups across multiple coal plants selling into the same electricity market. If plants j markups move in the same direction as other plants' markups, then my estimates for plant j's coal demand may be too large (small) at low (high) coal prices.

Figure 1.5.6: Coal Demand Estimation Results



Notes: These histograms report the distributions of estimated demand parameters ($\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, $\hat{\lambda}_{2j}$ from Equation (1.7)), and the empirical approximation of the comparative static $\frac{d\mu}{dZ}$ from Equation (1.8). Each histogram includes one observation per plant and applies nearest neighbor weights ($k = 3$). Matching criteria: up to k nearest neighbors, with a maximum distance of 200 miles; exact matches on coal rank; and removing non-utility and non-rail plants. The outermost bins are bottom-coded and top-coded, for ease of presentation. These outliers likely reflect idiosyncratic factors (other than the gas price) that have affected coal demand. Each histogram includes 86 captive plants and 96 matched non-captive plants.

1.5.4 Markup Changes

Changes in plant j's markups will depend on both its coal demand and the extent to which it faces market power in coal shipping. Recall the comparative static from Equation (1.3):

$$\frac{d\mu_{oj}}{dZ_j} = \frac{\frac{\partial P_{oj}}{\partial Z_j}(2 + E_{D_{oj}} - N_{oj}) - \frac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j} Q_{oj}\theta_{oj}}{2 + E_{D_{oj}}}$$

Using this theory as a guide, I can construct an empirical approximation for the change in plant j's average markup, μ_j, that should result from a change the gas price, Z_j:

$$\Rightarrow \quad M_j \equiv \frac{\hat{\lambda}_{0j}\left[D_j + \hat{\lambda}_{2j}(1 - W_j)(2 - D_j)^{-1}\right] - \hat{\lambda}_{1j}(1 - W_j)}{2 + \hat{\lambda}_{2j}(1 - W_j)(2 - D_j)^{-1}} \tag{1.8}$$

This variable M_j combines my estimated demand parameters ($\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, and $\hat{\lambda}_{2j}$) with data on plant j's transport market structure. I can translate the captiveness indicator D_j into a binary version of N_{oj}, where $\hat{N}_j = 2 - D_j$: for plants captive to a single carrier, $\hat{N}_j = 1$; for non-captive plants, $\hat{N}_j = 2$.[38] Likewise, the option to receive waterborne deliveries can serve as a (crude) empirical proxy for the conduct parameter $\hat{\theta}_j = 1 - W_j$, where W_j is an indicator of water access. For plants without a water option, $\hat{\theta}_j = 1$, consistent with Cournot oligopoly; for plants with the ability to receive barge deliveries that bypass rail carriers, $\hat{\theta}_j = 0$, consistent with a competitive fringe.

The bottom-right panel in Figure 1.5.6 reports the distributions of M_j, separately for captive and non-captive plants. This underscores two potentially important shortcomings of a difference-in-differences strategy that would split plants based on captiveness alone. First, M_j varies considerably across captive plants, with a median of 0.43 and an interquartile range of $[0.23, 0.70]$. This suggests that a captive plant at the 75th percentile of this distribution should have experienced three times larger changes in markups, compared to a captive plant at the 25th percentile. The binary indicator $D_j = 1$ obscures this key heterogeneity.[39] Second, while the distribution of M_j for non-captive plants has a large mass at 0, M_j is positive for 43 percent of non-captive plants. In fact, for many non-captive plants, M_j is larger than for their captive counterparts, implying that W_j, $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, and $\hat{\lambda}_{2j}$ combine to outweigh $D_j = 0$. This suggests that non-captive plants likely also experienced decreases in markups as gas prices fell, even though these plants face less rail market power.

[38] Given that the rail network is close to a duopoly in both the western and eastern U.S., I assign $\hat{N}_j = 2$ for non-captive plants. Using this formulation, $E_{D_j} \sim \hat{\lambda}_{2j}(1 - W_j)(2 - D_j)^{-1}$.

[39] In the absence of a theory model, one might interact D_j with the gas price to estimate a reduced-form difference-in-differences version of Equation (1.5). Appendix A.5.3 reports results from this model, which yields imprecise point estimates close to zero. This is unsurprising, given that these regressions rely on captiveness alone, while ignoring heterogeneity in coal demand.

Using M_j as cross-sectional variation, I estimate markup changes with a lagged difference-in-differences model:

$$
\begin{aligned}
P_{ojms} &= \tau M_j \cdot Z^{HH}_{m-L} + \sum_{\ell=0}^{L-1} \tau_\ell M_j \cdot \Delta Z^{HH}_{m-\ell} \quad \dots \\
&\quad + \beta_C \mathbf{C}_{ojms} + S(\mathbf{T}_{ojms}\,;\beta_T) + \beta_X \mathbf{X}_{jm} + \eta_{oj} + \delta_m + \varepsilon_{ojms}
\end{aligned}
\tag{1.9}
$$

Z^{HH}_m is the average Henry Hub spot price in month m, and $\Delta Z^{HH}_m = Z^{HH}_m - Z^{HH}_{m-1}$. The coefficient of interest τ captures the cumulative effect of a \$1/MMBTU change in gas price, over $L = 36$ months. Each lagged coefficient τ_ℓ captures the cumulative effect after ℓ months, for a plant with $M_j = 1$ relative to a plant with $M_j = 0$. I accommodate delayed effects of gas prices on markups because most coal deliveries occur on long-term contracts, which may be slow to adjust to changing market conditions.[40] Equation (1.9) includes origin-destination fixed effects η_{oj}, which control for the average markup of all shipments from county o to plant j. As in Equation (1.5), $\mathbf{C}_{ojms}$ and $S(\mathbf{T}_{ojms})$ control for the costs of each coal shipment, while $\mathbf{X}_{jm}$ controls for time-varying plant characteristics. I cluster standard errors by plant, and weight observations by the product of nearest-neighbor weights and shipment size.

I interpret $\hat{\tau}$ as the causal effect of gas price changes on coal-by-rail markups. The key identifying assumption is that gas price changes are uncorrelated with other factors affecting the differential trajectory of coal markups, after controlling for time-varying plant characteristics in $\mathbf{X}_{jm}$. Technological advances of the fracking boom were unrelated to coal mining costs; the Henry Hub spot price is also uncorrelated with U.S. diesel prices, which drive coal shipping costs.[41] A violation of parallel trends would occur if a coal plant unobservable that is correlated with coal prices (e.g., how electricity regulators monitor plants' coal purchase costs) changed differentially for plants with high vs. low M_j.[42]

[40] It is common to allow for delayed pass-through, in settings where price changes may not be instantaneous. Pouliot, Smith, and Stock (2017) use the same differenced lag structure to estimate delayed pass-through in the market for renewable fuel credits. This is algebraically equivalent to the standard (non-differenced) distributed lag model, $\sum_{\ell=0}^{L} \beta_\ell D_j \cdot Z^{HH}_{m-\ell}$, where $\sum_{\ell=0}^{L} \beta_\ell = \tau$. Coal prices for long-term contracts should be quite sticky, even though many coal contracts include flexible price-adjustment provisions that enable rail carriers to partially re-optimize markups before these contracts expire (Joskow (1988); Kosnik and Lange (2011)). I estimate Equation (1.9) separately for contract and spot market shipments.

[41] The fracking boom may have impacted local labor markets in certain coal mining regions, and oil-by-rail shipping increased congestion in certain portions of the rail network. However, my results are largely consistent across geographic regions, making these violations of parallel trends unlikely. During 2002–2015, the correlation between Henry Hub and U.S. average monthly diesel prices was -0.01. If these two price series were correlated, I would worry about multicollinearity between Z^{HH}_m and diesel prices in $\mathbf{T}_{ojms}$.

[42] Christian and Barrett (2017) show that even spurious time trends can induce bias for a difference-in-differences treatment variable that interacts a cross-sectional characteristic with an exogenous time series. In Appendix A.5.4.2, I show that my data exhibit parallel pre-trends in M_j. I also use interacted time fixed effects to help rule out potential time-varying confounders.

1.6 Results

1.6.1 Markup Levels

Table 1.6.2 reports results from estimating Equation (1.5), which demonstrate that captive plants indeed face higher markups than their matched non-captive counterparts. Point estimates of \$2–3 translate to average differential markups of 4–7 percent, on an average delivered price of \$37–40 per ton for non-captive plants. The implies that markups for captive plants contribute 13–24 percent of the spatial gap between mine-mouth prices (averaging \$22–24 per ton) and delivered prices. Given that the indicator D_j applies an arbitrary threshold to define captiveness, and non-captive plants likely face nonzero markups, these estimated *differentials* likely understate the average markup *level* faced by captive plants.

These results are robust to the number of nearest-neighbor matches, which I vary from $k = 1$ to $k = 5$. I construct my estimation sample to exclude the (very few) coal plants constructed after 2001, but many plants retired during my 2002–2015 sample period. If these plants' exit decisions were correlated with their delivered coal prices, which affected their ability to compete with low-cost natural gas, endogenous exit could bias my estimates in an unbalanced panel. Columns (4)–(6) in Table 1.6.2 restrict the sample to plants receiving at least 1 delivery in each year.[43] While this removes 31 percent of plants, point estimates remain statistically significant and increase slightly in magnitude.

Table 1.6.3 interacts rail captiveness with another pre-determined plant characteristic likely to affect markups: an indicator for access to waterborne shipments. These results reveal differential markups of \$2–5 per ton for captive plants with no coal-by-barge option, relative to plants with the most competitive shipping regimes (i.e. the omitted group with multiple rail carriers and barges). While these point estimates are more sensitive to the number of nearest neighbors and to balancing the panel, they show that the markup distortion may be as large as \$4–5 per ton, or 11–14 percent of the average delivered coal price. This implies markups of up to 30–41 above rail carriers' marginal shipping costs. My point estimates are consistent with the magnitudes of Busse and Keohane (2007), who estimate differential markups of \$4 per ton for coal shipped from Wyoming in the late 1990s. My analysis demonstrates that coal-by-rail price discrimination, due to geographic variation in market power, has persisted through recent years and across multiple coal-producing regions.[44]

[43] This is not a fully balanced panel. Coal shipments are lumpy, and many active plants do not report deliveries in each month. I "balance" the panel to mitigate any confounding effects from plant exit, not to take a stand on the timing of coal deliveries. Olley and Pakes (1996) demonstrate that bias due to endogenous exit may remain even after balancing the panel, if exit is correlated with unobserved firm productivity. This is not a concern in my setting, as I directly control for each plant's productivity (i.e. its heat rate).

[44] My estimates are consistent if I restrict the sample by origin or destination region. Appendix A.5.2 reports these results, along with additional sensitivity analyses.

Table 1.6.2: Markup Levels – Captive vs. Non-Captive Coal Plants

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\text{Captive}]_j$	2.190***	1.821***	1.637***	2.783***	2.301***	1.995***
	(0.748)	(0.612)	(0.579)	(0.806)	(0.655)	(0.607)
k nearest neighbors	1	3	5	1	3	5
Balanced panel				Yes	Yes	Yes
Coal county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	40.16	37.67	38.96	39.88	36.85	38.04
Plants	141	183	195	97	127	134
Captive plants	87	87	87	61	61	61
Plant-county-months	33,114	44,797	48,227	27,522	37,168	39,732
Observations	66,552	88,384	94,275	58,192	77,115	81,948

Notes: Each regression estimates Equation (1.5) at the coal shipment level, with delivered coal price ($ per short ton) as the dependent variable. I control for shipping costs using the 4-way interaction of rail distance, diesel price, tons shipped, and rail traffic density. Plant- and delivery-specific controls are listed in panels A and B of Table 1.5.1, respectively. I also control for the average annual coal price from the originating county, each plant's distance to its closest rail terminal, and baseload natural gas capacity in each plant's PCA. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for non-captive plants only. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 1.6.3: Markup Levels – Captiveness and Water Access

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}$[Captive, No Water]$_j$	3.728***	2.393***	2.331***	4.729***	3.102***	2.886***
	$[1.6, 5.9]$	$[0.7, 4.0]$	$[0.8, 3.8]$	$[2.6, 6.9]$	$[1.3, 4.9]$	$[1.3, 4.5]$
$\mathbf{1}$[Captive, Water]$_j$	2.104**	0.832	0.840	2.940***	1.501*	1.387*
	$[0.0, 4.2]$	$[-0.7, 2.4]$	$[-0.6, 2.3]$	$[0.9, 5.0]$	$[-0.2, 3.2]$	$[-0.2, 2.9]$
$\mathbf{1}$[Non-Captive, No Water]$_j$	2.150	0.268	0.555	2.737*	0.659	0.889
	$[-0.5, 4.8]$	$[-1.7, 2.2]$	$[-1.3, 2.4]$	$[-0.2, 5.6]$	$[-1.5, 2.8]$	$[-1.2, 2.9]$
k nearest neighbors	1	3	5	1	3	5
Balanced panel				Yes	Yes	Yes
Coal county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.74	34.56	35.32	33.31	34.25	35.16
Plants	141	183	195	97	127	134
Captive plants	87	87	87	61	61	61
Water-access plants	55	84	88	39	62	64
Plant-county-months	33,114	44,797	48,227	27,522	37,168	39,732
Observations	66,552	88,384	94,275	58,192	77,115	81,948

Notes: Each regression estimates Equation (1.5) at the coal shipment level, with delivered coal price ($ per short ton) as the dependent variable, and interacting rail captiveness with an indicator for water access (i.e. plants having the option to purchase coal via barge). I control for shipping costs using the 4-way interaction of rail distance, diesel price, tons shipped, and rail traffic density. Plant- and delivery-specific controls are listed in panels A and B of Table 1.5.1, respectively. I also control for the average annual coal price from the originating county, each plant's distance to its closest rail terminal, and baseload natural gas capacity in each plant's PCA. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for the omitted group (non-captive plants with a water option). Standard errors are clustered by plant, and 95 percent confidence intervals are in brackets. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

1.6.2 Markup Changes

Having established that markup *levels* vary across coal plants, I now estimate heterogeneous *changes* in markups due to changes in the gas price. I identify markup changes using cross-sectional variation in M_j, a predictor of how railroads reoptimize markups due to variation in both transport market power and plant-specific coal demand. Table 1.6.4 reports cumulative effects across 36 months, from estimating Equation (1.9). I find positive, statistically significant point estimates, which are qualitatively consistent with the predictions of my oligopoly model. As gas prices fell during the fracking boom, rail carriers reoptimized markups *heterogeneously* across plants, causing markups to decrease *more* for plants with greater M_j.

I estimate separate regressions for long-term contracts vs. spot-market shipments, as the timing of markup changes will likely differ by transaction type. Rail carriers should be able to reoptimize spot markups more quickly than contract markups; however, markup changes for relatively less flexible contracts should be more persistent, yielding larger cumulative effects. Figure 1.6.7 plots lagged coefficients $\hat{\tau}_\ell$ for each regression in Table 1.6.4, where each coefficient represents the cumulative effect through ℓ months. This reveals that contract markups begin to adjust 6 months after a gas price shock, with delayed effects that accumulate until month 18 and persist through month 36. By contrast, the effects for spot market shipments attenuate and lose significance after 36 months. This likely reflects a difference in transaction costs, as bilateral contract negotiations facilitate greater opportunity for price discrimination than posted spot shipping rates.[45] Table 1.6.5 reveals that differential changes in markups were more pronounced for bituminous coal and for plants in the South and East, even though my results remain statistically significant across coal ranks and plant regions.

Table 1.6.4 implies that for a \$1/MMBTU decrease in gas price, markups decreased by \$1.19–1.57/ton for a plant with $M_j = 1$, compared to plants with $M_j = 0$. However, a literal interpretation of $M_j = 1$ would suggest an effect size of \$1/MMBTU of coal, equivalent to roughly \$20/ton of coal.[46] This mismatch in magnitudes underscores the shortcomings of my oligopoly model, which does not account for railroad regulation or the threat of arbitrage. Suppose that binding regulation limited the maximum markup to \$5/ton, but unconstrained markups would have been \$14/ton before the fracking boom and \$4/ton after the fracking boom. In this case, I would only observe a \$1/ton decrease in markups, rather than the \$10/ton decrease predicted by an unconstrained model. This explains why M_j does not generate accurate *quantitative* predictions of how markups change.[47] However, my results in Table 1.6.4 demonstrate that M_j can generate accu-

[45] Railroads likely invest more resources in reoptimizing less flexible, longer-lived contracts. Whereas pooled and contract results are robust to alternative nearest-neighbor weights, fixed effects, and controls, spot market results are not. Appendix A.5.4 reports these sensitivities, and additional robustness checks.

[46] Recall that M_j is in units of \$/MMBTU of coal, but the dependent variable in Tables 1.6.4–1.6.6 is the coal price in \$/ton of coal. BTU content varies across coal shipments, with an average of 19.7 MMBTU/ton.

[47] My model also assumes that markups are independent across plants. If rail carriers jointly optimize markups across multiple markets, this would attenuate my estimates. M_j linearly extrapolates to a finite

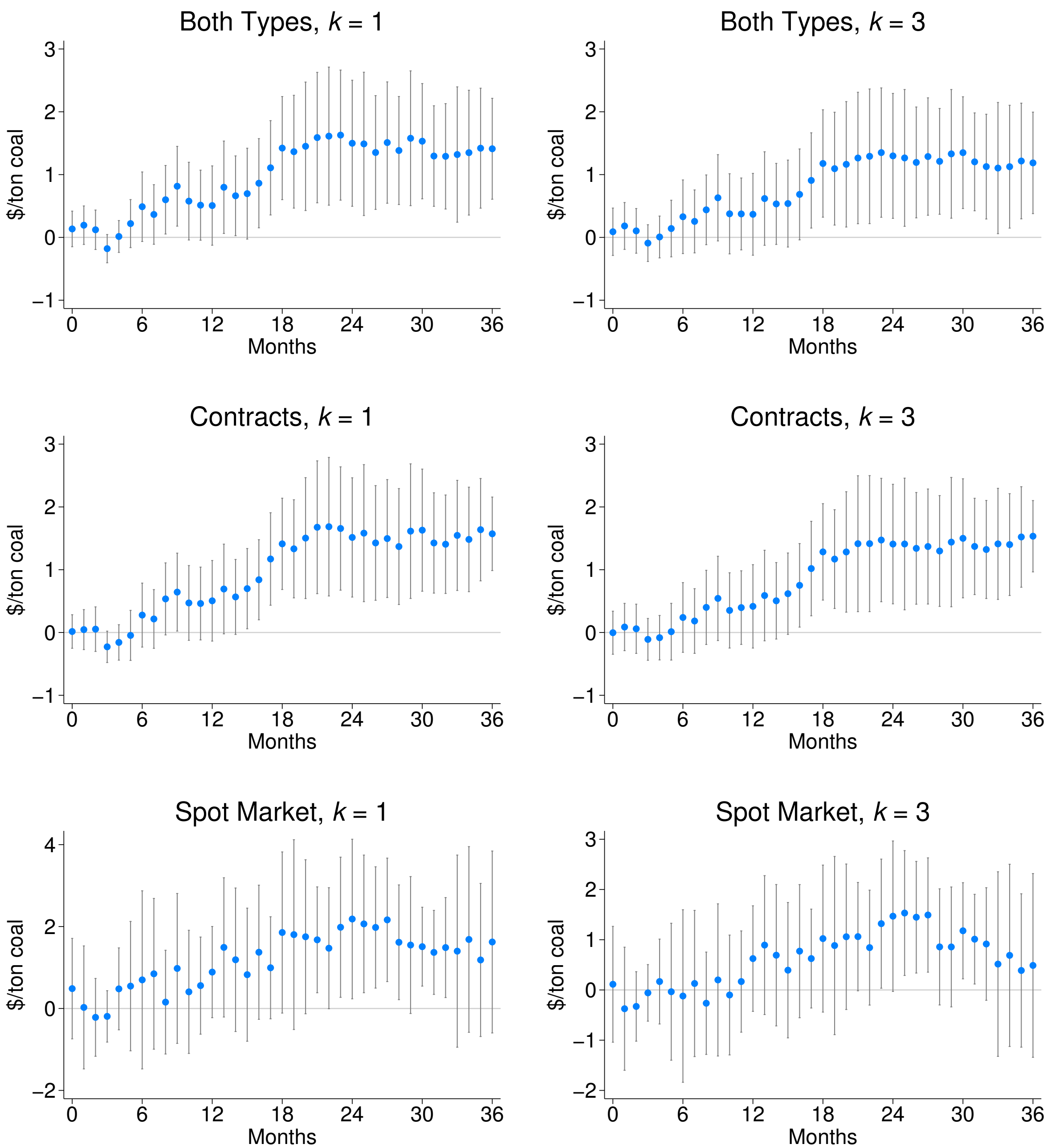
Table 1.6.4: Markup Changes – Demand Parameters Interacted with Gas Price

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\widehat{\Delta\text{Markup}})_j \times (\text{Gas Price})_m$	1.411***	1.187***	1.572***	1.533***	1.624	0.488
	(0.405)	(0.408)	(0.295)	(0.287)	(1.117)	(0.923)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant $\times$ county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.27	34.20	32.07	33.04	41.76	41.39
Plants	94	124	92	122	85	115
Plant-county-months	26,060	35,651	22,000	29,806	6,796	9,630
Observations	56,219	75,089	44,651	59,178	11,487	15,797

Notes: Each regression estimates Equation (1.9) at the coal shipment level, with delivered coal price ($ per ton) as the dependent variable. The first 2 columns pool long-term contracts and spot market shipments, while the middle and right columns split the sample by transaction type. The DD treatment variable interacts the empirical approximation M_j of the comparative static $\frac{d\mu}{dZ}$ (from Equation (1.8)) with the Henry Hub average monthly spot price for natural gas, using a lagged-differenced model with $L = 36$ lags. This table reports estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure 1.6.7 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. M_j is in units of $ per MMBTU of coal, and BTU content ranges from 14–30 MMBTU/ton. I control for shipping costs using the 4-way interaction of rail distance, diesel price, tons shipped, and rail traffic density. Plant- and delivery-specific controls are listed in panels A and B of Table 1.5.1, respectively. I also control for the average annual coal price from the originating county, and baseload natural gas capacity in each plant's PCA. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

‘

Figure 1.6.7: Markup Changes – Cumulative Effects



Notes: This figure plots 36 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{35}$) and $\hat{\tau}$, from 6 separate regressions of Equation (1.9) with $L = 36$ lags. Each panel corresponds to a column in Table 1.6.4, which reports $\hat{\tau}$ only (i.e. the rightmost point in each graph). Each coefficient estimates the interaction of M_j with the ℓ-month lagged difference in natural gas prices ($\Delta Z^{HH}_{m-\ell}$), such that each dot represents the cumulative effect through ℓ months. Whiskers denote 95 percent confidence intervals for each point estimate, with standard errors clustered by plant. See the notes below Table 1.6.4 for further details on the estimation.

Table 1.6.5: Markup Changes – Contract Shipments, Split Samples

	Split by Coal Grade		Removing Plants in Region		
	Bituminous	Sub-bituminous	West	Midwest	South/ East
	(1)	(2)	(3)	(4)	(5)
$(\Delta\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	2.115***	0.921**	1.461***	1.775***	1.059***
	(0.305)	(0.412)	(0.271)	(0.242)	(0.338)
k nearest neighbors	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes
Mean of dep var	48.89	22.02	40.49	30.40	28.98
Plants	76	69	85	65	94
Plant-county-months	18,693	10,910	23,182	17,640	18,872
Observations	31,595	27,215	44,501	35,329	38,614

Notes: Each regression estimates Equation (1.9) at the coal shipment level for contract shipments only, with delivered coal price ($ per ton) as the dependent variable. Columns (1)–(2) include only shipments of a given coal grade. I partition all plants into three regions, and Columns (3)–(5) each exclude plants from a given region. The DD treatment variable interacts the empirical approximation M_j of the comparative static $\frac{d\mu}{dZ}$ (from Equation (1.8)) with the Henry Hub average monthly spot price for natural gas, using a lagged-differenced model with $L = 36$ lags. This table reports estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. M_j is in units of $ per MMBTU of coal, and BTU content ranges from 14–30 MMBTU/ton. I control for shipping costs using the 4-way interaction of rail distance, diesel price, tons shipped, and rail traffic density. Delivery-specific controls are listed in panel B of Table 1.5.1. I omit plant-specific controls (except for ISO and scrubber indicators), in order to reduce the number of covariates to less than the number of plant clusters (pooled results in Table 1.6.4 are not sensitive to inclusion of these controls). I also control for the average annual coal price from the originating county, and baseload natural gas capacity in each plant's PCA. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

rate *qualitative* predictions, by capturing cross-sectional heterogeneity in both transport market structure and coal demand.

To better capture how magnitudes vary across the full range of coal plants, I discretize M_j into five indicator variables corresponding to the quintiles of its positive support.[48] Table 1.6.6 reports results for quintiles 2–5, revealing magnitudes that increase monotonically in M_j. The omitted category is the 41 percent of plants with $M_j \leq 0.22$, most of which are non-captive and located on navigable waterways. If I assume no markup changes for these omitted plants, the point estimates in Table 1.6.6 represent the average change in markups for plants in each quintile.[49] I find statistically significant decreases in contract markups, for the 43 percent of plants in quintiles 3–5 (i.e., $M_j > 0.35$). For the 14 percent of plants in quintile 5 (i.e., $M_j > 0.70$), a \$1/MMBTU gas price decrease caused average markups to fall by \$1.05–1.33/ton, and caused contract markups to fall by \$1.34–1.49/ton. Given that gas prices fell by \$4/MMBTU during the fracking boom, and that average markup *levels* were \$2–5/ton, these magnitudes imply that rail carriers have heterogeneously reoptimized markups, to eliminate most of the market power distortion for a subset of plants.

My results demonstrate that market power exists in coal transportation, and that rail carriers strategically reoptimize coal markups in a manner consistent with profit maximization. Rail market power arises primarily from coal's geographic specificity, as the production and consumption of coal are both highly locationally constrained. This affords rail intermediaries substantial bargaining power, and coal's low value-to-weight ratio increases the premium on transportation access. To identify market power, I exploit price dispersion due to the lack of spatial arbitrage in coal deliveries. While this feature is likely unique to coal markets, the features that generate market power in coal shipping—geographic specificity and high freight costs—exist in many other commodity markets (e.g., Covert and Kellogg (2017) on crude oil; Hortaçsu and Syverson (2007) on cement).

change in gas price, which may overstate markup changes if markups approach their lower bound of zero. Measurement error in M_j may also attenuate my estimates of $\hat{\tau}$, and I address this issue in Appendix A.5.4.3.

[48]Each "quintile" includes 14–16 percent of plants, because $M_j \leq 0$ for 28 percent of plants. Tables 1.6.4–1.6.6 omit 3 plants with extremely low/high M_j (i.e. $|M_j| > 2$), which almost certainly reflect errors in estimating these plants' demand parameters. Appendix A.5.4.3 reports results including these outliers, and my point estimates attenuate slightly but largely retain statistically significance.

[49]Non-captive plants with a water delivery option benefit from the most competitive shipping regimes. These plants likely faced markups close to zero, prior to the fracking boom. Hence, if gas price changes caused any markup decreases for omitted plants, such changes should have been relatively small.

Table 1.6.6: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{Gas Price})_m$	0.041	−0.026	0.243	0.274	−0.763	−1.269**
	(0.206)	(0.180)	(0.209)	(0.175)	(0.488)	(0.627)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{Gas Price})_m$	0.275	0.142	0.506**	0.476**	−0.231	−0.940
	(0.229)	(0.201)	(0.236)	(0.209)	(0.455)	(0.578)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{Gas Price})_m$	0.723***	0.561**	0.743***	0.684***	1.030	0.294
	(0.271)	(0.241)	(0.246)	(0.209)	(0.825)	(0.937)
$\mathbf{1}\Big[M_j \in (0.70, 2.00]\Big] \times (\text{Gas Price})_m$	1.334***	1.050***	1.492***	1.341***	1.367	0.098
	(0.493)	(0.376)	(0.466)	(0.367)	(1.123)	(0.981)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	31.46	35.03	30.84	34.88	33.55	35.54
Plants	94	124	92	122	85	115
Plant-county-months	26,060	35,651	22,000	29,806	6,796	9,630
Observations	56,219	75,089	44,651	59,178	11,487	15,797

Notes: Each regression estimates a modified Equation (1.9) at the coal shipment level, with delivered coal price ($ per ton) as the dependent variable, and $L = 36$ lags. Instead of interacting the L-month lagged gas price with a continuous M_j to estimate the coefficient of interest $\hat{\tau}$, I estimate four $\hat{\tau}$'s using indicator variables for quintiles of M_j's positive support. The omitted group is the first quintile, plus all plants with $M_j \leq 0$. This table reports the average cumulative change in markups caused by a \$1/MMBTU change in gas price, for plants in a given quintile relative to the omitted group. Lag-differenced coefficients $\hat{\tau}_\ell$ still use a continuous M_j interaction. The first 2 columns include long-term contracts and spot market shipments, while the middle and right columns split the sample by transaction type. I control for shipping costs using the 4-way interaction of rail distance, diesel price, tons shipped, and rail traffic density. Plant- and delivery-specific controls are listed in panels A and B of Table 1.5.1, respectively. I also control for the average annual coal price from the originating county, and baseload natural gas capacity in each plant's PCA. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

1.7 Implications for Climate Policy

1.7.1 Markup Size vs. External Costs of Coal

Given that coal intermediaries exercise market power, a carbon tax has the potential to restrict aggregate coal consumption below the social optimum. Buchanan (1969) demonstrates that if a market power distortion is sufficiently large relative to external costs, a Pigouvian tax could actually reduce welfare. Based on my estimates in Table 1.6.3, I can reject differential coal markups greater than \$7 per ton, relative to the "most-competitive" omitted category (i.e. plants with multiple rail carriers and a coal-by-barge option). \$7 per short ton of coal is roughly equivalent to \$2–5 per metric ton of CO_2, which is far below recent social cost of carbon estimates of \$50 per metric ton (Interagency Working Group (2016); Revesz et al. (2017)). Hence, coal markups are an order of magnitude smaller than the carbon externality.

This means that the welfare gains from Pigouvian taxation would likely dwarf any welfare loss from exacerbating the market power distortion (echoing Oates and Strassmann (1984)).[50] However, real-world carbon prices typically range from \$3–30 per metric ton of CO_2, which is far below estimated climate damages of \$50 per metric ton (Carl and Fedor (2016)). Under such a suboptimally low carbon tax, the presence of coal markups should increase welfare by internalizing an additional fraction of marginal damages. Even under a tax equal to marginal climate damages, markups could potentially *still* increase welfare by internalizing local air pollution damages from coal combustion (Levy, Baxter, and Schwartz (2009); Muller, Mendelsohn, and Nordhaus (2011)).

1.7.2 Pass-Through of Implicit Carbon Tax

A negative gas price shock makes coal plants less competitive in electricity supply, and a tax on CO_2 emissions similarly disadvantages coal, the more carbon-intensive fuel. Cullen and Mansur (2017) argue that under reasonable assumptions, the coal-to-gas price ratio is a sufficient statistic for CO_2 emissions from the electricity sector under a counterfactual carbon tax. If electricity demand is perfectly inelastic, and only coal or gas generators can be marginal in electricity supply, then a short-run change in relative fuel prices should yield the same emissions outcomes as the equivalent carbon tax.[51]

Using Cullen and Mansur's framework, a gas price change ΔZ yields the same fuel cost ratio (CR) as the carbon tax t (suppressing plant j subscripts):

$$CR = \frac{MC_{coal}}{MC_{gas}} = \frac{P}{Z + \Delta Z} = \frac{P + t\,E_{coal}}{Z + t\,E_{gas}} \tag{1.10}$$

[50] By contrast, the distortion above marginal cost pricing is large relative to pollution externalities in U.S. retail natural gas (Davis and Muehlegger (2010)), and cement markets (Fowlie, Reguant, and Ryan (2016)).

[51] Appendix A.1.2 provides further detail on the assumptions underlying this section, along with derivations of implied pass-through rates.

where P is the coal price paid by power plants, MC_{fuel} are marginal costs per MMBTU, and E_{fuel} are fuel-specific CO_2 emissions factors (in metric tons CO_2/MMBTU). My empirical results demonstrate that P is not fixed, and I can rewrite this expression to allow coal markups (μ) to endogenously respond to ΔZ:

$$CR = \frac{MC_{coal}}{MC_{gas}} = \frac{P + \Delta\mu}{Z + \Delta Z} = \frac{P + \rho\, t\, E_{coal}}{Z + t\, E_{gas}} \tag{1.11}$$

ρ is the pass-through rate of the implicit tax t. If markups do not change ($\Delta\mu = 0$), the cost ratio reflects full pass-through of the carbon tax, or $\rho = 1$ as in Equation (1.10). If markups decrease in response to a negative gas price shock ($\Delta Z < 0$ causing $\Delta\mu < 0$, consistent with Table 1.6.6), then pass-through of t is incomplete, and $\rho < 1$. By reoptimizing markups during the fracking boom, rail carriers effectively lowered the coal-to-gas cost ratio, which led to incomplete pass-through of the negative shock to coal demand.

I can rearrange Equation (1.11) to translate my point estimates from Table 1.6.6 into implied tax pass-through rates, setting $\Delta Z = 1$ and using average fuel prices from the start of the fracking boom. Table 1.7.7 reports pass-through rates for the five quintiles of M_j, assuming full pass-through ($\rho_j = 1$) for the omitted group in Table 1.6.6. Panel A reveals substantial heterogeneity both across and within plant groups. While most plants experience full pass-through, the 14 percent of plants with the highest M_j have pass-through ranging from $\rho_j = 0.42$ to $\rho_j = 0.90$, with an average rate of $\rho_j = 0.81$. Isolating long-term contracts implies even lower pass-through rates, due to larger changes in markups for contract shipments. Panel B calculates pass-through rates for a cost ratio inclusive of marginal environmental costs, to account for marginal costs of pollution abatement already incurred by coal and gas plants.[52] While pre-existing environmental policies reduce the size of the implicit carbon tax, implied pass-through rates increase only slightly. To my knowledge, this is the first empirical evidence that predicts heterogeneous and incomplete pass-through of a carbon tax in either U.S. coal markets or the U.S. electricity sector.[53]

My results contribute to a growing body of research finding heterogeneous pass-through of price-based climate policies. Previous work has shown that variation in market structure either across industries (Ganapati, Shapiro, and Walker (2016)), or across space within an industry (Pouliot, Smith, and Stock (2017)), can generate substantial heterogeneity in pass-through rates.[54] Similarly, I find that heterogeneous pass-through

[52]For plants covered by SO_2, NO_x, or CO_2 allowance trading regimes, I monetize each generating unit's emissions rates using prevailing allowance prices. I also include variable costs of operating pollution control devices, such as scrubbers.

[53]Chu, Holladay, and LaRiviere (forthcoming) estimate incomplete pass-through from coal spot prices to delivered coal prices; the authors caution that their analysis is not predictive of long-term price changes that would occur under a carbon tax. Kim, Chattopadhyay, and Park (2010) conceptually illustrate how variation in power plants' costs may lead to incomplete carbon tax pass-through.

[54]Ganapati, Shapiro, and Walker (2016) find heterogeneous energy cost pass-through for manufacturing industries under imperfect competition. Pouliot, Smith, and Stock (2017) find lower pass-through

Table 1.7.7: Heterogeneous Pass-Through of Implied Carbon Tax

	Plant Group (Quintile of M_j)				
	(1)	(2)	(3)	(4)	(5)
$\Delta\,\widehat{\text{Markup}}$ (M_j)	$(-2.00, 0.22]$	$(0.22, 0.35]$	$(0.35, 0.52]$	$(0.52, 0.70]$	$(0.70, 2.00]$
Share of plants	0.41	0.16	0.14	0.15	0.14
A. Fuel prices only					
ρ_j, all shipments	1.00	**1.00**	**0.97**	**0.90**	**0.81**
		[1.00, 1.01]	[0.91, 0.98]	[0.82, 0.94]	[0.42, 0.90]
		(0.90, 1.11)	*(0.88, 1.06)*	*(0.82, 0.99)*	*(0.68, 0.94)*
ρ_j, contracts only	1.00	**0.93**	**0.89**	**0.87**	**0.75**
		[0.83, 0.96]	[0.71, 0.94]	[0.77, 0.92]	[0.27, 0.87]
		(0.82, 1.03)	*(0.80, 0.98)*	*(0.80, 0.95)*	*(0.62, 0.88)*
B. Fuel + environmental costs					
ρ_j, all shipments	1.00	**1.00**	**0.97**	**0.91**	**0.83**
		[1.00, 1.01]	[0.93, 0.99]	[0.84, 0.96]	[0.51, 0.91]
		(0.91, 1.10)	*(0.89, 1.05)*	*(0.84, 0.99)*	*(0.72, 0.94)*
ρ_j, contracts only	1.00	**0.93**	**0.91**	**0.89**	**0.78**
		[0.85, 0.96]	[0.75, 0.95]	[0.79, 0.95]	[0.39, 0.88]
		(0.84, 1.02)	*(0.82, 0.99)*	*(0.83, 0.96)*	*(0.67, 0.89)*

Notes: This table converts point estimates from Table 1.6.6 into pass-through rates of an implied carbon tax, for $k = 3$ nearest neighbors. Average pass-through rates are in bold, and square brackets report the minimum and maximum pass-through rates for plants in each group. I report the 95 percent confidence intervals for the *average* (bolded) pass-through rates in parentheses and italics (calculated from the confidence interval of each $\hat{\tau}$ estimate). I rearrange Equation (1.11) to solve for ρ as a function $\Delta\mu$; assign P and Z their average prices from 2007–08, the beginning of the fracking boom; and set $E_{coal} = 0.095$ and $E_{gas} = 0.053$, their average emissions rates in metric tons CO_2 per MMBTU. For a \$1/MMBTU change in gas price (i.e. $\Delta Z = 1$) and assuming full pass-through ($\rho_j = 1$) in the omitted group, Table 1.6.6 estimates the average change in markups for each group (i.e. $\Delta\mu$, after converting from \$/ton to \$/MMBTU of coal). Whereas Panel A follows Equation (1.11) by only including fuel prices, Panel B includes environmental costs following Equation (A.21) in Appendix A.1.2.

of a carbon tax in U.S. coal markets would arise largely from spatial variation in the competitiveness of coal shipping. However, coal markups also adjust heterogeneously to plant-specific demand shocks; I am able to detect incomplete pass-through only after accounting for this second dimension of heterogeneity.

Muehlegger and Sweeney (2017) estimate incomplete pass-through of firm-specific cost shocks in petroleum refining, but full pass-through of cost shocks that are common across firms. Given that CO_2 emissions rates vary substantially across refineries, this implies that a carbon tax would likely lead to heterogeneous pass-through by inducing variation in refinery-specific costs. Coal power plants exhibit similar variation in CO_2 emissions rates, and I likewise find incomplete pass-through resulting from plant-specific shocks. My analysis is the first to show that pass-through of a carbon tax in the U.S. electricity sector may be heterogeneous and incomplete, in part due to variation in plants' sensitivity to relative cost shocks. By contrast, Fabra and Reguant (2014) estimate full pass-through of carbon prices in the Spanish wholesale electricity market, which they attribute to highly correlated emissions cost shocks among marginal plants. My results demonstrate that variation in upstream market power may weaken the correlation in cost shocks across plants, potentially leading to heterogeneous pass-through despite an *average* pass-through rate close to 1.[55]

1.7.3 Heterogeneous Tax Incidence

Weyl and Fabinger (2013) show how pass-through under imperfect competition is closely linked to economic incidence. In fact, the pass-through rate (ρ), conduct parameter (θ), and number of symmetric firms (N) are sufficient to characterize the incidence (I) of a tax (t):

$$I = \frac{dCS/dt}{dPS/dt} = \frac{\rho}{1 - (1 - \theta/N)\,\rho} \tag{1.12}$$

where CS and PS are consumer and producer surplus. Lower pass-through rates imply that consumers (i.e. coal plants) bear relatively less of the tax burden than producers (i.e. rail carriers). For a given pass-through rate ρ, a less competitive market structure (i.e. greater θ/N) implies that rail oligopolists bear a relatively greater tax burden.

rates of renewable fuel credits in less integrated market regions (see also Knittel, Meiselman, and Stock (2017); Li and Stock (2017)). Spatial and temporal variation in production capacity can also lead to heterogeneous pass-through in petroleum refining (Marion and Muehlegger (2011)); however incomplete pass-through caused by capacity constraints does not necessarily reflect market power (Borenstein and Kellogg (2014)).

[55]Fabra and Reguant (2014) also attribute their finding of full pass-through to inelastic aggregate electricity demand and high-frequency uniform-price auctions. Nazifi (2016) similarly predicts full pass-through of a carbon tax in the Australian electricity market. Stolper (2016) shows how the jurisdictional borders of energy taxes can also induce variation in firm-specific costs, resulting in heterogeneous pass-through.

Given the range of pass-through estimates in Table 1.7.7, the incidence of a carbon tax would likely vary substantially across coal plants. During the fracking boom, plants in the least competitive shipping regimes that were most sensitive to gas prices paid only 45 percent of the burden of low gas prices (i.e., $\rho_j = 0.80$, $\theta_j = 1$, $N_j = 1$). By contrast, plants with full pass-through and a water delivery option paid 100 percent of the lost surplus in coal shipping (i.e., $\rho_j = 1$, $\theta_j = 0$).[56] The complete incidence calculation would also include lost profits in electricity markets, which would depend in part on plants' ability to pass marginal emissions costs through to wholesale electricity prices (Fabra and Reguant (2014)).

My results add to a nascent body of evidence that the assumption of homogeneous incidence can obscure the true distributional impacts of energy taxes. Stolper (2017) uncovers heterogeneous tax incidence for Spanish transportation fuels, which renders a seemingly regressive tax unambiguously progressive. Similarly, Ganapati, Shapiro, and Walker (2016) show that a carbon tax appears less regressive after accounting for variation in the competitiveness of intermediate product markets. In my setting, heterogeneous incidence suggests that under a carbon tax, certain coal plants would stand to lose relatively less than others.[57] By shifting a share of the tax burden further upstream from electricity consumers, market imperfections in coal shipping may also reduce the regressivity of a carbon tax.

1.7.4 Counterfactuals

Figure 1.2.2 illustrates how U.S. electricity generation has shifted away from coal as gas prices have fallen, and several previous studies have sought to quantify the environmental benefits of coal-to-gas switching induced by the fracking boom. For example, Knittel, Metaxoglou, and Trindade (2015) estimate that the 70-percent drop in gas prices between 2008 and 2012 caused CO_2 emissions from electricity generation to fall by up to 19–33 percent.[58] My analysis is the first to show that coal markups have adjusted to partially offset this change in relative fuel prices. This suggests that if coal markups had not changed, the fracking boom could have yielded even greater reductions in CO_2 emissions from electricity generation.

To quantify how decreasing coal markups may have limited coal-to-gas switching in the short run, I first estimate a time series regression similar to Equation (1.6) for each

[56]The share of the burden borne by plant j is $\frac{I_j}{1+I_j} = \frac{\rho_j}{1+(\theta_j/N_j)\rho_j}$. Appendix A.1.3 contains a more detailed discussion of implied carbon tax incidence as it pertains to my theoretical framework.

[57]All coal plants would likely see profits decrease under a carbon tax, yet some plants would likely bear *relatively* less tax burden in the short run. Muehlegger and Sweeney (2017) find that a carbon tax on petroleum refiners would imply heterogeneous firm-specific cost shocks, also creating relative winners and losers.

[58]Holladay and LaRiviere (2017) estimate short-run changes in the marginal CO_2 emissions rates that vary substantially across electricity market regions. Fell and Kaffine (forthcoming) attribute the decline in coal generation to a combination of low natural gas prices and increased wind generating capacity. Wolak (2016) applies a general equilibrium framework to estimate the fracking boom's impact on global coal markets, and finds that increased U.S. coal exports have not led to increases in global CO_2 emissions.

coal generating unit. This calibrates a semi-parametric relationship between each unit's electricity generation and the coal-to-gas cost ratio. Next, I use this fitted model to infer predicted generation under two counterfactual cost ratios: (1) if the fracking boom never happened and gas prices had remained high; and (2) if the fracking boom did happen but coal markups had remained fixed. Converting predicted generation into predicted CO_2 emissions and summing across all coal units, I can calculate short-run CO_2 abatement from the fracking boom both with and without changes to coal markups.[59]

This exercise suggests that decreases in coal markups eroded roughly 8 percent of the fracking boom's short-run abatement potential. Based on these calculations, CO_2 emissions fell by 4.5 percent during the fracking boom, as a result of short-run coal-to-gas substitution alone. However, if coal markups had not changed, this would have been a 4.9-percent emissions reduction. These numbers capture only short-run changes on the intensive margin of fossil generation, and several other margins have contributed to the 20–25 percent decrease in CO_2 emissions from electricity.[60] Even so, they suggest that falling coal markups meaningfully reduced the environmental benefits of low natural gas prices, with unrealized CO_2 abatement equal to \$2.3 billion in climate damages.[61]

Extrapolating to future climate policy, decreases in coal markups may similarly erode the environmental benefits of a carbon tax. However, this countervailing effect would likely disappear if the tax were sufficiently large, as markups should not decrease below zero. This suggests that existing retrospective analyses may underestimate CO_2 abatement under future climate policy. By not accounting for incomplete pass-through in coal markets, these studies likely understate the amount of coal displacement that would occur if a sufficiently stringent climate policy pushed delivered coal prices down to marginal cost.

[59]This short-run exercise abstracts from changes to electric generating capacity. I assume that electricity demand is perfectly inelastic, with gas generation crowding out coal generation 1-for-1. Following Cullen and Mansur (2017), I include a cubic spline in the average cost ratio across all fossil generators in unit u's PCA. Unlike in Equation (1.6), I use the *average* cost ratio because I now want to allow unit u's generation to respond to changes in *other* units' coal prices. Appendix A.6 discusses these counterfactuals in further detail.

[60]Many coal plants have invested in medium-run efficiency improvements (Linn, Mastrangelo, and Burtraw (2014)). On the capacity margin, the fracking boom has spurred investment in new gas plants (Brehm (2017)), while accelerating coal plant retirements (Meng (2016)). Increases in non-fossil generation have also crowded out conventional fossil generation. This counterfactual exercise also ignores the long-run dynamic effects of relative fuel prices changes (Cullen and Reynolds (2016)).

[61]I monetize the difference between 4.5 and 4.9 percent abatement at \$50 per metric ton CO_2. Accounting for medium- and long-run margins would increase the value of unrealized abatement. Importantly, these calculations only consider the electricity sector. Low gas prices have increased CO_2 emissions from other end uses (e.g. residential space heating) and methane leaks from gas drilling. These factors may have combined to outweigh fracking-induced CO_2 abatement from electricity generation (Hausman and Kellogg (2015)).

1.8 Conclusion

This paper demonstrates that decreases in natural gas prices have caused decreases in coal markups. These effects vary substantially across coal-fired power plants, due to the interaction of heterogeneous transportation market structure and plant-specific shocks to coal demand. While previous studies have documented market power in coal shipping, my analysis is the first to show that oligopolist rail carriers reoptimize markups due to heterogeneous changes in plants' coal demand. I also show that pass-through of a carbon tax in the electricity sector may be heterogeneous and incomplete, as railroads would likely reduce markups to effectively buffer a subset of coal plants against the tax. This has the potential to significantly erode the environmental benefits of climate policy, and incomplete pass-through would likely reduce welfare under a carbon tax smaller than marginal climate damages.

My analysis highlights the need for further research estimating pass-through of environmental taxes under imperfect competition. Markets for energy or energy-intensive products tend to be highly concentrated, and the assumption of perfect competition can generate both misguided welfare estimates and biased policy counterfactuals. In order to more fully characterize welfare under climate policy, future research should incorporate market imperfections in both upstream fuel markets (Gillingham et al. (2016)) and downstream electricity markets (Bushnell, Mansur, and Saravia (2008)). My analysis also underscores how heterogeneous market imperfections can generate heterogeneous pass-through of environmental taxes. If pass-through varies across polluting firms, then a uniform carbon price may not incentivize an efficient allocation of CO_2 abatement (Montgomery (1972)), and the optimal second-best climate policy may feature a non-uniform carbon tax.

Future research should also investigate how coal-by-rail market power impacts climate policy outcomes in the medium-to-long run. For example, a carbon tax may incentivize investment in new coal transportation infrastructure, which would mitigate market power and reduce dispersion in delivered coal prices. My analysis largely ignores the coal mining sector, and it is important to consider how a carbon tax might disproportionately hurt labor markets in coal communities (Lobao et al. (2016)). Finally, similar market imperfections in coal transportation likely exist outside the U.S., due to coal's geographic specificity and high transportation costs. Hence, market power in coal shipping may impact climate policy outcomes in the developing world, where coal consumption continues to rise (Wolfram, Shelef, and Gertler (2012)).

Chapter 2

Out of the Darkness and Into the Light? Development Effects of Rural Electrification[1]

2.1 Introduction

Approximately 1.1 billion people around the world still lack access to electricity. These people are overwhelmingly rural, and live almost exclusively in Sub-Saharan Africa and Asia. In recent years, developing countries have made large investments to extend the electricity grid to the rural poor. The International Energy Agency estimates that approximately \$9 billion was spent on electrification in 2009, which it expects to rise to \$14 billion per year by 2030 (International Energy Agency (2011)). This is not surprising, given that electrification is widely touted as an essential tool to help alleviate poverty and spur economic progress; universal energy access is one of the UN's Sustainable Development Goals (UNDP (2015), World Bank (2015)). While access to electricity is highly correlated with GDP at the national level, there exists limited evidence on the causal effects of electricity access on rural economies.

[1]This chapter is coauthored with Fiona Burlig. The original version can be found online at `https://ei.haas.berkeley.edu/research/papers/WP268.pdf`. We thank Michael Anderson, Maximilian Auffhammer, Jie Bai, Kendon Bell, Susanna Berkouwer, Joshua Blonz, Fenella Carpena, Steve Cicala, Lucas Davis, Meredith Fowlie, James Gillan, Michael Greenstone, Solomon Hsiang, Kelsey Jack, Katrina Jessoe, Erin Kelley, Aprajit Mahajan, Shaun McRae, Edward Miguel, Brian Min, Paul Novosad, Nicholas Ryan, Elisabeth Sadoulet, Jacob Shapiro, Andrew Stevens, Adam Storeygard, Matt Woerman, Catherine Wolfram, and seminar participants at University of California, Berkeley, University of Michigan, Colorado School of Mines, NEUDC 2015, PacDev 2016, the Energy Institute at Haas, University of Chicago, and the 2017 NBER Summer Institute in Energy and Environmental Economics for valuable comments and suggestions. We benefited from conversations with officials at the Indian Ministry of Power, the Rural Electrification Corporation, and the JVVNL Distribution Company in Jaipur. George Fullerton and Puja Singhal assisted us in acquiring the data that made this project possible. All remaining errors are ours.

Recovering causal estimates of the effects of electrification is challenging, since energy infrastructure projects target relatively wealthy or quickly-growing regions. Selection of this kind would bias econometric estimates of treatment effects toward finding large economic impacts. Previous work has relied on instrumental variables strategies to circumvent this problem, and has tended to find large positive effects of electrification. Posited mechanisms for these gains include structural transformation, which in turn changes employment opportunities (Rud (2012)); female empowerment (Dinkelman (2011)); increased agricultural productivity (Chakravorty, Emerick, and Ravago (2016)); health improvements as households switch from kerosene and coal to electricity (Barron and Torero (2017)); and greater educational attainment (Lipscomb, Mobarak, and Barham (2013)).

This paper documents that while large-scale rural electrification causes a substantial increase in energy access and power consumption, it leads at best to small changes in economic outcomes in the medium term. We exploit quasi-experimental variation in electrification generated by a population-based eligibility cutoff in India's massive national rural electrification program, Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY). The "Prime Minister's Rural Electrification Program" was launched in 2005 to expand electricity access in over 400,000 rural Indian villages across 27 states. In order to cap program costs, the Central Government introduced a population-based eligibility cutoff based on the size of village neighborhoods ("habitations").[2] When the program was introduced, only villages with constituent habitations larger than 300 people were eligible for electrification under RGGVY.

We pair detailed geospatial information with rich administrative data on the universe of Indian villages and use a regression discontinuity (RD) design to test for the village-level effects of RGGVY eligibility on employment, asset ownership, household wealth, village-wide outcomes, and education. This design relies on relatively weak identifying assumptions, and we provide evidence that these assumptions are satisfied below. We estimate effects using a main sample of nearly 30,000 villages across 22 states. We demonstrate that RGGVY led to statistically significant and economically meaningful increases in electric power availability and consumption that is visible from space. We then show that despite these gains, electrification led to at most modest changes in economic outcomes. More specifically, we are able to reject even small changes, of 0.26 of a standard deviation, across a range of outcomes, including employment, asset ownership, the housing stock, village-wide outcomes, household wealth, and school enrollment. Taken together, these results suggest that the causal impact of large-scale rural electrification on economic development may be substantially smaller than previously thought.

We show that these small effects do not simply reflect issues with the timing or quality of RGGVY project implementation. Our results are quantitatively similar for villages electrified near the beginning and near the end of our sample period, meaning that any

[2]In the 2001 Indian Census, the village was the lowest-level administrative unit. Villages are composed of habitations (or "hamlets"), which correspond to the inhabited areas of a village. South Asian villages typically have one or more inhabited regions surrounded by agricultural land. India's 600,000 villages contain approximately 1.6 million unique habitations.

confounding rollout effects are unlikely. Likewise, we find quantitatively similar results for the subset of states with above-average power supply reliability, which suggests that even in places with relatively infrequent power outages, the economic impacts remain quite small. We also employ an alternative identification strategy, difference-in-differences (DD), which reveals that our RD results appear to generalize to villages far from our 300-person population cutoff. Using this DD approach, we find treatment effects that are broadly consistent with our RD strategy, across the full support of Indian village populations. Our main RD results also stand up to a battery of placebo tests, falsification exercises, and robustness checks.

This paper makes three key contributions to the existing literature. First, our results contrast starkly with the large economic impacts of electrification found in earlier work. They apply directly to rural villages across 27 states in India, representing the world's largest unelectrified population. Perhaps more importantly, we use a regression discontinuity design to quantify the effects of electrification; this necessitates substantially weaker identifying assumptions than the instrumental variables approaches of the prior literature. Second, we add to the knowledge on the economic effects of infrastructure in developing countries. Existing work in this area has tended to find large positive impacts of infrastructure investments.[3] We provide evidence that electricity infrastructure may not necessarily spur large-scale economic growth. Third, our results contribute to a small but growing literature on energy use in the developing world.[4] We demonstrate that while electrified villages are consuming power, this energy use does not appear to be transforming rural economies.

The remainder of the paper proceeds as follows: Sections 2.2, 2.3, and 2.4 describe rural electrification in India, our empirical strategy, and the data used in our analysis. Section 2.5 presents our main empirical results, which we discuss and interpret in Section 2.6. Section 2.7 concludes.

2.2 RGGVY

At the time of its independence in 1947, only 1,500 of India's villages had access to electricity (Tsujita (2014)). By March 2014, that number had risen to 576,554 out of 597,464 total villages. This massive technological achievement is largely attributable to a series of national electrification programs, the first of which began in the 1950s. The flagship program of India's modern electrification efforts was Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), or the Prime Minister's Rural Electrification Plan. Prior to RGGVY, over 125,000 (21 percent) of rural villages had no access to power whatsoever. Many of the remaining villages had extremely limited power access; 57 percent of all rural households lacked access to electricity, with the majority of unelectrified households falling

[3] See, for example, Donaldson (2018) on the effects of railroads on trade costs and welfare in India and Banerjee, Duflo, and Qian (2012) and Faber (2014) on roads in China.

[4] See Gertler et al. (2016) on income growth and energy demand, Allcott, Collard-Wexler, and O'Connell (2016) on power outages, and McRae (2015) on energy infrastructure.

below the poverty line. Substantial opportunities remained to expand electricity access in rural communities.

RGGVY was launched in 2005 with the goal of extending power access to over 100,000 unelectrified rural villages in 27 Indian states. The program also set out to provide more intensive electrification to over 300,000 "under-electrified" villages. RGGVY's primary mandate was to install and upgrade electricity infrastructure — specifically transmission lines, distribution lines, and transformers — in order to support electric irrigation pumps, small-to-medium industries, cold chains, healthcare, schooling, and information technology applications. Such infrastructure investments aimed to "facilitate overall rural development, employment generation, and poverty alleviation" (Ministry of Power (2005)). RGGVY also extended electric connections to public places, including schools, health clinics, and local government offices. While the program focused on providing electricity infrastructure to support growing village economies, RGGVY was also charged with extending household electricity access by offering free grid connections to all households below the poverty line.[5] RGGVY investments occurred primarily on the intensive margin, upgrading existing infrastructure to have the capability to power growing rural economies. The majority of RGGVY works, including new grid connections, occurred in villages with some degree of household electrification prior to 2005.

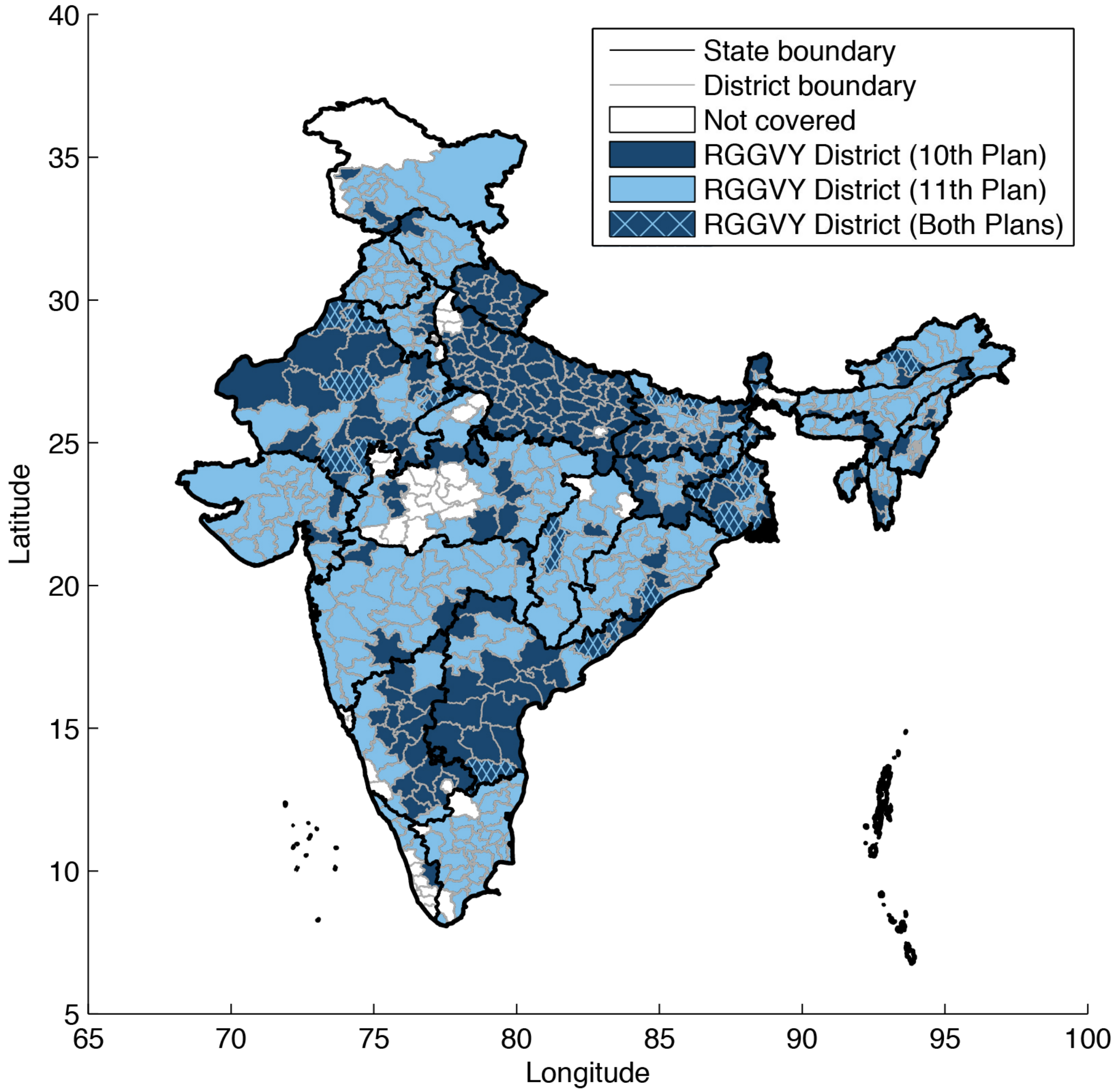
In order for a village to be electrified under RGGVY, its state government had to submit an implementation proposal to the Rural Electrification Corporation (REC), a public-private financial institution overseen by the national government's Ministry of Power. These district-specific proposals, or Detailed Project Reports (DPRs), were based on village-level surveys carried out by local electric utilities, covering both unelectrified villages and partially electrified villages in need of "intensive electrification." Each DPR proposed a village-by-village implementation plan, which included details on new electricity infrastructure to be installed and the number of households and public places to be connected. The REC reviewed DPR proposals, approved projects, and disbursed funds to states.

Funding for RGGVY came from India's 10th (2002–2007) and 11th (2007–2012) Five-Year Plans.[6] Districts were sorted into Plans on a first-come, first-serve basis: the first group of approved DPRs were allocated funding under the 10th Plan, and the next group were allocated funding under the 11th Plan. Under the 10th Plan, all villages with habitation populations above 300 were eligible for RGGVY electrification. Under the 11th Plan, this threshold was decreased to 100. Approximately 164,000 (267,000) villages in 229 (331) districts in 25 (25) states were slated for electrification under the 10th (11th) Plan, which also targeted 7.5 million (14.6 million) below-poverty-line households for free

[5] Above poverty line households were able to purchase connections. All households were required to pay for their own power consumption. The program did not subsidize the consumption of electricity for any household, but Indian retail electricity tariffs are heavily subsidized, and average 2.4 rupees (4 U.S. cents) per kilowatt-hour.

[6] Midway through the 12th Plan, RGGVY was subsumed into Deendayal Upadhyaya Gram Jyoti Yojana (DDUGJY); the remaining projects are slated to be finished by the end of the 13th Plan. As of 2016, all villages are eligible for electrification under DDUGJY, regardless of size.

Figure 2.2.1: Indian Districts by RGGVY Implementation Phase



Notes: This map shows 2001 district boundaries, shaded by RGGVY coverage status. Navy districts are covered under the 10th Plan, light blue districts are covered under the 11th Plan, cross-hatched districts were covered under both the 10th and 11th Plans, and white districts are not covered by RGGVY. As of 2001, India had 584 districts across its 28 states and 7 Union Territories. RGGVY covered 530 total districts in 27 states (neither Goa nor the Union Territories were eligible), with 30 districts split between the 10th and 11th Plans.

connections. Funding for the 10th Plan was disbursed between 2005 and 2010, with over 95 percent of funds released before 2008. The 11th Plan distributed funds between 2008 and 2011.[7]

Figure 2.2.1 shows the spatial distribution of RGGVY districts covered by the 10th and 11th Plans, highlighting the program's broad scope. The vast majority of eligible districts received RGGVY funding under exactly one Five-Year Plan, and 23 out of 27 states contain both 10th- and 11th-Plan districts. We focus our empirical analysis on the districts that received RGGVY funding under the 10th Plan, because electrification in these districts was completed earlier, giving us a longer post-electrification sample period.

2.3 Empirical approach

2.3.1 Regression Discontinuity Design

In this paper, we aim to estimate the causal effect of rural electrification on development. Because energy infrastructure programs are large-scale investments, and because governments allocate funds to specific regions or groups of people in ways that are likely correlated with economic outcomes of interest, it can be challenging to disentangle the impact of electrification from other observed and unobserved factors that affect development. Furthermore, since the electricity grid is spatially integrated, a national-scale rollout of electrification is likely to have different effects than can be observed by a randomized controlled trial that impacts a few hundred rural villages.[8] To overcome these challenges, we implement a regression discontinuity design, allowing us to identify the causal effect of electrification at scale.

Under the RGGVY program rules, villages in 10th-Plan districts were eligible for treatment if they contained habitations with populations of 300 or above. Our RD analysis includes only villages whose districts received funding under the 10th Plan, and we restrict our sample to villages with exactly one habitation. This allows us to use an RD to estimate local average treatment effects for villages with habitation populations close to this 300-person cutoff. In this sharp RD design, eligibility for treatment changes discontinuously from 0 to 1 as village population (our running variable) crosses the 300-person threshold, allowing us to identify the effects of eligibility for RGGVY on both observed changes in electrification and on village-level economic outcomes.[9]

This design necessitates two main identifying assumptions. First, we must assume continuity across the RD threshold for all village covariates and unobservables that might be correlated with our outcome variables. While this assumption is fundamentally untestable,

[7]We downloaded data on RGGVY implementation from `http://www.rggvy.gov.in`, since replaced with `http://www.ddugjy.gov.in`. Appendix B.3 describes the RGGVY program in greater detail, along with additional background on the history of rural electrification in India.

[8]Lee, Miguel, and Wolfram (2016) are implementing a randomized controlled trial of household electrification in 150 rural communities in Western Kenya.

[9]See Imbens and Lemieux (2008) and Lee and Lemieux (2010) for further detail about the formal assumptions underlying RD analysis, and practical issues in applying RD designs.

we support it with evidence from several key village characteristics.[10] We know of no other Indian social program with a 300-person eligibility threshold. Second, we assume that our running variable, 2001 Census population, is not manipulable around the threshold. Because our running variable predates the announcement of RGGVY in 2005, we are confident that our population data were not influenced by the future existence of RGGVY. Figure 2.4.3 shows no evidence of bunching of villages around this 300-person population cutoff.

Given these assumptions, our RD design provides a consistent estimate of the effect of eligibility for treatment on outcomes of interest for the set of single-habitation villages located in districts that received RGGVY funding under the 10th Plan. Formally, we estimate:

$$\begin{aligned} Y_{vs}^{2011} &= \beta_0 + \beta_1 Z_{vs} + \beta_2(P_{vs} - 300) + \beta_3(P_{vs} - 300) \cdot Z_{vs} + \beta_4 Y_{vs}^{2001} + \eta_s + \varepsilon_{vs} \\ &\text{for } \; 300 - h \leq P_{vs} \leq 300 + h\,, \quad \text{where } \; Z_{vs} \equiv \mathbf{1}[P_{vs} \geq 300]\,. \end{aligned} \tag{2.1}$$

Y_{vs}^{2011} represents the outcome of interest in village v in state s in 2011, P_{vs} is the 2001 village population, Z_{vs} is the RD indicator equal to one for villages above the cutoff, h is the RD bandwidth, Y_{vs}^{2001} is the 2001 value of the outcome variable, η_s is a state fixed effect, and ε_{vs} is an idiosyncratic error term.[11] We cluster our standard errors at the district level to allow for arbitrary dependence between the errors of villages within the same district. This accommodates both implementer-specific correlations within a district's DPR (RGGVY's unit of project implementation) and natural spatial autocorrelation between nearby villages. We use a preferred RD bandwidth of 150 people on either side of the 300-person cutoff; this allows us to include a large sample of villages, while remaining confident that villages away from the discontinuity are similar to those at the 300-person cutoff.[12]

2.3.2 Economic Outcomes

Economic theory suggests that electrification could impact village economies through several channels. First, as electricity becomes available, we should expect small firms to invest in new capital equipment that uses power. This in turn would raise the marginal product of labor in the non-agricultural sector, drawing workers to new employment opportunities (Rud (2012)). On the other hand, electrification could spur agricultural mechanization, which would improve farm productivity (Chakravorty, Emerick, and Ravago (2016)).[13]

[10] We find no evidence to suggest that pre-period covariates change discontinuously across the 300-person cutoff. These results are available in Appendix B.2.4.5, as well as in Figure 2.5.5 below.

[11] Neither the 2001 value of the outcome variable nor the fixed effects are necessary for identification, but they improve the precision of our estimates (see Lee and Lemieux (2010)).

[12] We perform bandwidth sensitivity checks in Appendix B.2.1.2, including calculating the Imbens and Kalyanaraman (2012) optimal bandwidth; our results are not sensitive to bandwidth choice.

[13] In the Indian context, one potential use of electricity in agricultural production is to power irrigation tubewells.

This could either increase or decrease employment in agriculture.[14] However, because the marginal product of labor would unambiguously increase in both the agricultural and non-agricultural sectors, this should increase wages, incomes, and expenditures.

Next, electricity access may lead to gains for women. New employment opportunities, like those described above, could enable more women to work outside the home. Alternatively, newly-electrified households could invest in labor-saving devices, which could decrease the time required for women to complete household duties. This could also lead to increased female employment, either outside the home or in microenterprises. Dinkelman (2011) uses an instrumental variables approach in South Africa, and finds that electrification substantially raises female employment through this latter channel.

Rural electrification may also bring substantial health benefits. Kerosene is widely used throughout the developing world as a fuel for both lighting and cooking, and Indian households also commonly cook with coal and biomass. Combustion of these fuels produces harmful indoor air pollution, which is especially detrimental to young children and infants in utero. Access to electricity may foster investment in electric lights and electric cookstoves, which would likely reduce indoor air pollution and improve child health outcomes (Barron and Torero (2017)). Electrification may also indirectly improve health outcomes, through higher incomes and improved access to health care.

Finally, electrification could impact educational attainment through several channels. On the extensive margin, total school enrollment may increase if electrification leads to income gains, making households less reliant on child labor earnings. On the other hand, rising wages may draw students out of school and into the labor force. Alternatively, we might expect electricity access to change the education production function. Lighting or computing facilities in schools may improve learning in the classroom, and children in homes with electric lighting will likely develop more effective study habits. If electrification improves student performance, it could affect the intensive margin of schooling as students tend to stay in school longer, causing enrollment in upper grades to increase. Using instrumental variables, Lipscomb, Mobarak, and Barham (2013) find that rural electrification increases the number of years that students attend school.

2.4 Data

Our empirical analysis uses data from four main sources. First, we link satellite images of nighttime brightness to village boundary shapefiles, yielding a panel of village brightness. Next, we use several large administrative datasets published by three different Indian government entities, which contain village populations and a broad set of economic indicators. Armed with a wealth of data on Indian villages, we can test the channels through which we expect electrification under RGGVY to impact economic development.

[14] The potential for changes in agricultural employment depends on several factors, including the excess supply of labor, the excess supply of farmland, the degree to which farm mechanization and agricultural labor are complements or substitutes, and the effect of electricity access on agricultural commodity prices.

2.4.1 Nighttime Lights Data

In order to understand the economic effects of electrification resulting from RGGVY, we must first demonstrate that RGGVY led to a meaningful increase in electricity access and consumption in rural Indian villages. A binary indicator of the presence of electricity infrastructure would be insufficient, since it would mask heterogeneity in power quality, electricity consumption, and connection density. There exists no comprehensive dataset of power consumption at the village level across India, but we are able to construct a measure of electricity consumption using remotely-sensed data.

As an indicator of electrification under RGGVY, we use changes in nighttime brightness as observed from space. The National Oceanic and Atmospheric Administration's Defense Meteorological Satellite Program–Operational Line Scan (DMSP-OLS) program collects images from U.S. Air Force satellites, which photograph the earth daily between 8:30pm and 10:00pm local time. After cleaning and processing these images, NOAA averages them across each year and distributes annual composite images online.[15] Each yearly dataset reports light intensity for each 30 arc-second pixel (approximately 1 km^2 at the equator) on a 0–63 scale, which is proportional to average observed luminosity.[16] Figure 2.4.2 shows nighttime brightness in India in 2001 and 2011.

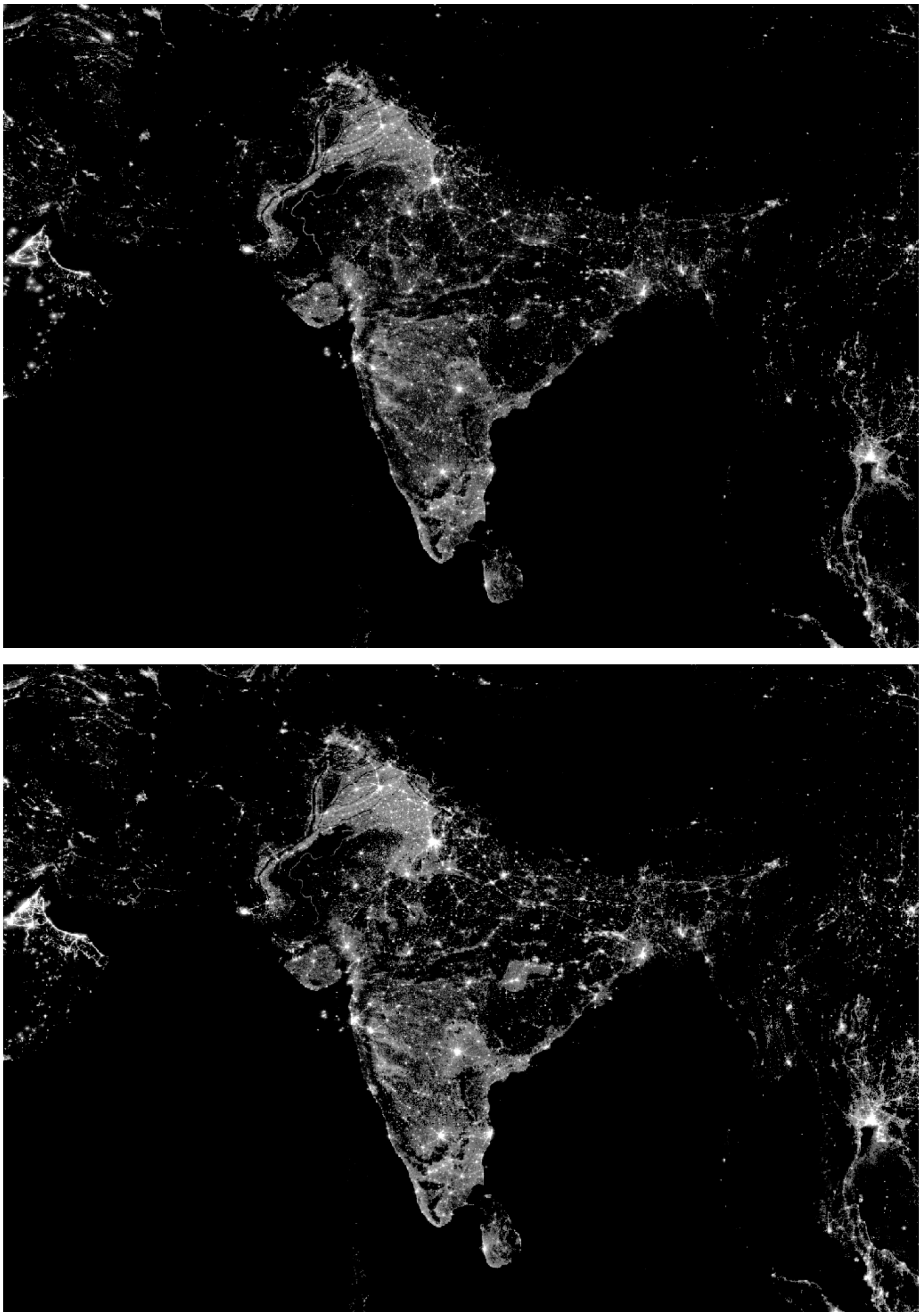
Economists frequently use these nighttime lights data as proxies for economic activity, as popularized by Chen and Nordhaus (2011) and Henderson, Storeygard, and Weil (2012). Existing work demonstrates that nighttime brightness can also be used to detect electrification, even at small spatial scales: Min et al. (2013) find evidence of a statistically detectable relationship between NOAA DMSP-OLS brightness and the electrification status of rural villages in Senegal and Mali. Min and Gaba (2014) show that a similar correlation between electrification and nighttime brightness also exists in rural Vietnam. Chand et al. (2009) show a direct relationship between nighttime lights and electric power consumption in India, while Min (2011) finds a strong correlation between brightness and district-level electricity consumption in Uttar Pradesh. We build on this research by using nighttime brightness to demonstrate that RGGVY successfully increased village electricity access, where nighttime lights serve an objective measure of realized energy consumption in these villages.

Importantly, these satellite images represent a lower bound on electricity consumption. While nighttime brightness data record light output (including lighting from houses, public spaces, and outdoor streetlights), they do not directly measure electricity consumed for other purposes. Because all electricity end-uses rely on the same power grid, we treat increases in nighttime brightness as necessary indicators of investments in electricity infrastructure. Likewise, if total electricity consumption increases, we should expect

[15]This cleaning removes any sunlit hours, glare, cloud cover, forest fires, the aurora phenomena, and other irregularities. Nighttime lights data are available for download at `http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`. We use the average lights product in our main analysis. See Appendix B.1.3 for further discussion.

[16]Chen and Nordhaus (2011) detail the relationship between physical luminosity and brightness in the nighttime lights images.

Figure 2.4.2: Nighttime Lights in India, 2001 and 2011



Notes: This figure shows the DMSP-OLS nighttime brightness data for India. The top panel shows nighttime lights in 2001, and the bottom panel shows nighttime lights for 2011. The $\approx$1km^2 pixels in this image range in brightness from 0 to 63, covering the full range of the DMSP-OLS data.

nighttime brightness to increase as well, as more power reaches rural villages. A potential concern with using nighttime lights to proxy for total electricity consumption is that we could mistake new sources of outdoor lighting for increases in electricity access. However, RGGVY's primary mandate was to expand and improve electricity infrastructure, and there is no mention of streetlight installation in 10th-Plan program documentation.[17] Hence, an observed increase in nighttime brightness as a result of RGGVY would very likely be driven not by new streetlights alone, but rather by village-wide increases in access to energy services.

We construct a village-level panel of nighttime brightness by overlaying annual NOAA DMSP-OLS images with 2001 village shapefiles.[18] Our preferred measure of a village's lighting is the maximum brightness of any pixel whose centroid lies within its borders.[19] We use the brightest pixel because Indian villages are typically organized such that there are centralized populated areas surrounded by fields. This targets our electrification measure at the populated parts of villages, while avoiding measurement error from brightness averaged across unlit agricultural land.[20] In performing this calculation, we are forced to drop 10 states from our sample. We are missing shapefiles for five states, which represent fewer than 3 percent of the total villages covered by RGGVY. We also exclude five states because we believe these shapefiles to be of extremely low quality: the correlation between the village area implied by the shapefiles and village area recorded by the Indian Census, the entity in charge of defining village boundaries, is below 0.35.[21] We are left with a nighttime lights sample of 370,689 villages across 15 states. We do not impose these sample restrictions for any other outcome variables.

2.4.2 Census of India

We combine several village-level datasets published by the Census of India from the 2001 and 2011 decennial Censuses.[22] The Primary Census Abstract (PCA) contains village

[17]RGGVY 11th-Plan documentation did discuss streetlights in the context of a small carve-out for microgrids targeted at extremely remote villages. Because this carve-out did not exist under the 10th Plan, the 300-person eligibility cutoff did not apply for these villages.

[18]Indian villages have official boundaries, which are recorded by the Census Organization of India. Every square meter in India (excluding bodies of water and forests) is contained in a city, town, or village. We use shapefiles of village boundaries published by ML InfoMap, Ltd.

[19]We calculate this level in ArcGIS, using the standard `Zonal Statistics as Table` operation. For villages too small to contain a pixel's centroid, we assign the brightness value of the pixel at the village centroid.

[20]Our results remain largely unchanged if we use the mean lights value rather than the maximum value. We also undertake a procedure to remove measurement error from the nightlights data via linear projection. See Appendix B.1.3 for details.

[21]The five states with missing shapefiles are Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, and Sikkim. The five states with low-quality shapefiles and village areas are Assam, Himachal Pradesh, Jammu and Kashmir, Uttar Pradesh, and Uttarakhand. The remaining states in the sample all have correlations between datasets above 0.6. See Appendix B.1.2 for further discussion.

[22]These data are all publicly available at `http://www.censusindia.gov.in`. Because our research design relies on observing a large number of villages with populations around 300, we are unable to use

population data, and a detailed breakdown of labor allocation by gender and job type. In particular, the PCA reports the number of men and women that are working in agriculture; "household industry workers" (engaged in informal production of goods within the home); and "other workers" that engage in all other types of work.[23] Examples of "other workers" include government servants, municipal employees, teachers, factory workers, and those engaged in trade, commerce, or business. These data allow us to test for sectoral shifts in employment due to RGGVY electrification, either away from agriculture (consistent with structural transformation) or into agriculture (consistent with increased agricultural productivity). We also test for effects on female employment. Because we observe the share of women engaged in economic activity both outside and within the home, these data are well-suited to capture potential impacts of electrification on female labor.

The Houselisting Primary Census Abstract (HPCA) provides extensive data on living conditions, household size, physical household characteristics, and asset ownership. These data report the fraction of households that own a variety of assets, including radios, mobile phones, bicycles, motorcycles, and televisions. RGGVY may have contributed both directly and indirectly to asset ownership, if households purchased electric appliances to take advantage of improved power availability, or if potential income games from electrification enabled increased household expenditures on durable goods. Physical housing characteristics such as floor and roof materials are indicators of household wealth. If RGGVY spurred increases in household expenditures, we expect to observe medium-run investments to improve the housing stock. The HPCA also allows us to examine the health channel, as this dataset reports the fraction of households that cook with electricity and that use kerosene as a main source of lighting.

Finally, the Village Directory (VD), another Census dataset, contains detailed information on village amenities.[24] In particular, the VD includes data on the presence of education and medical facilities; banking facilities and agricultural credit societies; the existence and quality of road network connections and the presence of bus services; and communications access, including postal services and mobile phone networks. We use these data to test for the effects of RGGVY on village amenities. The VD also includes information on village electrification, in the form of binary indicators of electric power availability in each village, separately for the agricultural, domestic, and commercial sectors. These indicators are coded as "1" if *any* electric power was available for a given end use anywhere in the village, and as "0" otherwise. Two-thirds of RGGVY 10th-Plan villages met this criterion at baseline (i.e. were coded as "1" for electric power availability), making these variables particularly poorly suited to analyze the effects of RGGVY. The main goals of RGGVY were to upgrade energy infrastructure and increase the penetra-

---

additional Indian survey datasets such as the NSS or ASI. These datasets do not include a sufficient number of small villages to support our RD analysis, and are not designed to be representative below the district level.

[23]The agriculture category is decomposed further into "cultivators" (on their own land) and "agricultural laborers" (on others' land).

[24]In 2001, the VD was a separate Census product. In 2011, it was bundled into the District Census Handbook (DCHB).

tion of electricity access within each village. The VD data contain no information on the intensity of electrification within a village, and therefore do not reflect the vast majority of RGGVY works.[25] We instead turn to the nighttime lights data, which allow us to track intensive-margin changes in energy consumption.

We combine the PCA, HPCA, and VD data into a two-wave village-level panel. The 2001 PCA also reports the official 2001 population of each village, which was the population of record for the RGGVY program, and which we use as our RD running variable.[26] However, RGGVY implementing agencies were instructed to determine eligibility based on 2001 *habitation* (sub-village neighborhood) populations. To the best of our knowledge, the only nation-wide habitation census in existence was conducted by the National Rural Drinking Water Program.[27] We use a fuzzy matching algorithm, modified from Asher and Novosad (2018), to link this habitation census to our village panel and identify the 50 percent of villages with exactly one habitation.[28] For these single-habitation villages, habitation populations are equivalent to village populations—meaning that 2001 village population should exactly correspond to the population that determined RGGVY eligibility for these villages.

The main dataset for our analysis contains the 2001–2011 Census, nighttime brightness, RGGVY program implementation details, and the number of habitations in each village. The subsample of single-habitation, 10th-Plan villages comprises 20 percent of Indian villages.[29] After restricting this 20 percent sample to our preferred RD bandwidth of 150 people above and below the 300-person threshold, we are left with 29,765 10th-Plan single-habitation villages from 22 states.[30] The top panel of Figure 2.4.3 displays a histogram of village populations, showing that the modal village lies within our RD window of 150–450 people. The bottom panel demonstrates how our two sample restrictions reduce the size of our RD sample, and shows that our running variable, 2001 village population, is smooth across the RD threshold.

Table 2.4.1 reports 2001 summary statistics for three sets of villages with populations between 150 and 450: all Indian villages, all villages in 10th-Plan districts, and all villages in 10th-Plan districts that have only one habitation. On average, villages in 10th-Plan districts are geographically smaller and less electrified than the national average, but

[25] The 2011 Village Directory also reports the average hours of electricity available per day, by sector. Because electricity is distributed over an integrated grid, it is unlikely that RGGVY's infrastructure upgrades would have any effect on these measures of electricity access.

[26] RGGVY ledgers we observed in Rajasthan were pre-printed with 2001 Census populations.

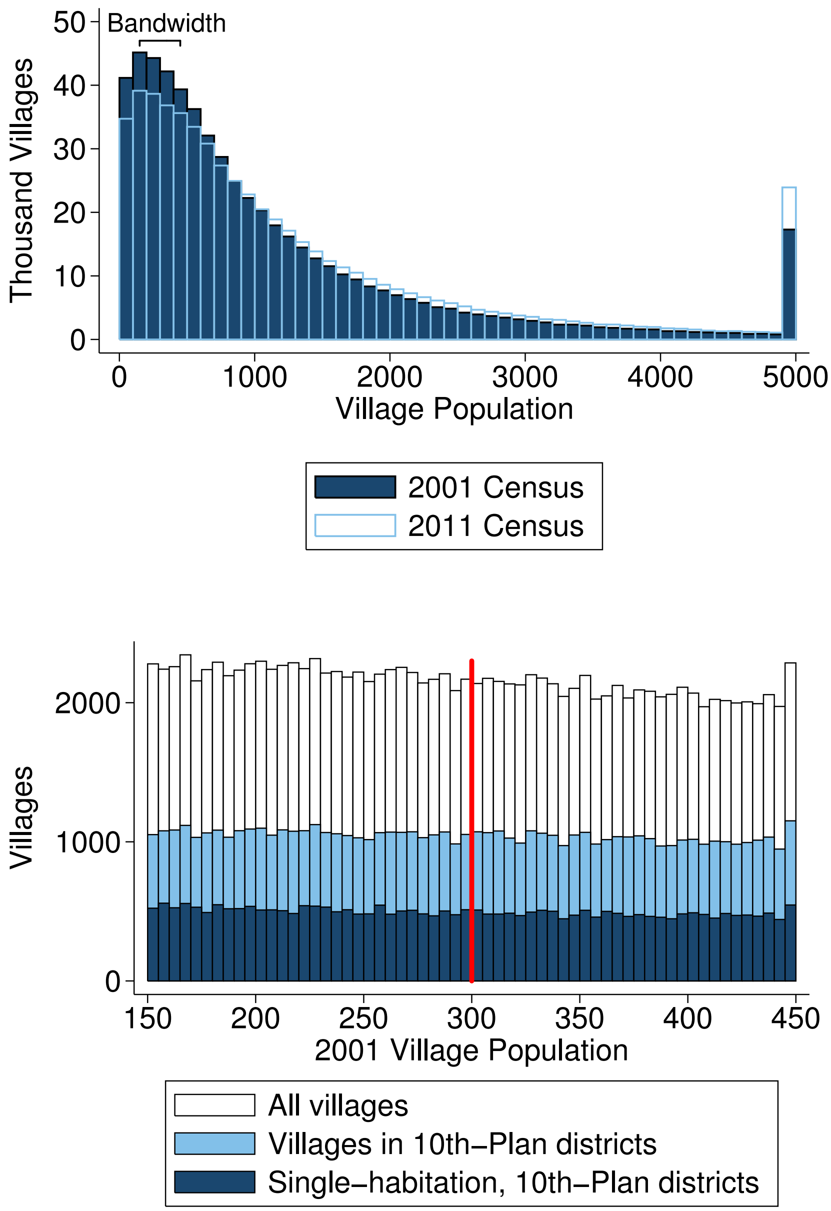
[27] Administered by the Ministry of Drinking Water and Sanitation, this census of habitations was collected in 2003 and 2009, and is available at `http://indiawater.gov.in`.

[28] We thank the authors for sharing their code. Appendix B.1.5 details our matching algorithm.

[29] 50 percent of villages are in districts eligible under RGGVY's 10th Plan, 86 percent of villages match to the habitation census, and 52 percent of matched villages in 10th-Plan districts have one habitation. Our analysis excludes villages that match to the habitation census but have populations that disagree by over 20 percent across datasets, as these matches are likely erroneous. In Appendix B.2, we show that including these villages slightly attenuates our RD point estimates as expected, yet they remain statistically significant.

[30] Three small states with 10th-Plan districts (Manipur, Kerala, and Tripura) are excluded from our final regression because they have no villages that meet these criteria.

Figure 2.4.3: Density of RD Running Variable



Notes: This figure summarizes the distribution of Indian village populations. The top panel shows the population distribution of villages in India in 2001 (solid navy) and 2011 (hollow blue). The bottom panel zooms in on the set of villages used in our RD analysis, within a 150–450 population window around the 300-person cutoff. Our RD sample of single-habitation 10th-Plan villages is shown in navy, relative to all Indian villages (white) and all villages in 10th-Plan districts (light blue).

similar across a range of other covariates. 10th-Plan villages with only one habitation are very similar on observables to average 10th-Plan villages.

2.4.3 Socioeconomic and Caste Census

We draw on individual-level microdata from the Socioeconomic and Caste Census (SECC) for measures of income and alternative employment data. The SECC was collected between 2011 and 2012, with the goal of enumerating the full population of India. We obtained a subset of these data from the Ministry of Petroleum and Natural Gas, whose liquid petroleum gas subsidy program, Pradhan Mantri Ujjwala Yojana, uses SECC data to determine eligibility.[31] As a result, we observe the universe of rural individuals that are eligible for this fuel subsidy program. This includes all individuals living in households that satisfied at least one of seven poverty indicators, and that did not meet any of fourteen affluence criteria.[32] This yields a dataset of data of 332 million individuals from 81 million households, representing roughly half of all households in rural India.

For this selected sample, we observe individual-level data on age, gender, employment, caste, and marital status; and household-level data on the housing stock, land ownership, asset ownership, and income sources. We use the SECC to test for the effects of RGGVY on wealth, using three main indicators. First, we test for the fraction of households with at least one poverty indicator (and no affluence indicators), as measured by the fraction of 2011 Census households that appear in our SECC dataset. Next, the SECC contains an indicator for whether the main income earner in each household earns at least 5,000 rupees per month.[33] This represents the highest-resolution measure of household income in a large-scale Indian dataset, enabling us to directly, albeit coarsely, test the effect of electrification on income. We also use SECC data to test for the effects of RGGVY on the fraction of households that own land or have at least one salaried laborer, two additional wealth indicators. Finally, we construct SECC employment variables that are analogous to the Census's village-wide measures, allowing us to test for distributional employment effects among the subset of households with poverty indicators.

2.4.4 District Information System on Education

In order to estimate the effects of electrification on education, we include data on the universe of Indian primary and upper primary schools from the 2005–2006 school year

[31] The Ministry of Rural Development, who collected the SECC, are in the process of making the full dataset publicly available. As of now, only district-level aggregates are posted at `http://secc.gov.in/welcome`. We downloaded our data in Excel format from `http://lpgdedupe.nic.in/secc/secc_data.html`.

[32] The sample also excludes the less than 1 percent of the population that met one of five destitution indicators. See Appendix B.1.6 for more details on the inclusion and exclusion criteria. We are missing data from six rural districts, which represent less than 1 percent of Indian villages.

[33] All households whose primary earner made over 10,000 rupees per month were ineligible for the fuel subsidy program, and are not included in our SECC dataset.

Table 2.4.1: Summary Statistics – Villages with Populations Between 150 and 450

2001 Village Characteristics	All Districts	10th-Plan Districts	10th-Plan Districts Single-Habitation
Village area (hectares)	199.74	177.98	173.53
	(462.39)	(561.29)	(661.57)
Share of area irrigated	0.23	0.30	0.35
	(0.30)	(0.33)	(0.34)
Agricultural workers / all workers	0.39	0.37	0.37
	(0.16)	(0.16)	(0.15)
Other workers / all workers	0.06	0.06	0.06
	(0.08)	(0.08)	(0.08)
Employment rate	0.46	0.44	0.44
	(0.14)	(0.14)	(0.14)
Literacy rate	0.45	0.44	0.45
	(0.18)	(0.17)	(0.17)
Education facilities (0/1)	0.66	0.58	0.58
	(0.47)	(0.49)	(0.49)
Medical facilities (0/1)	0.13	0.12	0.12
	(0.34)	(0.32)	(0.32)
Banking facilities (0/1)	0.01	0.01	0.01
	(0.11)	(0.11)	(0.10)
Agricultural credit societies (0/1)	0.03	0.03	0.03
	(0.18)	(0.16)	(0.16)
Electric power (0/1)	0.68	0.62	0.64
	(0.46)	(0.49)	(0.48)
Share households with indoor water	0.21	0.21	0.25
	(0.17)	(0.17)	(0.19)
Share households with thatched roofs	0.27	0.27	0.28
	(0.27)	(0.24)	(0.24)
Share households with mud floors	0.78	0.79	0.77
	(0.17)	(0.16)	(0.17)
Average household size	5.36	5.53	5.56
	(0.58)	(0.61)	(0.60)
Number of villages	129,438	62,638	29,765

Notes: This table shows village-level summary statistics from the 2001 Census, for three sets of villages with 2001 populations between 150 and 450: all villages, villages in 10th-Plan districts, and single-habitation villages in 10th-Plan districts. This third group corresponds to the sample of villages used in our RD analysis. We present workers by sector as the share of total workers in the village; "other" workers are classified as non-agricultural, non-household workers. The employment rate divides the number of workers by village population. Binary variables are labeled (0/1). Standard deviations in parentheses.

through the 2014–2015 school year.[34] These data come from the District Information System on Education (DISE), which reports annual school-level snapshots on a variety of student, teacher, and school building characteristics. We collected these data at the school level and construct a 10-year panel dataset containing information from 1.68 million unique schools.[35] This panel is strongly unbalanced, and the average school appears in 7 out of a possible 10 years. Given that the reporting of school characteristics varies considerably across years, we focus our analysis on village-wide enrollment counts, which are consistently reported by gender and grade level. We test for effects of RGGVY on total enrollment, enrollment by gender, and enrollment by grade level, which allows us to measure how electrification impacted both the extensive and intensive margins of schooling.

2.5 Regression Discontinuity Results

2.5.1 Electrification

In order to demonstrate that RGGVY had a meaningful effect on electrification in eligible villages, we examine the effects of eligibility for RGGVY on nighttime brightness. Specifically, we use Equation (2.1) to estimate the effect of having a 2001 population above the RGGVY cutoff on village brightness in 2011. After removing states with low-quality or missing shapefiles, we are left with a sample of 18,686 single-habitation villages, in RGGVY 10th-Plan districts across 12 states, with populations in our RD bandwidth of 150–450 people.

Figure 2.5.4 presents the results from our preferred RD specification graphically, while Table 2.5.2 reports the corresponding numerical results. We find that 2011 nighttime brightness increased discontinuously at the 300-person threshold by 0.15 units of brightness. This jump is statistically significant at the 5 percent level, with a p-value of 0.015.[36] Appendix B.2.1 demonstrates that this is robust to a range of alternative bandwidths, functional forms, and specifications.

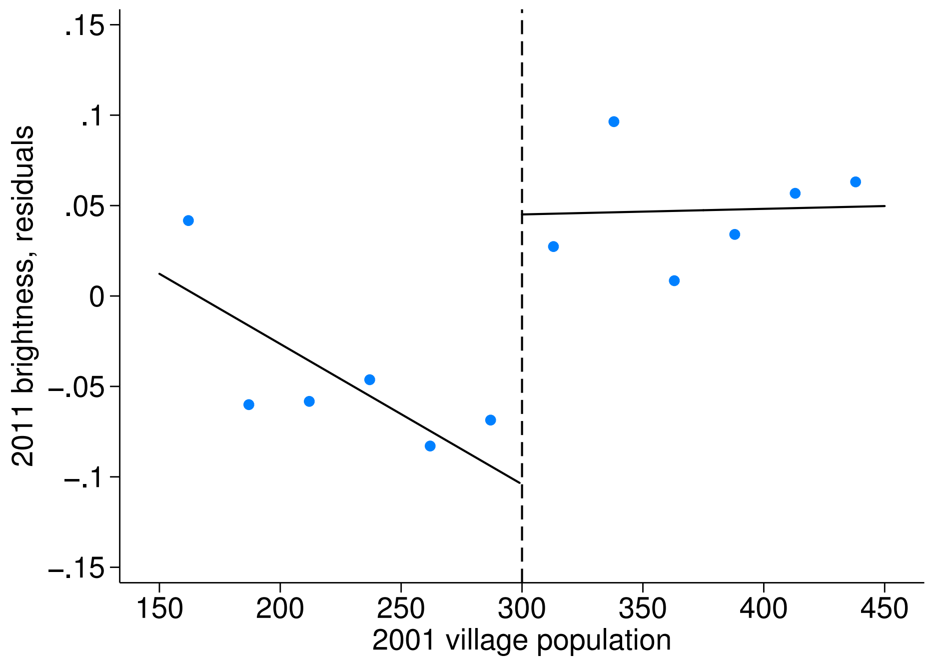
Though this point estimate might seem small, these results in fact demonstrate that RGGVY eligibility led to a substantial increase in brightness for barely-eligible villages as compared to barely-ineligible villages. To interpret these effects, we turn to the remote

[34] While we use the full time series to match DISE schools to Census villages, we restrict our analysis to the 2010–11 school year, for consistency with our other outcome variables.

[35] We downloaded these data from `http://schoolreportcards.in/SRC-New/`. See Appendix B.1.7 for details.

[36] These results include a control for 2001 nighttime brightness. Due to substantial cross-sectional heterogeneity, conditioning on the pre-period level dramatically improves the signal-to-noise ratio. This is common practice with remote sensing data (see also Jayachandran et al. (2017)). If we restrict the RD sample to include only villages that had electric power availability, according to the 2001 Census, we recover a nearly identical result ($\hat{\beta}_1 = 0.16$ with a p-value of 0.046). This suggests that the Census's 1/0 indicator variable for electric power availability masks substantial changes in electricity access under RGGVY, which we are able to detect using nighttime lights.

Figure 2.5.4: RD – 2011 Nighttime Brightness



Notes: This figure shows RD results using maximum 2011 nighttime brightness as a dependent variable, as reported in Table 2.5.2. Blue dots show average residuals from regressing the 2011 maximum nighttime brightness on 2001 maximum nighttime brightness and state fixed effects. Each dot contains approximately 1,600 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for 18,686 single-habitation villages between 150–450 people, in 10th-Plan districts. The point estimate on the level shift is 0.149, with a p-value of 0.015. Neither slope coefficient is significant at conventional levels.

Table 2.5.2: RD – Nighttime Brightness

	2011 village brightness
1[2001 pop ≥ 300]	0.1493**
	(0.0603)
2001 population	−0.0008
	(0.0007)
1[2001 pop ≥ 300] × 2001 pop	0.0008
	(0.0008)
2001 Control	Yes
State FEs	Yes
RD bandwidth	150
Number of observations	18,686
Number of districts	130
Mean of dependent variable	6.370
R^2	0.766

Notes: This table shows results from estimating Equation (2.1), which corresponds to Figure 2.5.4. We define village brightness based on the brightest pixel contained within the village boundary. This regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

sensing literature. The magnitude of the effect we observe is consistent with ground-truthed estimates by Min et al. (2013), who find that electrification is associated with a 0.36-unit increase in nighttime brightness in rural villages in Senegal.[37] Our point estimate of 0.15 is on the same order of magnitude but smaller, which is to be expected, given that villages in our RD bandwidth are significantly smaller than the villages studied in Senegal. In a similar exercise, Min and Gaba (2014) find that a 1-unit increase in brightness corresponds to 60 public streetlights or 240–270 electrified homes in Vietnamese villages.[38]

Extrapolating these results to the Indian context, our estimated 0.15-unit increase translates to roughly 9 additional streetlights per village. This represents a substantial increase in nighttime luminosity, especially considering that RGGVY did not install streetlights. Alternatively, if we extrapolate the (weaker) household relationship to our setting, a 0.15-unit increase would translate to roughly 38 newly electrified homes, or 68 percent of households in the average village in our RD sample. These estimates from Senegal and Vietnam suggest that our effect size in India is consistent with a substantial increase in village electrification under RGGVY, especially given that many electricity end-uses that RGGVY sought to enable are not captured by the nighttime brightness proxy.[39]

We perform a series of validity tests in order to demonstrate that this increase in brightness is, in fact, attributable to the RGGVY program. First, we estimate Equation (2.1) using 2005 nighttime brightness as the dependent variable. Because RGGVY was announced in 2005 and nearly all project implementation began in subsequent years, we should not expect to find an immediate effect of program eligibility on brightness. The top-left panel of Figure 2.5.5 shows no visual evidence of a discontinuity in 2005 brightness at the 300-person threshold. The point estimate in this regression is 0.031, with a standard error of 0.020, and is not statistically significant at conventional level. This demonstrates that nighttime brightness was smooth at the 300-person cutoff prior to RGGVY.[40]

Next, we conduct a placebo test using 801 placebo RD "thresholds" between 151 and 1000.[41] For each threshold, we re-estimate Equation (2.1) and save $\hat{\beta}_1$. We plot the distribution of these placebo coefficients in the top-right panel of Figure 2.5.5. We

[37]This result uses the same average annual DMSP–OLS product that we use, unlike many of the other results reported in the paper, which rely on monthly composites that are not publicly available. We exclude the Mali results described in Min et al. (2013) because the authors exclude them from their main regression estimates.

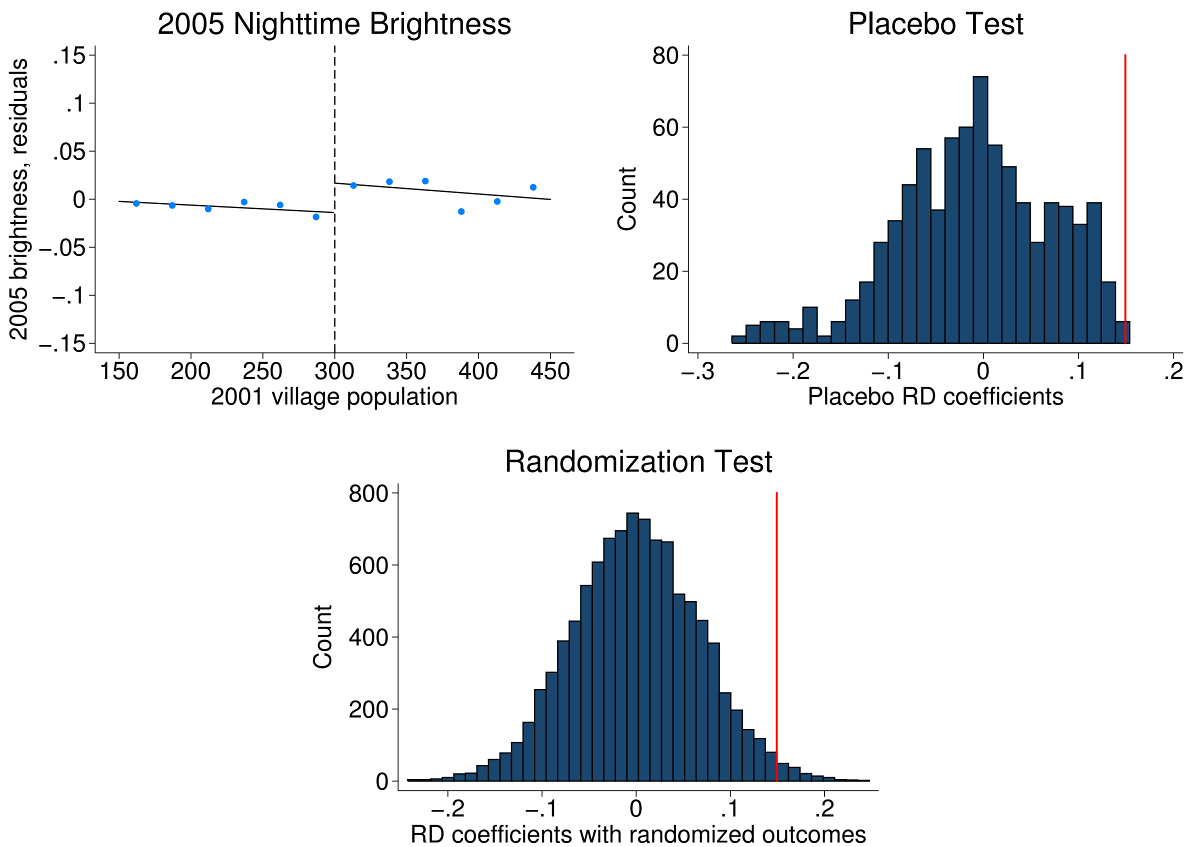
[38]The relationship between nighttime brightness and streetlights is predictably stronger than the relationship between nighttime brightness and electrified homes.

[39]While many factors could cause the relationship between household electrification and nighttime brightness to differ between India and West Africa or Vietnam, Min et al. (2013) and Min and Gaba (2014) provide evidence that the magnitude of our RD point estimate is consistent with what we might expect from a substantial increase in electricity access in these small villages.

[40]We perform a variety of additional pre-period covariate smoothness checks in Appendix B.2.4.5, and find no evidence of discontinuities prior to RGGVY. Appendix B.2.3 demonstrates that the discontinuity in brightness steadily increases from 2006 onward.

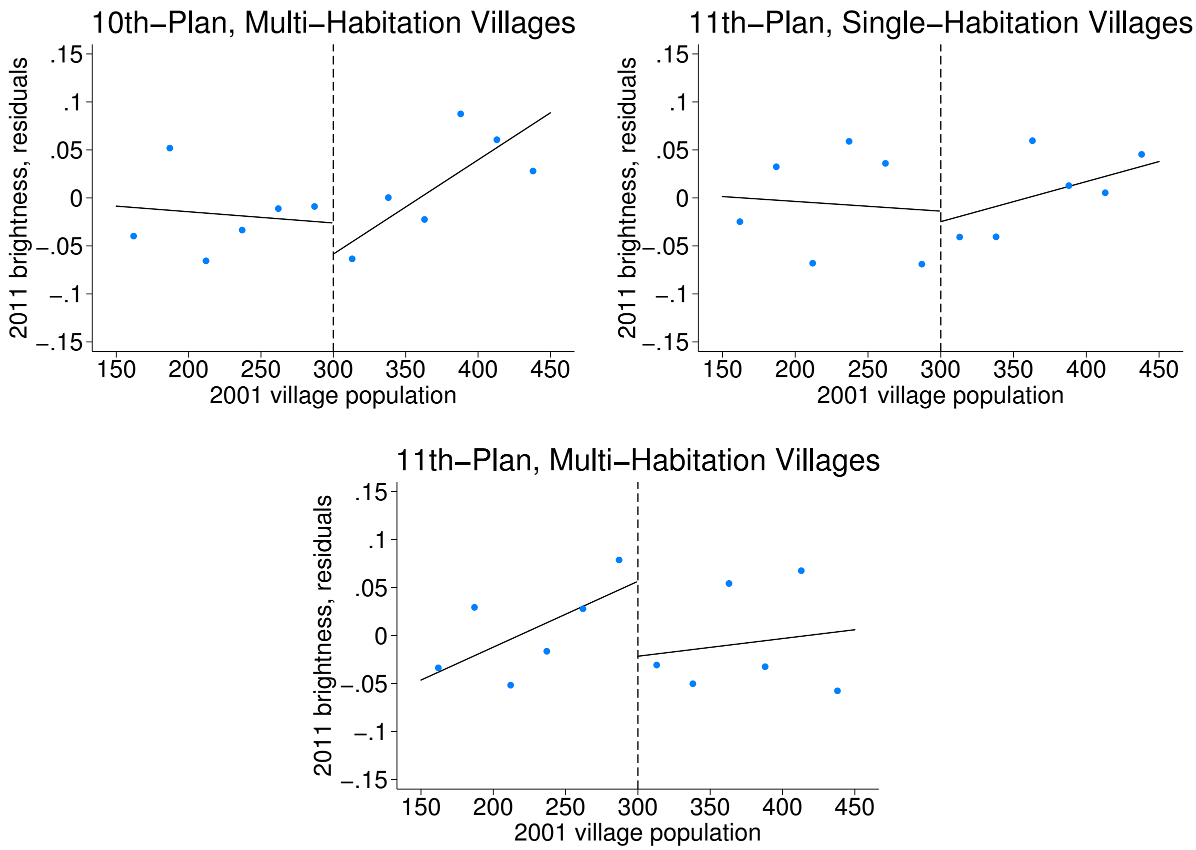
[41]We test all 801 integer values in $[151, 275] \cup [325, 1000]$, which is asymptotically equivalent to simulating placebo draws across this discrete support. We omit thresholds between 275 and 325 to avoid

Figure 2.5.5: Nighttime Brightness – Validity Tests



Notes: This figure presents results from three RD validity checks. The left panel displays results from estimating our main specification using 2005 brightness as the dependent variable; the point estimate is 0.031 with a standard error of 0.020. The center panel was generated by estimating Equation (2.1) on 801 placebo RD thresholds, representing all integer values in $[151, 275] \cup [325, 1000]$. We omit placebo thresholds within 25 people of the true 300-person threshold to ensure that placebo RDs do not detect the true effects of RGGVY eligibility, and we exclude thresholds below 151 due to our 150-person bandwidth. The right panel was generated by scrambling village brightness 10,000 times and re-estimating Equation (2.1). The red lines represent the RD coefficient from the actual data at the correct 300-person threshold. Our RD point estimate falls above the 99th percentile of the placebo distribution and above the 98th percentile of the randomization distribution.

Figure 2.5.6: Nighttime Brightness – Falsification Tests



Notes: This figure presents three falsification tests for our RD on nighttime brightness. The top-left and bottom panels include only villages with multiple habitations, for which the running variable of village population did not determine village eligibility. The top-right and bottom panels include only villages in districts that became eligible for RGGVY under the 11th Plan, for which the appropriate eligibility cutoff was lowered from 300 to 100 people. Blue dots show average residuals from regressing 2011 nighttime brightness on 2001 brightness and state fixed effects. Each dot contains approximately 900–1,600 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for villages within the 150–450 population bandwidth. Supplementary Table B.2.10 reports the regression results that correspond to these figures.

also perform a randomization inference exercise, by scrambling the relationship between nighttime brightness and village population 10,000 times.[42] For each iteration, we estimate Equation (2.1), and the bottom panel of Figure 2.5.5 shows the resulting distribution of RD point estimates. The red lines indicate our estimate of $\hat{\beta}_1$, which falls above the 99th percentile of the placebo distribution and above the 98th percentile of the randomization distribution. This provides evidence that our RD estimates do not simply reflect spurious volatility in the relationship between nighttime lights and village population data.

We also perform a falsification exercise based on the implementation details of the RGGVY program. Our RD sample includes only villages that were eligible for RGGVY under the 10th Plan, for which the relevant eligibility cutoff was 300 people. It also includes only those villages confirmed to have exactly one habitation, for which 2001 village population is the appropriate running variable. We should not find effects at the 300-person cutoff on nighttime brightness for villages eligible under the 11th Plan, for which the relevant eligibility cutoff was moved from 300 to 100 people. Similarly, we should not find any RD effects for villages comprising multiple habitations, because these villages' populations do not correspond to the habitation populations that determined RGGVY eligibility. Figure 2.5.6 presents RD results estimated using these alternative samples: as expected, none exhibits evidence of a discontinuity at the 300-person cutoff. This provides strong evidence that RGGVY, rather than spurious effects or other programs, is causing these effects.

2.5.2 Economic Outcomes

We now turn to the effects of RGGVY eligibility on village economies, and test for impacts of electrification via each of the potential channels discussed in Section 2.3.2. We estimate Equation (2.1) using outcome variables from six broad categories: employment, asset ownership, housing stock characteristics, village-wide outcomes, household income, and education. Each RD regression uses a dependent variable from 2011, while controlling for 2001 population as the running variable, state fixed effects, and the 2001 level of the dependent variable (unless otherwise noted).

First, we test for employment effects by estimating Equation (2.1) using the total number of male (female) workers in a given category divided by the total male (female) population of a village as the dependent variable.[43] Figure 2.5.7 summarizes these workforce results graphically for each gender and sector, and Panel B of Table 2.5.3 reports them numerically. We find that eligibility for RGGVY caused a 0.7 percentage point decrease in the share of men working in agriculture, on a mean of 42 percent. In contrast, the percentage of men in non-agricultural, non-household labor increased by 0.5 percentage points, on a mean of 10 percent. While these sectoral shifts are statistically

possible contamination of the placebo results with the real threshold. We also avoid placebo thresholds below 151, to ensure positive values across the full 300-person RD window.

[42] We assign lights values to each village by sampling $\{Y_v^{2001}, Y_v^{2011}\}$ pairs without replacement.

[43] 2011 population does not change discontinuously at the 300-person threshold. See Panel A of Table 2.5.3, where we find that RGGVY caused no meaningful changes in village demographics.

significant and in a direction consistent with the structural transformation hypothesis, these effect sizes are very small: we can reject changes in male labor allocation larger than 1.3 percentage points. We find no statistically significant effects of electrification on the share of women working in any sector, and similarly narrow confidence intervals allow us to reject changes in female employment larger than 1.3 percentage points.[44]

We next test for effects of RGGVY eligibility on the share of households with a variety of different assets and housing stock characteristics. Figure 2.5.8 depicts RD results for the percent of households that own a telephone, own a television, own a motorcycle, have kerosene lighting, have mud floors, and are categorized as "dilapidated" by the Census. We see no strong graphical evidence of discontinuous changes in any of these dependent variables at the 300-person cutoff. Table 3 presents these results numerically in Panels C and D, while also reporting on the share of households with radios, bicycles, and without assets; the share of households cooking with electricity or gas; and the share of households with thatched roofs. Consistent with the graphical evidence in Figure 2.5.8, these results show that RGGVY did not lead to economically meaningful investments in electricity-using assets, non-electricity-using assets, or the housing stock in the medium term. We can reject increases larger than 1 percentage point in all cases. This suggests that RGGVY is unlikely to have contributed to significant increases in household expenditures. The program is also unlikely to have led to meaningful reductions in indoor air pollution, since we see no effects on the share of households with kerosene lighting or electric/gas cooking.

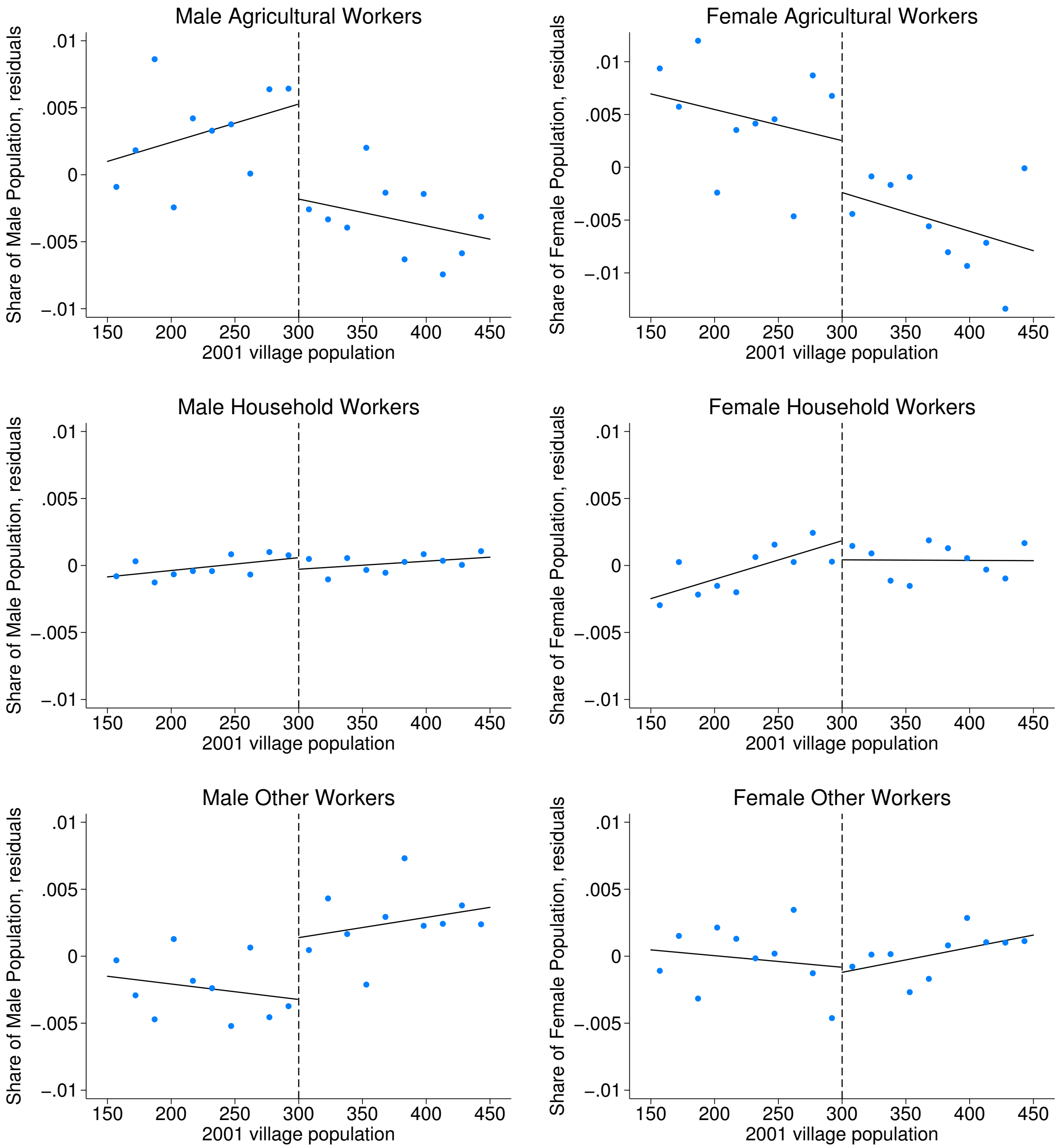
In Panel E of Table 2.5.3, we present RD results for village-level outcomes, including mobile phone coverage, the presence of agricultural credit societies, and the presence of irrigation tubewells, and the share of village area planted and irrigated.[45] These results are not statistically significant, and even the upper bounds on the 95 percent confidence intervals represent economically insignificant changes (smaller than 2 percentage points) in these outcomes. Taken together, these result imply that if RGGVY did lead to increases in agricultural productivity, farmers did not respond by increasing either the scale of irrigation or total farmland.

Next, we test for effects of RGGVY eligibility on economic outcomes among households with at least one poverty indicator. We estimate Equation (2.1) using the fraction of households with at least one poverty indicator (and zero affluence indicators) as the dependent variable. We also test for effects on the fraction of this subset of households for which the main income earner earns at least 5,000 rupees per month. The top row of Figure 2.5.9 presents these results graphically, revealing no evidence that RGGVY led to changes in these outcomes. Panel A of Table 2.5.4 reports the corresponding regression results, along with RD estimates for the fraction of households that report salaried em-

[44]These results focus on the extensive margin of employment (i.e., number of workers). We also test for effects the intensive margin of employment in Appendix B.2.5 (i.e., share of workers working at least six months of the year). We find no evidence of statistically significant or economically meaningful changes on the intensive margin.

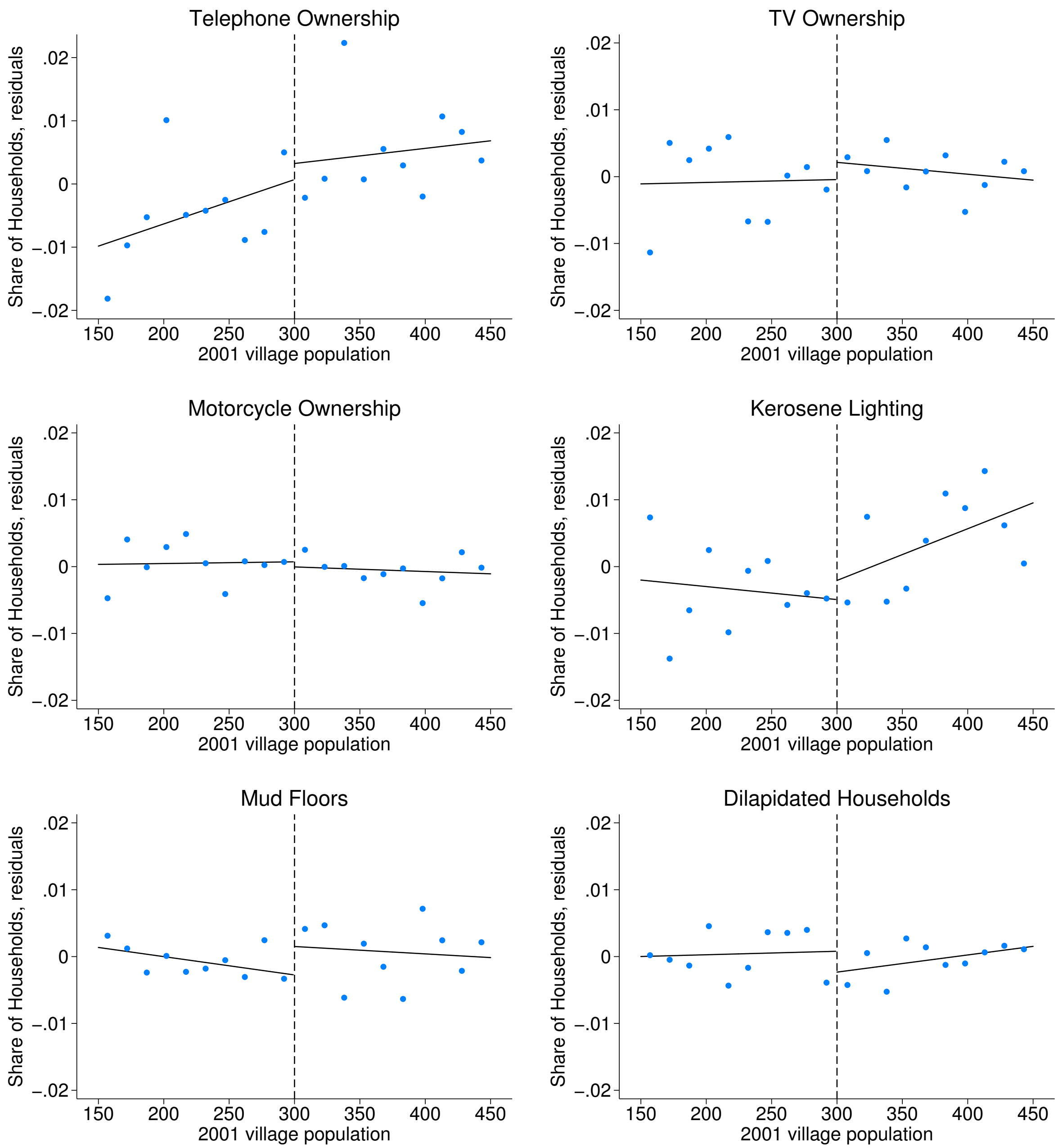
[45]Tubewells are deep wells used for groundwater extraction, which are a common means of irrigation throughout rural India. Electric pumps improve the efficiency of tubewells.

Figure 2.5.7: RD – Labor Outcomes



Notes: This figure shows the results from our preferred RD specification (Equation (2.1)), as reported numerically in Panel B of Table 2.5.3. Blue dots show average residuals from regressing the 2011 percentage of the male/female population classified in each labor category on the corresponding 2001 percentage and state fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Figure 2.5.8: RD – Housing and Asset Ownership



Notes: This figure shows the results from our preferred RD specification (Equation (2.1)), as reported numerically in Panels C and D of Table 2.5.3. Blue dots show average residuals from regressing the 2011 percentage of households owning each asset (or with each characteristic) on the corresponding 2001 percentage and state fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Table 2.5.3: RD – Census Outcomes

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	-0.8647	(2.528)	$[-5.820, 4.091]$	271.09
0–6 cohort / total population	0.0009	(0.001)	$[-0.001, 0.002]$	0.14
Average household size	-0.0051	(0.013)	$[-0.030, 0.020]$	5.13
Literacy rate	-0.0025	(0.002)	$[-0.007, 0.002]$	0.57
B. Labor outcomes				
Male agricultural workers / male pop	-0.0071^{**}	(0.003)	$[-0.013, -0.002]$	0.42
Female agri. workers / female pop	-0.0049	(0.004)	$[-0.013, 0.003]$	0.29
Male household workers / male pop	-0.0009	(0.001)	$[-0.002, 0.000]$	0.01
Female household workers / female pop	-0.0014	(0.001)	$[-0.004, 0.001]$	0.01
Male other workers / male pop	0.0046^{**}	(0.002)	$[0.001, 0.008]$	0.10
Female other workers / female pop	-0.0004	(0.002)	$[-0.004, 0.004]$	0.05
C. Asset ownership				
Share of households with telephone	0.0025	(0.006)	$[-0.008, 0.013]$	0.54
Share of households with TV	0.0026	(0.004)	$[-0.005, 0.010]$	0.26
Share of households with bicycle	-0.0015	(0.004)	$[-0.010, 0.007]$	0.50
Share of households with motorcycle	-0.0008	(0.003)	$[-0.006, 0.004]$	0.13
Share of households without assets	0.0039	(0.004)	$[-0.004, 0.012]$	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	0.0005	(0.003)	$[-0.005, 0.006]$	0.07
Share of households w/ kerosene lighting	0.0029	(0.006)	$[-0.009, 0.015]$	0.48
Share of households with mud floors	0.0043	(0.004)	$[-0.003, 0.012]$	0.73
Share of households with thatched roof	-0.0034	(0.005)	$[-0.013, 0.007]$	0.23
Share of households dilapidated	-0.0031	(0.003)	$[-0.009, 0.002]$	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	-0.0008	(0.011)	$[-0.023, 0.021]$	0.75
1/0 Post office in village	0.0018	(0.004)	$[-0.005, 0.009]$	0.03
1/0 Ag credit societies in village	0.0013	(0.004)	$[-0.006, 0.009]$	0.02
1/0 Water from tubewell in village	-0.0023	(0.011)	$[-0.024, 0.019]$	0.44
Share of village area irrigated	-0.0057	(0.005)	$[-0.016, 0.004]$	0.35
Share of village area planted	0.0015	(0.006)	$[-0.010, 0.013]$	0.58

Notes: Each row represents a separate regression estimating Equation (2.1) on the outcome variable. The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. All specifications also include state fixed effects. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

ployment and that own land. For each outcome, we can reject increases larger than 1.6 percentage points, at 95 percent confidence, suggesting that eligibility for RGGVY did not have economically meaningful effects on household poverty or wealth.

Using the SECC dataset, we also can test whether RGGVY eligibility had different employment impacts among individuals of lower socioeconomic status. We construct sector-specific labor shares that are analogous to Panel B of Table 2.5.3, except that they include only adults living in households with at least one poverty indicator.[46] We report these results graphically in the bottom row of Figure 2.5.9 and numerically in Panel B of Table 2.5.4. While the SECC sample differs notably from the village averages in the PCA, these results are broadly consistent with our main labor results, and visual evidence suggests a small decrease (increase) in agricultural (other) employment for adult men. We can reject 2 percentage point shifts across all six labor categories, which suggests that the average employment effects of RGGVY were similar to the effects on less wealthy households.

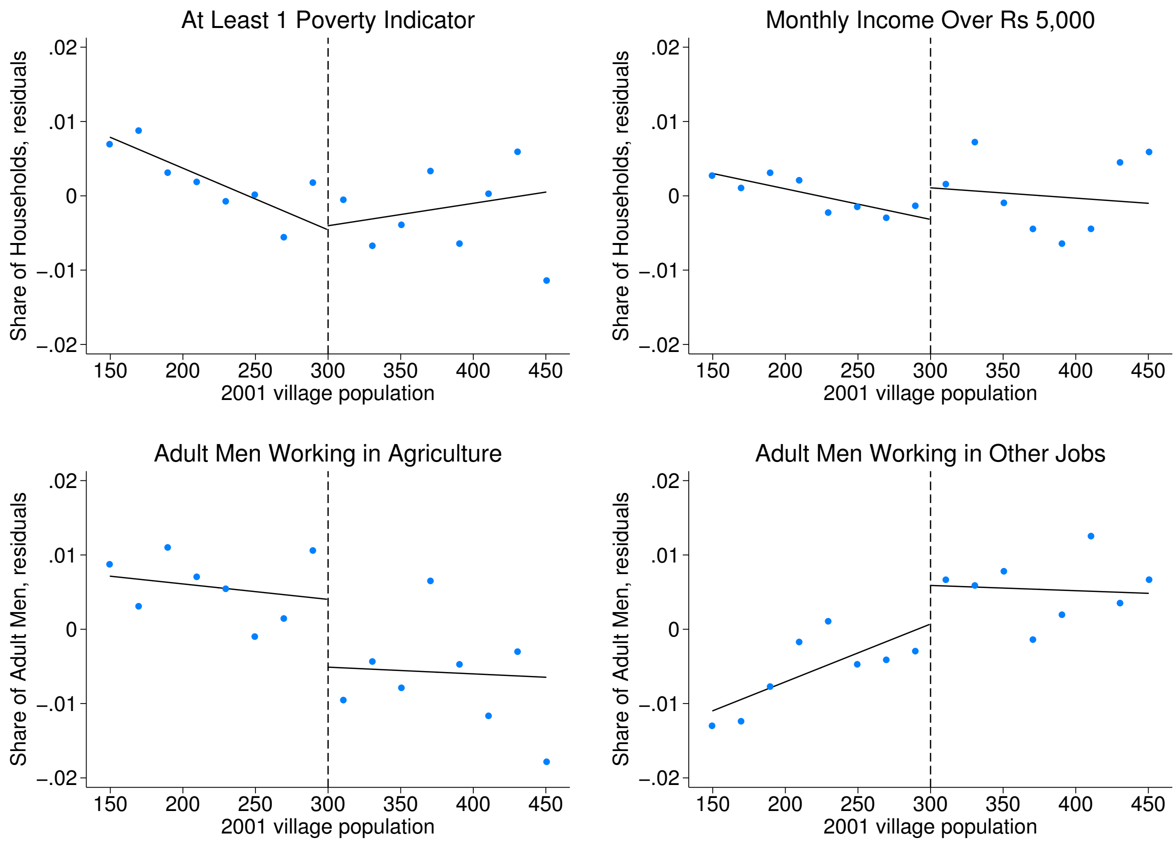
Finally, we test for the effects of RGGVY eligibility on education. We estimate Equation (2.1) using village-wide enrollment for grades 1–8, both pooled and separately by gender, as the dependent variable. We also test for separate effects for primary (grades 1–5) and upper primary (grades 6–8) enrollment, where the latter reflects changes on the intensive margin of schooling. We report these results in Figure 2.5.10 and Table 2.5.5, which show no statistically significant changes in enrollment at the 300-person threshold.[47] As with our other results, our 95 percent confidence intervals can reject even moderate changes in enrollment on either the intensive or extensive margins.

Taking these results together, we conclude that while the provision and consumption of electricity substantially increased as a result of RGGVY eligibility, we detect no economically meaningful changes in labor outcomes, asset ownership, the housing stock, village-level outcomes, household income, or school attendance. Our RD results are precisely estimated, enabling us to rule out even modest effect sizes for these outcomes. This suggests that eligibility for RGGVY did not lead to structural transformation, increased agricultural productivity, female empowerment, reductions in indoor air pollution, improved education, or poverty reductions.

[46] See Appendix B.1.6 for further information on how we constructed these categories from the SECC data.

[47] These regressions control for the 2005 level of the outcome variable, which is the earliest year of enrollment data available. Also, we not that because these village-level enrollment regressions aggregate enrollment across all schools in each village, they might confound changes in within-school attendance with changes in enrollment due to new school construction over time. Appendix B.2.8 repeats these same regressions using school-level enrollment observations, while conducting additional sensitivity analysis.

Figure 2.5.9: RD – SECC Village-Level Outcomes



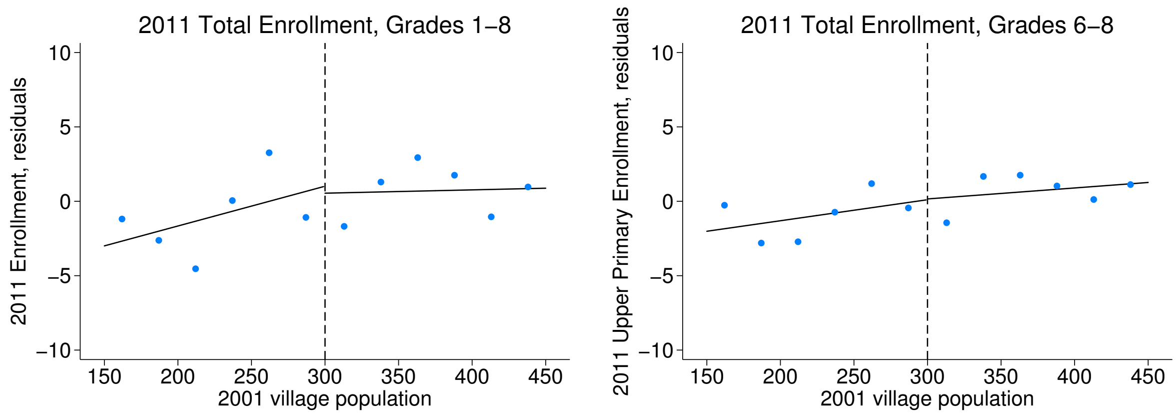
Notes: This figure shows the results from our preferred RD specification (Equation (2.1)), as reported numerically in the first two rows of Table 2.5.4. The upper-left panel reports the proportion of total village households with at least one poverty indicator in 2011, while the upper-right panel reports the proportion of households with a poverty indicator that had a maximum monthly income over Rs 5,000 in 2011. The lower panels report the share of adult men in households with a poverty indicator with occupations in each category. Blue dots show average residuals from regressing the 2011 share of households on state fixed effects. Each dot contains approximately 1,600 villages, averaged in 20-person population bins. Lines are estimated separately on each side of the 300-person threshold for 25,942 villages, i.e. all 10th-Plan single-habitation villages within our 150–450 population RD bandwidth, that match to the SECC dataset.

Table 2.5.4: RD – SECC Village-Level Outcomes

2011 Outcome	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Share of households				
At least one poverty indicator	0.0006	(0.006)	$[-0.011, 0.012]$	0.48
Monthly income > Rs 5,000	0.0043	(0.004)	$[-0.004, 0.013]$	0.08
One member holding salaried job	0.0030	(0.002)	$[-0.002, 0.008]$	0.02
Owning any land	−0.0005	(0.008)	$[-0.017, 0.016]$	0.44
B. Adult employment				
Male agricultural workers / adult men	−0.0091*	(0.005)	$[-0.019, 0.001]$	0.29
Female agri. / adult women	−0.0039	(0.005)	$[-0.013, 0.006]$	0.08
Male household workers / adult men	0.0008	(0.001)	$[-0.002, 0.004]$	0.01
Female household workers / adult women	−0.0015	(0.008)	$[-0.016, 0.013]$	0.51
Male other workers / adult men	0.0052	(0.006)	$[-0.007, 0.017]$	0.42
Female other workers / adult women	0.0054	(0.005)	$[-0.005, 0.016]$	0.16

Notes: Each row represents a separate regression estimating Equation (2.1) on a different SECC village-level outcome. The first row of Panel A is coded as the share of *total* households in the village with at least one poverty indicator. Other outcomes in Panel A are coded as the proportion of this *subset* of households (with poverty indicators) that meet each criterion. Panel B outcomes are coded as the share of adult men (women) with an occupation in each subcategory, for the sample of adults in households with at least one poverty indicator. (We treat all invididuals over 16 years of age as adults.) The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications include state fixed effects, but they do not include any additional baseline control variables. The RD bandwidth includes 25,942 villages with 2001 populations between 150 and 450. These regressions contain fewer villages than regressions in Table 2.5.3 because only 87 percent of 10th-Plan, single-habitation, 150–450 villages match to the SECC dataset. Standard errors are clustered at the district level with 222 clusters, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 2.5.10: RD – School Enrollment



Notes: This figure shows the results from our preferred RD specification (Equation (2.1)), as reported numerically in the first and last rows of Table 2.5.5. Blue dots show average residuals from regressing the 2011 number of (total, grades 6–8 only) students on the corresponding 2005 enrollment counts and state fixed effects. Each dot contains approximately 1,000 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for 12,251 single-habitation villages between 150 and 450 people, in 10th-Plan districts, with school-village matches and nonmissing 2005 and 2011 enrollment data.

Table 2.5.5: RD – School Enrollment

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
Total enrollment, grades 1–8	−0.472	(3.93)	[−8.18, 7.24]	74.05
Male enrollment, grades 1–8	0.197	(2.00)	[−3.72, 4.11]	37.60
Female enrollment, grades 1–8	−0.650	(2.02)	[−4.61, 3.31]	36.45
Total enrollment, grades 1–5	−0.408	(2.95)	[−6.19, 5.37]	60.58
Total enrollment, grades 6–8	0.051	(1.50)	[−2.89, 2.99]	13.47

Notes: Each row represents a separate regression estimating Equation (2.1) on a different enrollment count, aggregating schools enrollment up to village-level observations. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2005 level of the outcome variable and state fixed effects. The RD bandwidth includes 12,251 village observations with 2001 populations between 150 and 450, with a single habitation, in RGGVY 10th-Plan districts. These regressions contain fewer villages than regressions in Table 2.5.3 because only 51 percent of 10th-Plan, single-habitation, 150–450 person villages match to a school, and only 76 percent of these matched villages contain schools that report nonmissing enrollment values for 2011 and 2005. Standard errors are clustered at the district level, with 215 clusters, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2.6 Interpretations and Extensions

2.6.1 Scaling

The above regressions recover intent-to-treat estimates: they show the effect of being *eligible* for RGGVY on our outcomes of interest. In order to compute average treatment effects, we need to scale these estimates such that we recover the effect of electrification on development.[48]

We propose several methods of scaling our estimates. First, we consider inflating our outcomes based on the proportion of villages within our bandwidth that RGGVY claims to have treated. This is akin to the scale factor we would apply with a traditional instrumental variables estimator. RGGVY's district-level aggregate data suggest that between 56 and 82 percent of eligible villages were treated by the program.[49] This implies that our estimates should be inflated by approximately a factor of 1.5 in order to recover the causal effects of treatment under RGGVY.

Alternatively, we can calibrate a scaling factor to the magnitude of the increase in nighttime brightness, which we estimate to be 0.15 units of brightness. Min et al. (2013) suggest that when villages in Senegal were electrified, they experienced increases of approximately 0.4 nighttime brightness points. If, alternatively, we apply Min and Gaba (2014)'s estimates of a 1-unit increase in brightness corresponding to 240–270 electrified households, then full electrification of the average village in our RD sample with 56 households would imply an increase of 0.2 brightness points.[50] This suggests that our RD estimates should be inflated by a factor of between 1.3 and 3 to recover the average effect of RGGVY electrification.[51]

Scaling the point estimates reported in Tables 2.5.3–2.5.5 by a factor of 3 does not yield adjusted estimates that are economically meaningful. For the vast majority of outcomes, we see no visual evidence of a discontinuity, suggesting that these upper bounds are quite conservative. Even after inflating the 95 percent confidence intervals by these factors, we can still reject 4 percentage point changes in labor outcomes, 4 percentage point changes in asset ownership, 5 percentage point changes in the housing stock, and 8 percentage point changes in village-level outcomes. We can also reject 6 percentage point changes in outcomes in Table 2.5.4, as well as 21 student (30 percent) increases in total school enrollment. Scaling by a factor of 3, we can rule out effects larger than 0.26 of one standard deviation in all outcomes presented in Tables 2.5.3–2.5.5.

[48] We do not scale via two-stage least squares because we do not have access to a binary "RGGVY electrification" variable, nor would this variable capture different levels of energy access and consumption across villages treated under RGGVY, as discussed above.

[49] RGGVY's aggregate village counts in 10th-Plan districts sum to 56 percent of the total number of villages in these districts, and 82 percent of villages with 2001 populations over 300.

[50] These increases of 0.4 and 0.2 are internally consistent; the average villages in Min et al. (2013) and Min and Gaba (2014) are larger than the villages in our RD bandwidth.

[51] We do not propose a scale factor based on Min and Gaba (2014)'s streetlights estimate, since we do not have data on the number of streetlights per village, and because RGGVY did not install streetlights.

Even if we were to scale our estimates by an extremely conservative factor of 10, we can still reject effect sizes consistent with the previous literature.[52] Dinkelman (2011) finds that electrification caused 9–9.5 percentage point increases in female employment; we can reject 2 percentage point increases in total female employment.[53] Lipscomb, Mobarak, and Barham (2013) likewise find large effects of electrification on total employment rates; we can reject 1 percentage point increases in the village-wide employment rate even after applying a conservative scaling factor of 10.[54] Chakravorty, Emerick, and Ravago (2016) find that rural electrification leads to a 56 percent decrease in a deprivation index, and a 38 percent increase in household expenditures; scaling our Table 2.5.4 results by 10, we can reject an 11 percentage point decrease in the share of households with at least one poverty indicator, and a 13 percentage point increase in the share of households (with at least one poverty indicator) with monthly incomes greater than 5,000 rupees.

2.6.2 Heterogeneous Effects

It is possible that our results mask heterogeneity in the quality of energy services experienced by RGGVY villages. In particular, India faces major electricity shortages, which vary across locations (Allcott, Collard-Wexler, and O'Connell (2016)). If half of the villages in our sample experienced frequent power outages while the other half received consistent power, our average intent-to-treat estimate across both groups would be small even if RGGVY led to large economic effects in places with high-quality energy supply. We test for this by re-estimating all of our RD results using the subset of states with above-average power availability (Central Electricity Authority (2011)).[55] In this subsample, our estimated RD coefficient on nighttime brightness increases from 0.15 to 0.25, statistically significant at the 1 percent level. However, the results for labor, asset ownership, the housing stock, village-level outcomes, and household wealth are quantitatively similar to those estimated using the full RD sample.[56]

This suggests that poor power quality in a subset of states is not attenuating our estimate of the average effect across the full sample. Moreover, our main RD results

[52] In order to arrive at factor of 10, which we believe to be the most conservative interpretation of our results, we assume that RGGVY *only* impacted household electricity end-uses. Our nighttime brightness effect of 0.15 is comparable to the change in brightness associated with a 10 percentage point increase in the share of households with electric lighting, a proxy for household power consumption, at the mean of our RD sample. This suggests a scaling factor of 10 to translate this into an increase from 0 to 100 percent of households.

[53] We estimate Equation (1) pooling female employment across all three sectors, resulting in an RD point estimate of -0.0067 with the upper end of our 95 percent confidence interval of 0.0015, which we multiply by 10. We can similarly reject increases of 3 percentage points in female agricultural employment, 1 percentage point in female household employment, and 4 percentage points in female other employment.

[54] If we pool all six labor outcomes in Panel B of Table 2.5.3, the resulting RD point estimate is -0.0053 with an upper 95 percent confidence interval of 0.0002.

[55] These seven states are (in decreasing order of 2011 power quality): Chhattisgarh, Orissa, Karnataka, West Bengal, Gujarat, Haryana, and Rajasthan.

[56] Appendix B.2.10 reports regression results for both split-sample exercises discussed in this section. The schooling results are qualitatively similar, but somewhat less robust.

reflect the realized implementation of a large-scale national rural electrification program in the developing world. Even if we had found substantial positive effects for a subset of states, the overall treatment effect would be indicative of the degree to which future rural electrification programs might be limited by the supply reliability.

It is also possible that we do not detect large effects because the benefits of electrification take many years to accrue. While we cannot rule this possibility out completely, our 2011 outcome data were collected between three and five years after 95 percent of villages in our sample received RGGVY funding.[57] Even if there were significant delays in implementation, this is much longer than the time span over which development interventions are typically studied. Nevertheless, we recover quantitatively similar RD point estimates when we restrict our RD sample to districts with early RGGVY funding. Therefore, it is unlikely that our small results are driven by villages that failed to take advantage of the full set of possible medium-run benefits of electric power before being surveyed by the 2011 Census.

2.6.3 Difference-in-differences

Finally, we might be concerned that villages close to the 300-person RD threshold stand little to gain from electrification. Perhaps these small villages are simply too poor, too credit-constrained, or too economically isolated to translate increased electricity access into new employment or income-generating opportunities. We employ a second identification strategy, difference-in-differences (DD), to test for the effects of RGGVY eligibility on larger villages far from our RD threshold. Recall that there were two major phases of RGGVY implementation: the 10th-Plan phase and the 11th-Plan phase. The majority of 11th-Plan electrification projects had not been completed before the 2011 Census. We can therefore use 10th-Plan districts as a "treated" group and 11th-Plan districts as a "control" group in a DD framework.[58] We estimate the following fixed effects specification on our two-decade village panel:

$$Y_{vst} = \gamma_0 + \sum_b \gamma_1^b \mathbf{1}\left[\text{10th} \times \text{Post}\right]_{vt} \times \mathbf{1}\left[P_v \in \text{Bin}_b\right] + \delta_t + \eta_v + \varepsilon_{vt} \tag{2.2}$$

where $\mathbf{1}\left[\text{10th} \times \text{Post}\right]_{vt}$ is an indicator equal to one if village v was eligible for RGGVY under the 10th Plan and the year t is 2011, $\mathbf{1}\left[P_v \in \text{Bin}_b\right]$ are 2001 village population bins along the full support of populations (shown in Figure 2.4.3), δ_t are year fixed effects, and η_v are village fixed effects. This necessitates stronger identifying assumptions than our RD specification, namely that villages in 10th-Plan districts were trending in parallel to 11th-Plan villages prior to RGGVY. Village-level data are not available for the 1991 Census, therefore we are unable to directly test this assumption.[59]

[57] Over 70 percent of villages in our RD sample are in districts that received RGGVY funding before the end of 2006. See Appendix Table B.2.26.

[58] Selection into the different plans was non-random. It is plausible that 10th-Plan districts were more administratively capable than 11th-Plan districts, likely biasing our DD estimates upward.

[59] In Appendix B.2.11, we test for differential pre-trends using district-level data. These trends are not statistically zero, suggesting that our DD results should be interpreted with some caution.

Figure 2.6.11 compares our main RD results with DD results from estimating Equation (2.2) with 300-person population bins, for nighttime brightness and male agricultural workers. For both outcomes, the RD point estimates lie within the DD confidence intervals. Moreover, the DD effect of RGGVY on nighttime lights increases nearly monotonically in population, while the DD effect for male agricultural labor is close to constant as population increases. This suggests that our small RD results for male agricultural employment are likely to be externally valid outside of our RD bandwidth. Other economic outcome variables show similarly constant DD coefficients across small and large villages.[60]

These DD results are broadly consistent with our RD results, despite using a much larger population of villages (from 10th- and 11th-Plan districts, including multi-habitation villages) and using 11th-Plan villages as counterfactuals (as opposed to barely ineligible 10th-Plan villages). Beyond allowing us to extend our RD results to larger villages, the DD results are encouragingly similar to the RD. Relying on alternative identifying assumptions on a different sample of villages, we again demonstrate that RGGVY caused nighttime brightness to increase, but has not meaningfully improved the economic outcomes that we observe.

2.6.4 Costs and Benefits

We do not have direct estimates of village-level program costs, incomes, or expenditures.[61] However, we perform several back-of-the-envelope calculations based on our RD results. This enables us to better understand the overall economics of RGGVY, while also quantifying the costs and benefits of electrification.

First, we consider the per-village costs of RGGVY implementation. In 2005, RGGVY was expected to cost 634.2 billion rupees, or approximately \$17.2 billion.[62] Given the stated scope of the program detailed in Section 2.2, this suggests a cost per village of approximately 1,470,000 rupees, or \$36,000 in 2015 USD.[63]

We can apply average Indian rural wage rates to estimate the income differential that might have resulted from the (small) sectoral shift from agricultural to non-agricultural employment we observe under RGGVY. According to India's National Sample Survey Office, the average 2011 wage for male (female) non-agricultural workers was 196 (116) rupees per day, which was 26 (0.9) percent higher than the average agricultural wage of 155 (115) rupees per day. To compute the average increase in village-level income, we scale the lower bound of our confidence interval for male (female) agricultural labor from Table 2.5.3, -0.013 (-0.013), by a factor of 3. This converts our intent-to-treat estimate

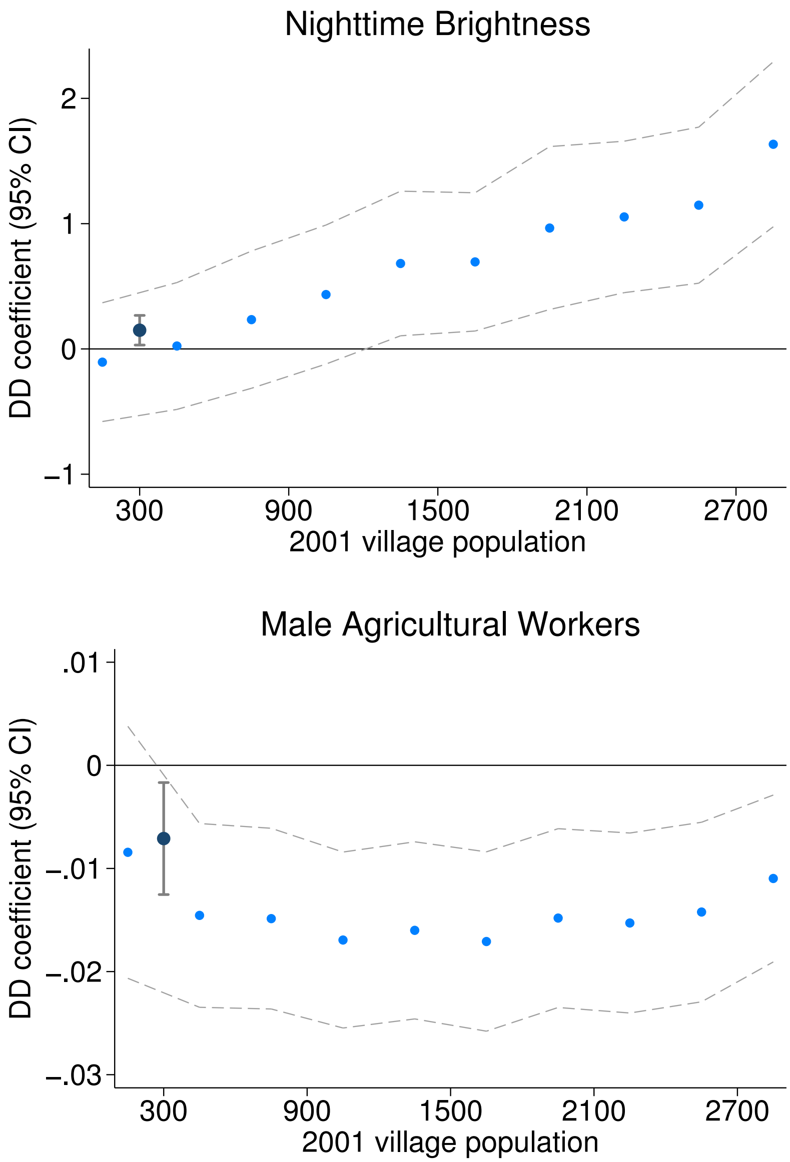
[60] We report additional DD results in Appendix B.2.11.

[61] The SECC income data indicate whether households' main income earners earned more or less than 5,000 rupees per month, and comes from a selected subset of households. Hence, we exclude these data from the subsequent cost-benefit analysis.

[62] We use the 2005 exchange rate of 44 rupees per dollar, and convert to 2015 USD.

[63] This is comparable to Chakravorty, Emerick, and Ravago (2016), who report average electrification costs of \$42,000 per village in the Philippines.

Figure 2.6.11: Difference-in-Differences Results



Notes: This figure compares the reduced form effects from our preferred RD specification (Equation (2.1)) to the results from our DD specification (Equation (2.2)), using 300-person population bins. Navy blue dots show the RD coefficients, with whiskers indicating 95 percent confidence intervals. Light blue dots and dashed lines show the binned DD point estimates and 95 percent confidence intervals. The top panel shows the effects for nighttime lights, as measured by maximum village brightness. The bottom panel shows the effects for male agricultural workers. The RD results are statistically significant at the 5 percent level and the 1 percent level, respectively. The pooled DD point estimates are 0.45 and −0.008; both are statistically significant at the 10 percent level (Appendix B.2.11 reports these results in a regression table). DD regressions for lights and labor include 629,778 and 994,802 village-year observations, respectively.

into an average treatment effect, and it implies a maximum shift out of male and female agricultural employment of 3.9 percentage points.[64] If all of these men (women) shifted from agriculture into non-agriculture employment, then total daily male (female) village wage earnings would have increased by approximately 293 (7) rupees. If each employed person worked 365 days per year, this would translate into a total annual village income increase of approximately 109,000 rupees, or an upper bound of 1.4 percent.

Alternatively, we can use our RD estimates on household asset ownership to infer changes in expenditures resulting from electrification. Scaling the upper confidence intervals in Panel C of Table 2.5.3 by 3, we can reject increases in asset ownership of greater than 3.9 percent for mobile phones, 3.0 percent for televisions, 2.1 percent for bicycles, and 1.2 percent for motorcycles. Monetizing these upper bounds using asset prices from ICRISAT's Village Dynamics in South Asia dataset, this implies a maximum average household expenditure of 572 rupees.[65] Supposing that only 10 percent of RGGVY-driven expenditure increases were spent on these four durable goods implies a maximum increase in per-household expenditure of 5,720 rupees, or a total village-wide increase of around 398,000 rupees. These asset purchases occurred during the 3–6 year period after electrification; if we conservatively assume that they all occurred within 3 years of electrification, this would represent at best a 2.1 percent increase in annual village expenditures.[66]

Our back-of-the-envelope estimates suggest that annual village income increased by a maximum of 109,000 rupees, that annual village expenditures increased by a maximum of 133,000 rupees, and that RGGVY electrification came at a cost of approximately 1,470,000 rupees per village. These results are quite conservative: though we do not measure all possible benefits from electrification, the benefits we do use in performing this calculation come almost entirely from regression estimates where we cannot reject zero; our assumptions in performing this calculation also make it biased towards finding large effects. Using the larger expenditure estimate and applying a conservative 3 percent discount rate, this translates into a payback period of approximately 12 years.[67]

At best, we find that RGGVY increased annual incomes by 1.4 percent and annual expenditure by 2.1 percent, despite causing a substantial shift in nighttime lights. This suggests exercising caution when using nighttime brightness as a proxy for income or expenditures. The DMSP-OLS dataset measures light emissions. Because brightness relates directly to energy consumption through lighting, it serves as a useful indicator of electrification. Since electrification should lead to increased brightness even absent a corresponding increase in incomes, we do not use the DMSP-OLS data as a proxy for

[64] In keeping with Section 2.6.1, we apply a scaling factor of 3 throughout this section.

[65] The average prices for durables commonly purchased after electrification are: Rs 2,796 for cell phones, Rs 4,166 for televisions, Rs 1,259 for bicycles, and Rs 25,922 for motorcycles.

[66] India's average rural monthly per capita expenditures were 1,430 rupees for 2011–2012.

[67] This starkly contrasts with Chakravorty, Emerick, and Ravago (2016), who find a payback period of approximately 1 year; however, it is corroborated by evidence from Lee, Miguel, and Wolfram (2016), who use revealed preference results to suggest that the costs of electrification are much larger than the benefits.

income/expenditures, and caution others against doing so when evaluating programs that directly increase light emissions.

Importantly, our results do not speak directly to the effects of RGGVY on welfare. It is quite possible that electrification has dramatically increased average quality of life for rural Indians. Indeed, since villagers are using more power as a result of RGGVY, revealed preference suggests that they benefit from the program. Even though we measure a wide range of outcome variables which are typical of large-scale administrative datasets, there may be important utility benefits that we cannot measure. Our results highlight the need to incorporate additional non-market measures into future administrative data collection efforts.

2.7 Conclusion

In this paper, we evaluate the medium-run effects of electrification on development using a regression discontinuity (RD) design which exploits a population eligibility threshold in India's national rural electrification program, RGGVY. We find that eligibility for RGGVY led to substantial changes in nighttime brightness and power availability. Despite this increase in energy access, we find that electrification did not have economically meaningful impacts on a range of development outcomes.

These results hold when we rescale our reduced form estimates to account for the proportion of eligible villages that underwent treatment. We see similar effects on development among states with high and low average reliability of electricity supply. We also find similar effects when we restrict our analysis to the earliest districts to obtain RGGVY funding, suggesting that our results do not depend on the timing of our post-intervention data. Finally, we apply a difference-in-differences strategy, which relies on alternative identifying assumptions and includes a larger sample of villages well outside our RD bandwidth. These results support the main conclusions from our RD analysis that while nighttime lights, and therefore power consumption, increased substantially with RGGVY electrification, other development outcomes that we observe did not. Our cost-benefit calculations suggests a much longer payback period than previously estimated.

These results are the first to suggest that electrifying rural villages may not cause sizable economic gains in the medium term. Our regression discontinuity strategy relies on much less stringent identifying assumptions than the instrumental variables approaches of previous work, allowing us to measure effects of a natural rollout of rural electrification, at scale. In contrast to the existing literature, we find that electrification did not yield even modest changes in labor, income, household wealth, asset ownership and expenditures, village-level outcomes, and education. These null results come from the world's largest unelectrified population, and appear to generalize to over 400,000 villages across rural India.

Nevertheless, electrification may lead to large economic benefits in certain contexts, and may have important positive effects on human well-being that we are unable to quantify. An important direction for future work will be to understand when, where, and after how long electricity access and power availability have the greatest economic impact. For example, electrification may lead to substantial gains in economic productivity in urban settings, or in regions with budding local industries. There may also be substantial long-run effects of electrification, and more research is necessary to identify these benefits. Finally, we encourage future research on quantifying the non-market benefits from electrification that frequently go unmeasured.

Chapter 3

Costs of Misallocation in Indian Electricity Supply[1]

3.1 Introduction

Electricity is an essential input to modern economic activity, and energy consumption is highly correlated with GDP in the global cross-section. Electricity demand in the developing world is projected to rise dramatically over the coming decades, as households move out of poverty and purchase electric appliances (Wolfram, Shelef, and Gertler (2012); Gertler et al. (2016)). Governments and development agencies invest billions of dollars annually to expand access to cheap, reliable electricity in low- and middle-income countries (Burlig and Preonas (2016)). Meeting this key development goal will require these countries to develop well-functioning electricity supply sectors.

At the same time, electricity supply is notoriously complicated due to the physical requirements of producing and transporting electric power. Electricity generation has sharply increasing returns to scale, and power plants incur extremely high upfront investment costs relative to the marginal value of electricity they produce. Because electricity is extremely expensive to store and supply must instantaneously meet demand for markets to clear, regulators must allow generators to earn sufficient revenue to finance high fixed costs while still incentivizing them to supply cheap, reliable power to consumers.[2] Traditionally, regulators have allowed electric utilities to operate as vertically integrated

[1]This chapter is coauthored with Fiona Burlig and Akshaya Jha. We thank Severin Borenstein, Steve Cicala, Michael Greenstone, Ryan Kellogg, Koichiro Ito, Nick Ryan, Matt Woerman, Catherine Wolfram, and seminar participants at the University of Chicago, Columbia SIPA, Camp Resources, the Heartland Environmental and Resource Economics Workshop, and the UC Berkeley Economics Department for helpful comments and suggestions. Erin Kelley provided invaluable data acquisition support, and Jessica Jiang and Xiner Xu served as excellent research assistants. All remaining errors are our own.

[2]This balancing act—simultaneously providing low cost electricity to consumers while still incentivizing investment in electricity generating capacity—is known as the "missing money" problem (see Joskow (2006); Joskow (2008)).

natural monopolists, while restricting electricity prices to be just high enough for utilities to recover both fixed and variable costs of electricity generation and transmission.

In developing countries, electric utilities have historically tended to be state-owned monopolies responsible for the generation, transmission, and distribution of electricity. These government-run utilities have typically struggled to provide reliable power, often facing frequent supply shortages and blackouts while operating on the edge of bankruptcy. In response, more than 70 low- and middle-income countries have implemented electricity market reforms since the early 1990s. A typical reform package has sought to establish a competitive wholesale market for electricity, while encouraging private entry of electric power plants and unbundling the generation of electricity from its transmission and distribution. Despite the importance of the power sector for economic growth, there remains little empirical evidence on the economic efficiency of electricity markets in the developing world (Jamasb, Nepal, and Timilsina (2015)).

The U.S. electricity sector underwent similar reforms in the 1990s, and the transition towards wholesale electricity markets yielded large welfare gains (e.g., Fabrizio, Rose, and Wolfram (2007); Davis and Wolfram (2012); Cicala (2017)). However, the California electricity crisis of 2000–2001 serves as a stark reminder that such reforms can also create substantial welfare losses absent sufficient regulatory oversight (Borenstein, Bushnell, and Wolak (2002); Borenstein (2002)). In low- and middle-income countries, relatively weaker institutions and limited regulatory capacity have the potential to undermine the welfare gains from electricity market reforms. As these countries continue to rapidly expand electricity supply, understanding the drivers of economic inefficiencies in electricity markets is critical to supporting economic growth in the developing world.

In this paper, we quantify the costs of short-run misallocation in Indian electricity supply. We begin by assembling a novel dataset on electricity production and marginal costs of generation for each utility-scale power plant in the country, between 2013 and 2017. We use these data to estimate short-run supply misallocation by creating a least-cost counterfactual. That is, we "dispatch" power plants in order of lowest to highest cost, until there is enough aggregate supply to meet aggregate demand and clear the market. We compute the total short-run cost under this "least-cost" counterfactual, and we also calculate the total observed short-run costs of electricity generation based on each plant's actual production. By comparing the costs calculated under least-cost versus observed dispatch scenarios, we are able to quantify the cost gap between factual and counterfactual electricity supply. After accounting for transmission constraints, the remaining cost gap represents the short-run misallocation wedge (Hsieh and Klenow (2009)).

Using administrative data on marginal costs, we find that least-cost dispatch has total short-run variable costs of approximately \$24 billion per year, while the observed dispatch has costs of approximately \$29 billion per year—more than \$4.6 billion (16 percent) higher. We also construct our own measure of marginal costs using detailed information on plant efficiency and input prices; using these constructed cost data, we likewise find that observed dispatch is over 17 percent more costly in the short run than least-cost dispatch. In order to conservatively account for both interregional and intraregional transmission constraints, we impose autarky within each subregion of the Indian electricity grid. The remaining cost gap is $3.2 billion per year (11 percent), which represents a lower bound on the short-run misallocation wedge.

We investigate several potential drivers of this misallocation wedge: market power, political economy, and market design. We find that market power plays at most a small role in Indian electricity supply, consistent with the fact that the majority of plants face rate-of-return regulation. We do find evidence that political factors impact misallocation, using a regression discontinuity design around close election outcomes. Finally, we argue that market design is an important driver of misallocation in the Indian market.

This paper makes three main contributions to the literature on electricity supply in developing countries.[3] First, we assemble a novel dataset on market operations and marginal costs of electricity supply in India. Armed with these data, we provide among the first estimates of the short-run cost of electricity supply in a developing country. Second, we extend the U.S.-focused literature on deviations from first-best outcomes in electricity markets, by quantifying the extent of misallocation in the second-largest electricity market in the developing world. Finally, we contribute to the broader literature on electricity markets by studying the mechanisms that generate misallocation in power supply.

This paper proceeds as follows. Section 3.2 describes the Indian electricity market. Section 3.3 presents our data. Section 3.4 outlines our empirical approach to estimating costs of generation in India, and Section 3.5 displays the results of this exercise. In Section 3.6, we provide a discussion of mechanisms behind the deviation between actual market operations and short-run cost minimization. Section 3.7 concludes.

3.2 The Indian Electricity Market

In this section, we provide an overview of the major features of the Indian electricity market: power market reforms from 2003 to today, today's generation mix, long-term contracts, and transmission.

Power market reforms The majority of electricity generation capacity in India is owned by the central or state governments. In 1948, the Electricity Supply Act gave rise to State Electricity Boards (SEBs), responsible for electricity generation, transmission, and distribution, as well as tariff-setting. The tariffs set by these SEBs were too low to recover costs, leading to supply shortages, a lack of investment in generation capacity, and bankruptcy among state electricity companies. In response, the Indian electricity sector was charged with reform, beginning with the introduction of State Electricity Regulatory Commissions, starting in 1996.

[3]There is a rapidly growing literature on electricity in the developing world. See for example Lee, Miguel, and Wolfram (2016), Burlig and Preonas (2016), and Dinkelman (2011) on rural electrification; Allcott, Collard-Wexler, and O'Connell (2016) and Abeberese (Forthcoming) on the impacts of power shortages and electricity prices on firm performance; and Ryan (2017) on the role of transmission constraints in Indian electricity supply, among many others.

The most significant reform in recent history has been the 2003 Electricity Act, which called for an overhaul of the Indian power market. The stated goal of the Act was to facilitate competition in the supply of electricity through many different changes to the industry, including opening up access to electricity transmission grids as well as electricity distribution as well as removing the licensing requirement for the generation and distribution of electricity. In the wake of the Act, private investment in generation increased dramatically. Today, the Indian electricity sector includes over 330 GW of capacity, roughly 56 percent of which is government owned.

Fuel The majority of electricity generation in India comes from coal-fired generation sources. As of 2017, coal makes up 58.4 percent of this capacity, gas makes up 7.6 percent of this capacity, hydro makes up 13.5 percent of this capacity, nuclear makes up 2.1 percent of this capacity, and renewables makes up 18.2 percent of this capacity. Most coal- and gas-fired power plants procure input fuels via long-term Fuel Supply Agreements (FSAs), which are governed by fuel price regulations. Typically, fossil fuel plants are required to sign an FSA before they are allowed to begin operations. The majority of coal-fired power plants operate via linkages with Coal India, Ltd., the government-owned coal supply monopoly.[4] Recent shortages in coal and natural gas supply have been anecdotally linked with rolling blackouts and intermittent power supply.

Long-term contracts In spite of recent market reforms, the short-term power market comprised only 10 percent of the total electricity procured in India in 2016–17. The vast majority of electricity in India is sold via medium- or long-term contracts between generation companies and wholesale consumers (e.g. distribution companies). The Central Electricity Regulatory Commission (CERC) approves the electricity tariffs received by generators on these bilateral contracts; this approval process is based on each plant's fixed and variable costs.

Transmission As with any electricity market, transmission plays an important role in the Indian electricity sector. India is split into five electricity transmission grids, with inter-regional transmission lines between these five regions. The total interregional transmission capacity has increased rapidly in recent years, from 14,050 MW in 2007 to 75,050 MW in 2017. Expansions in transmission, including integrating all of India's transmission regions into one grid under the One Nation, One Grid scheme has been a major goal of the current government.

Transmission capacity is allocated by the National Load Despatch Centre in conjunction with regional and state Load Despatch Centres, using an administrative process that prioritizes long-term contracts, medium-term contracts, and finally the short-term electricity market (see Ryan (2017) for more details). Importantly, by prioritizing the

[4] After our sample period, the government introduced the "Scheme to Harness and Allocate Kolya (Coal) Transparently in India" (a.k.a. Shakti) policy allocates *new* coal contracts to generation units based on an auction mechanism; the first auction ran in September 2017.

transmission rights of long-term contracts, this limits the ability of generators to arbitrage differences between contract prices and short-term trading prices. In other words, a long-term contract serves as both a forward financial commitment between traders *and* a de facto commitment to a physical allocation of electricity generation.

In this paper, we seek to estimate the costs of short-run misallocation in electricity supply in India. We estimate both the total costs of observed electricity supply and the costs associated with a least-cost counterfactual, in which we re-dispatch units according to marginal cost. Importantly, we hold demand, transmission, costs, and generating capacity fixed, such that our estimates reflect consequences of short run reallocation only.

3.3 Data

In order to quantify the costs of misallocation in power generation in India, we digitized and assembled a novel dataset on the Indian electricity supply sector. Our empirical analysis draws on four main types of data. First, we use detailed daily plant-level generation and operation data, reported for the majority of India's utility-scale power plants by the Central Electricity Authority (CEA). Second, we use data on plants' reported operating costs from India's Ministry of Power. Third, we construct our own measure of power plant operating costs based on plant-specific heat rates and fuel consumption published by the CEA (a plant's heat rate is the amount of thermal energy in kcal required to produce one MWh of electricity). Fourth, we incorporate detailed data on fuel prices to convert heat rates into marginal costs. These include fuel prices from the Ministry of Coal and Ministry of Petroleum and Natural Gas, hand-geocoded plant coordinates, and geospatial data on India's coalfields from the USGS—which we combine to calculate the shortest distance between each coal-fired power plant and the nearest coalfield.

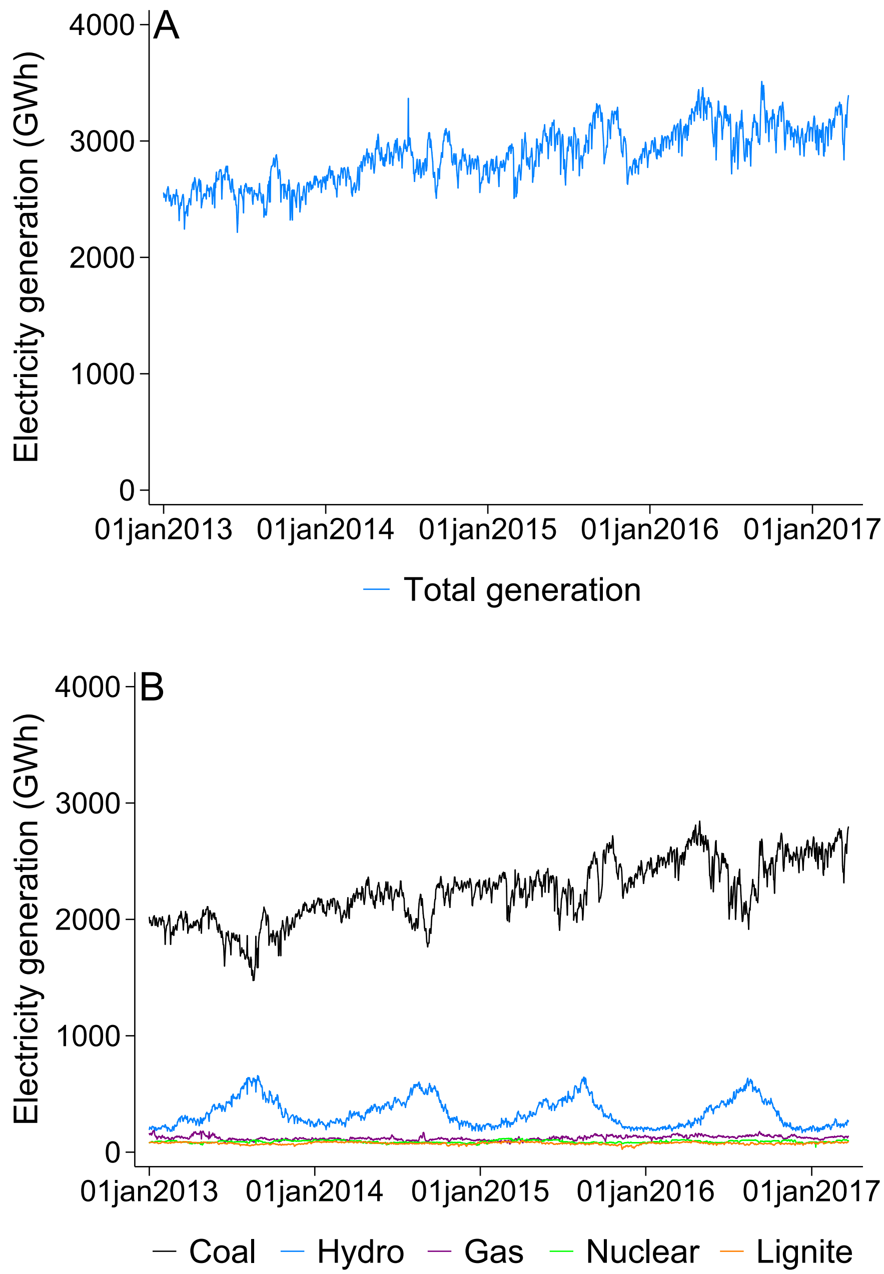
Daily generation data The Central Electricity Authority (CEA) monitors the operations of all utility-scale (greater than 50 MW) fossil, hydroelectric, and nuclear power plants in India.[5] This include centrally owned, state-owned, and privately owned facilities. We obtain the CEA's Daily Generation Reports from January 1, 2013 (the earliest date available) to March 23, 2017.[6] Each report enumerates operational capacity, scheduled generation, and actual generation for each plant in each day-of-sample. The 485 plants in the CEA data make up 291 GW of India's 330 GW of electric generating capacity, with an average total generation of approximately 2.9 TWh per day (i.e. 2,900 GWh).

Panel A of Figure 3.3.1 shows daily total generation summed over the plants in our sample while Panel B breaks this generation down by source. The vast majority of generation in the CEA data comes from the 196 coal-fired power plants, which average 2.2 TWh

[5] Wind and solar resources instead fall under the Ministry of Renewable Energy. To the best of our knowledge, there exists no publicly available daily generation data from renewable sources. Non-hydro renewables account for less than 20 percent of India's electricity generation capacity. These technologies are non dispatchable and have extremely low marginal cost.

[6] We expect our final analysis period to range from January 1, 2013 to December 31, 2017.

Figure 3.3.1: Daily Electricity Generation in India



Notes: This figure displays data from the Central Electricity Authority's Daily Generation Reports. Panel A shows total electricity generation in GWh from January 1, 2013 through March 23, 2017. The 485 plants in these data average 2.9 TWh per day. Panel B breaks these plants out by fuel type. The 196 coal-fired power plants generate by far the most electricity, averaging 2.2 TWh per day. The remainder of generation comes from 193 hydroelectric plants (336 GWh per day, but highly seasonal); 62 natural gas plants (121 GWh per day); 7 nuclear plants (90 GWh per day); 9 lignite plants (75 GWh per day); and 18 liquid-fuel-based plants (4.8 GWh per day).

per day. The remainder comes hydroelectric (193 plants, 336 GWh per day), natural gas (62 plants, 121 GWh per day), nuclear (7 plants, 90 GWh per day), lignite (9 plants, 75 GWh per day), and liquid fuel-based generation (18 plants, 4.8 GWh per day).

Ministry of Power operating costs As part of a new transparency initiative, Merit Order Dispatch of Electricity for Rejuvenation of Income and Transparency ("MERIT"), the Ministry of Power began collecting and reporting power plant operating costs in October 2017. We scraped these data from the MERIT website, and matched the plants in these data to our main CEA sample. MERIT reports fixed and variable costs per kilowatt hour, which it obtains from the State Load Despatch Centre (similar to a electricity transmission company) in each of India's states. Assuming constant marginal costs, as is standard in the electricity literature, these reported variable costs are equivalent to the marginal cost of operations at each plant. These variable costs include fuel costs, any variable operations and maintenance, and within-state transmission charges. Because we only observe these data for the end of 2017, we must assume that these costs are static over time. We are able to match 273 plants in the CEA data with MERIT data, and these plants constitute the majority of capacity (225 GW out of 291 GW) and generation (2.4 TWh per day out of 2.9 TWh per day) in the CEA data. This matched sample of MERIT data includes 133 coal plants, 87 hydroelectric plants, 36 gas plants, 7 nuclear plants, 9 lignite plants, and 1 liquid fuel-based plant.

Constructed operating costs As an alternative to the Ministry of Power's reported variable cost, we construct our own measure of marginal operating costs using a variety of data sources. The CEA provides more detailed information on plant operations in the annual *Review of Performance of Thermal Power Stations*. In particular, each annual *Review* contains data on operating heat rates, a standard measure of electric generator efficiency measured in kilocalories per kilowatt-hour, for a large subset of coal-fired power plants. We obtained the 1997–2009 data from Chan, Cropper, and Malik (2014), and digitized the 2012–2014 *Reviews*, the most recent available at the CEA.[7] Because our analysis spans 2013 to 2017, we assign each plant its most recent heat rate that we observe. There are just 16 plants appearing in the *Reviews* for which the most recent heat rate was reported prior to 2012; for these plants, we obtained more recent heat rate data from tariff petitions to the Central Electricity Regulatory Commission. Indian coal prices vary by coal grade (i.e. coal's gross calorific value in kcal per kilogram), and we compute plant-specific coal grades by combining CEA heat rates with monthly coal consumption data from the CEA's Monthly Coal Reports. Ultimately, we have heat rate and coal grade data for 84 coal-fired plants and 7 lignite-fired plants, representing approximately 50 and 80 percent of each fuel's respective generating capacity from the CEA's daily generation data.[8]

[7]We thank the authors for sharing these data. We obtained access to hard copies of the the 2012–2014 *Reviews* on a recent trip to India.

[8]Centrally owned and state-owned plants face stricter reporting requirements than private power producers (Chan, Cropper, and Malik (2014)). Hence, we observe heat rates for a non-random sample

We require data on fuel prices to convert plant-specific heat rates in to plant-specific marginal costs. Transportation costs contribute a substantial share of the total costs of coal paid by power plants. In order to estimate these transportation costs, we geocoded every power plant in our CEA sample by hand using a combination of CEA data and Google Maps. We also obtained a GIS map of the major coalfields in India from the U.S. Geological Survey (USGS), and match each coal-fired power plant with the closest coalfield. Then, we estimate the shortest as-the-crow-flies distance between each plant and its closest coalfield. Since the majority of plants are either pit-head plants, or transport coal via railway, we use coal-specific freight rates from the Freight Operations Information System of the Ministry of Railways to compute approximate transportation costs per kilogram of coal for each plant.[9]

Fuel prices are strictly regulated in India; state-owned coal suppliers set coal prices by grade of coal, which must be approved by India's Ministry of Coal. We digitized pithead coal and lignite prices from the Ministry of Coal's annual *Coal Directory of India*, which contains coal-grade specific prices.[10] We calculate coal costs as pithead prices plus transportation costs as well as royalties and other taxes.

For natural gas-fired power plants, we use heat rate data from the CEA's monthly gas report. Each monthly report presents total generation and natural gas consumption. We follow the Ministry of Natural Gas and Petroleum in assuming that the calorific content of natural gas is 10,000 kCal per standard cubic meter.[11] These data enable us to construct heat rates for 58 of the 62 gas plants in our daily CEA sample. We digitized natural gas prices from the Ministry of Petroleum and Natural Gas' annual *Petroleum and Natural Gas Statistics*, which yields a fuel price time series similar to our annual times series of grade-specific coal prices.

Election data To investigate the extent to which political economy considerations impact misallocation, we incorporate data on the outcomes of India's state, or Legislative Assembly, elections. These data are publicly available from the Election Commission of India. We use Election Commission data from 2014–2017 on votes for each Legislative Assembly election candidate to determine (i) which party holds the most seats in each state legislature, and (ii) how the ruling party fares in each local election outcome.[12] This

of power plants. In our empirical approach, we simply exclude plants for which do not have data from the calculation of costs for both observed and least-cost scenarios; thus, to the extent that unobserved plants are dispatched in the wrong order, we are under-stating the extent of the cost gap between these two scenarios. We discuss this and other measurement error issues in Section 3.4.

[9]We are currently working to incorporate geospatial information on the rail network to calculate transportation costs *along the rail network*. This will modestly inflate our rail costs relative to those computed using as-the-crow-flies distances.

[10]Pithead (a.k.a. mine-mouth) prices are the commodity price of coal at (close to) the point of extraction.

[11]Monthly Gas Reports are available for 2012 and 2016–2017. We assign each plant its average observed heat rate over this period.

[12]We focus on Legislative Assembly elections because they occur more frequently than general elections, and because they occur on different dates in different states.

also allows us estimate the impact of aligned state and local party control on misallocation in electricity generation, using a close-election regression discontinuity design.

3.4 Empirical Approach

Our empirical analysis consists of three main steps. First, we construct the marginal operating cost for each power plant in our sample. Second, we compute the costs of marginal-cost-based—or “least-cost”—electricity dispatch. Next, we compute the costs of observed dispatch, and compute the total short run costs of observed deviations from least-cost dispatch. Finally, we conservatively account for both interregional and intraregional transmission constraints, in order to isolate the remaining misallocation wedge between least-cost and observed dispatch.

3.4.1 Marginal Costs

To quantify marginal costs, and eventually aggregate variable costs, we invoke several standard assumptions in electricity markets. We assume that power plant i may generate any feasible quantity of electricity at time t, such that its generation Q_{it} falls between 0 and its binding capacity constraint $\overline{Q}_i$.[13] We also assume that for any quantity of generation $Q_{it} \in [0, \overline{Q}_i]$, plant i faces constant marginal costs MC_{it}. This is a reasonable approximation for plants' true marginal costs, which are driven largely by (1) its costs of purchasing fuel inputs P_{it}^f (e.g., the price of coal inclusive of transportation costs, which is independent of Q_{it}); and (2) its heat rate HR_{it}, or the calorific content of fuel needed to produce 1 MWh of electricity.[14]

We can express plant i's marginal costs per MWh at time t as:

$$MC_{it} = P_{it}^f \cdot HR_{it} + VOM_{it} \tag{3.1}$$

VOM_{it} represents variable operations and maintenance costs, including labor and non-fuel materials, which are typically much smaller than fuel costs P_{it}^f (Fabrizio, Rose, and Wolfram (2007); Cicala (2017)).[15] Nuclear plants face relatively lower fuel costs and non-trivial costs of operations and maintenance, compared to conventional fossil fuel gen-

[13] We compute $\overline{Q}_i$ as the 98th percentile of plant i's observed generation in each year, following Davis and Hausman (2016), to ensure that our capacity constraints are within the space of feasible plant operations. As a robustness check, we also define $\overline{Q}_i$ as the 80th percentile of plant i's observed generation.

[14] Heat rates are the inverse of thermal efficiency, and plants with lower heat rates will tend to have lower marginal costs, all else equal. Energy economists typically assume that heat rates are constant across a plant's feasible quantities of generation (e.g. Cicala (2017)).

[15] In the U.S. or Europe, plants may face (implicit or explicit) taxes on the emissions of sulfur dioxide (SO_2), nitrous oxides (NO_x), or carbon dioxide (CO_2). However, during our sample, Indian power plants did not face emissions taxes, so we exclude them from our marginal cost measures here.

erators. Hydroelectric plants typically incur close to zero marginal costs but face complex dynamic operating constraints.[16]

Variable cost data from India's Ministry of Power's MERIT data include both fuel costs and non-fuel variable costs. Our main results use these MERIT data for plants' marginal operating costs. However, we also construct our own measures of marginal cost using plant-specific heat rates and input fuel prices (including transportation costs), as described in Section 3.3.

3.4.2 Least-cost and Observed Dispatch

Armed with a measure of each plant's marginal cost, we leverage two additional features of electricity markets in order to compute the costs of least-cost electricity dispatch. First, electricity demand is extremely inelastic in the short run, and we assume perfectly inelastic demand. This assumption is common in the literature and facilitates counterfactual comparisons that hold total electricity generation fixed (Cullen and Mansur (2017)). Second, electricity is a homogeneous commodity, and electrons move across the transmission grid at close to the speed of light. This means that absent transmission constraints, geographically dispersed power plants produce an identical product to sell into a single geographically integrated market. We begin by abstracting away from transmission constraints, before incorporating interregional and intraregional constraints in an extension to our main analysis.

We calculate the cost of least-cost dispatch by finding the cost-minimizing allocation of electricity generation that could potentially meet observed market demand (or "load"):

$$COST_t^{LC} = \min_{\mathbf{Q}_t^*} \sum_i MC_{it} \cdot Q_{it}^* \quad \text{s.t.} \quad \sum_i Q_{it}^* = \sum_i Q_{it} \ , \quad 0 \leq Q_{it}^* \leq \overline{Q}_i \ \forall i \tag{3.2}$$

Here, Q_{it} denotes *observed* generation for plant i in time t, while $\mathbf{Q}_t^*$ is an $[I \times 1]$ vector of *counterfactual* least-cost generation Q_{it}^* for each of the same I plants. The solution to this constrained minimization problem simply sorts power plants from lowest-to-highest MC_{it}, and then "dispatches" each plant at its full capacity (i.e. $Q_{it}^* = \overline{Q}_i$) until aggregate idealized generation is equal to aggregate observed generation. We can calculate $COST_t^{LC}$ at the daily level, for each day t in our dataset. We start by assuming a fully unconstrained transmission grid, such that the idealized allocation of generation $\mathbf{Q}_t^*$ is not restricted by individual plant locations. Below, we describe how we incorporate transmission constraints into our least-cost counterfactuals.

Equation (3.2) effectively calculates the area under an idealized electricity supply curve, up to a given aggregate quantity $\sum_i Q_{it}$. Each level of $\sum_i Q_{it}$ implies a unique cost-minimizing vector $\mathbf{Q}_t^*$, which simply ranks plants in order of lowest-to-highest marginal

[16]Large hydroelectric dams often optimize multiple objectives, only one of which relates to electricity production (e.g., other water uses, dam release limits). Nuclear, coal, and (to a lesser extent) natural gas plants also face less extreme dynamic operating constraints, which we ignore in our analysis (Mansur (2008); Reguant (2014)). Run-of-the-river plants store little to no water and provide non-dispatchable electricity at (virtually) zero marginal cost.

cost. This ranking, known as the "merit order", serves as an efficient allocative benchmark and implies a weakly monotonically upward-sloping supply curve for electricity generation. Deviations from this idealized merit order increase total variable costs of electricity generation.

Calculating the cost of observed generation is more straight-forward, as we simply multiply each plant's observed generation (Q_{it}) by its marginal costs:

$$COST_t^{OBS} = \sum_i MC_{it} \cdot Q_{it} \tag{3.3}$$

Figure 3.4.2 illustrates the intuition behind calculating costs for both least-cost and observed dispatch. Gray shaded areas show the total variable cost associated with each blue dispatch curve. Observed dispatch meets market demand (vertical dashed lines) at greater total variable costs than least-cost dispatch.

3.4.3 Cost Gaps

We calculate the difference between realized vs. idealized costs for each day in our sample as:

$$COST_t^{DIFF} = COST_t^{OBS} - COST_t^{LC} = \sum_i MC_{it} \cdot (Q_{it} - Q_{it}^*) \tag{3.4}$$

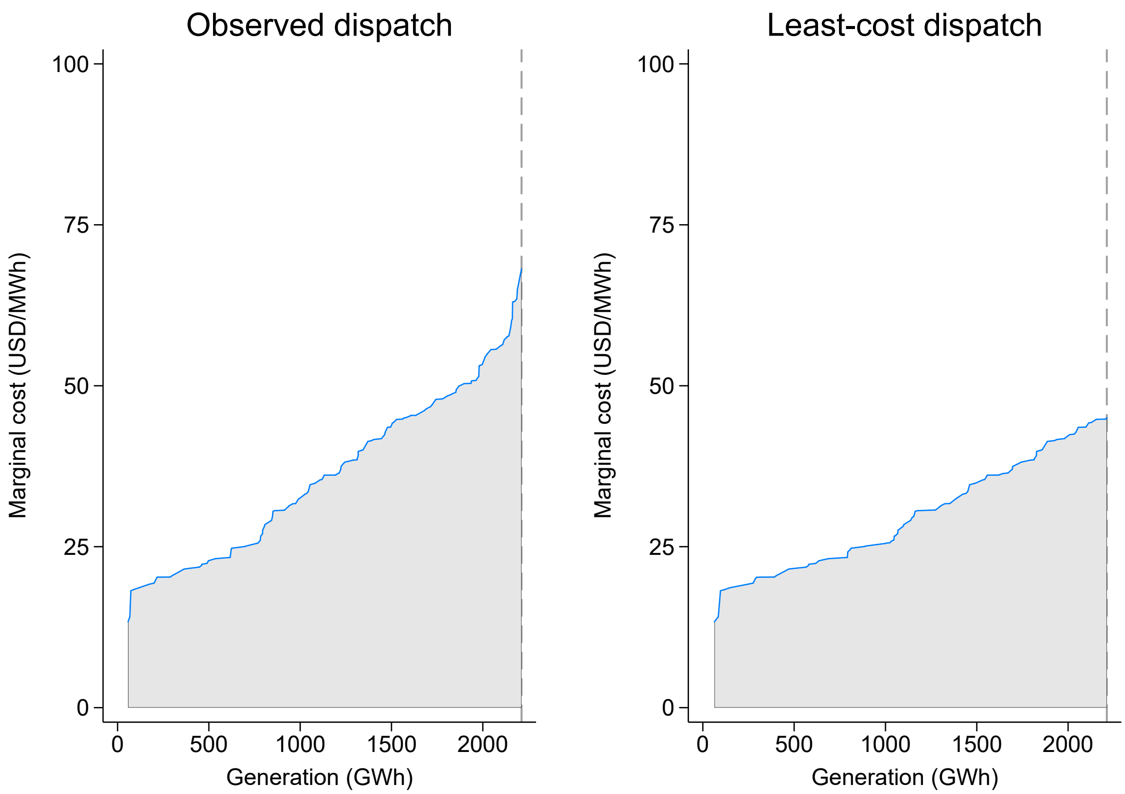
$COST_t^{DIFF}$ summarizes the economic cost of deviations from least-cost electricity supply, or the potential decrease in total variable cost for day t. Importantly, the set of plants $i = \{1, \ldots, I\}$ entering into each cost calculation is a subset of the full population of electricity generating plants. We omit plants for which we do not observe data on marginal costs, meaning that $\sum_i Q_{it}$ is less than full market-wide electricity load for day t. By excluding a subset of plants from $COST_t^{DIFF}$, we implicitly assume that unobserved plants are efficiently dispatched and offer no potential gains from reallocation (i.e. $Q_{jt}^* = Q_{jt}$ for any plant $j \notin \{1, \ldots, I\}$). We also omit hydroelectric plants from our preferred cost calculations, due to the complex nature of characterizing theoretically "optimal" hydroelectric generation.[17] Hence, our calculations represent a lower bound on deviations from least-cost dispatch, if there exist additional cost-reducing potential reallocations that we do not observe.

3.4.4 Transmission Constraints

The above cost difference calculation assumes unconstrained electricity transmission, yet physical transmission constraints may limit the flow of power between regions of the elec-

[17] If we included hydro plants, then the solution to Equation (3.2) would assume maximum hydro generation in all time periods due to hydro's extremely low marginal costs. In reality, dams cannot generate at 100 percent capacity due to several dynamic constraints—most notably, a finite supply of water.

Figure 3.4.2: Costs of Observed vs. Least-cost Dispatch



Notes: This figure illustrates the empirical approach underlying our main estimates using data from March 12, 2015. In the "Observed dispatch" panel, we take all the plants that were observed to be running on this date, and plot the quantity they generate in increasing order of marginal cost. The blue line shows the "supply curve" of plants, as ordered by marginal cost. The gray area is our estimate of total observed costs—in this case, \$78.9 million. The dashed line shows the total load on this day, of 2.2 TWh. The "Least-cost dispatch" panel, by contrast, orders all plants from lowest to highest marginal cost, and dispatches plants to their capacity in this order until we reach 2.2 TWh of generation. Again, the gray area represents the total costs of meeting this load: \$65.3 million, or \$13.6 million lower than the cost of observed dispatch. It is clear from this figure that total costs under least-cost dispatch are lower than those under observed dispatch, and that this is because more low-cost plants are being dispatched to meet load under the least-cost-dispatch scenario.

tric grid. For example, consider two electricity market regions (A and B) separated by a single transmission line with a finite flow capacity, but with otherwise unconstrained intraregional transmission. The efficient allocation of generation will dispatch plants in order of lowest-to-highest cost (ignoring plant regions) *unless* the AB transmission constraint binds. However, if the constraint binds and limits the interregional flow of electricity, then an efficient allocation of generation may force relatively high cost plants to operate in region A (in order to satisfy region A's demand) while relatively low cost plants in region B sit idle. In reality, India has five major transmission regions: the Northern Region, Western Region, Eastern Region, Southern Region, and North Eastern Region. Figure 3.4.3 displays these regions, along with the locations of all major power plants.

To quantify the extent to which transmission constraints might explain the total cost gap $COST_t^{DIFF}$, we compute the costs of least-cost dispatch for each of India's five major electricity market regions (indexed by r):

$$COST_{rt}^{LC} = \min_{\mathbf{Q}_{rt}^*} \sum_{i \in r} MC_{irt} \cdot Q_{irt}^* \quad \text{s.t.} \quad \sum_{i \in r} Q_{irt}^* = \sum_{i \in r} Q_{irt} \ , \quad 0 \leq Q_{irt}^* \leq \overline{Q}_{ir} \ \forall i \tag{3.5}$$

Then, we sum across regions to calculate the total cost of least-cost dispatch:

$$COST_t^{LC_R} = \sum_r COST_{rt}^{LC} \tag{3.6}$$

This allows us to calculate the cost difference between observed dispatch and least-cost dispatch under interregional autarky:

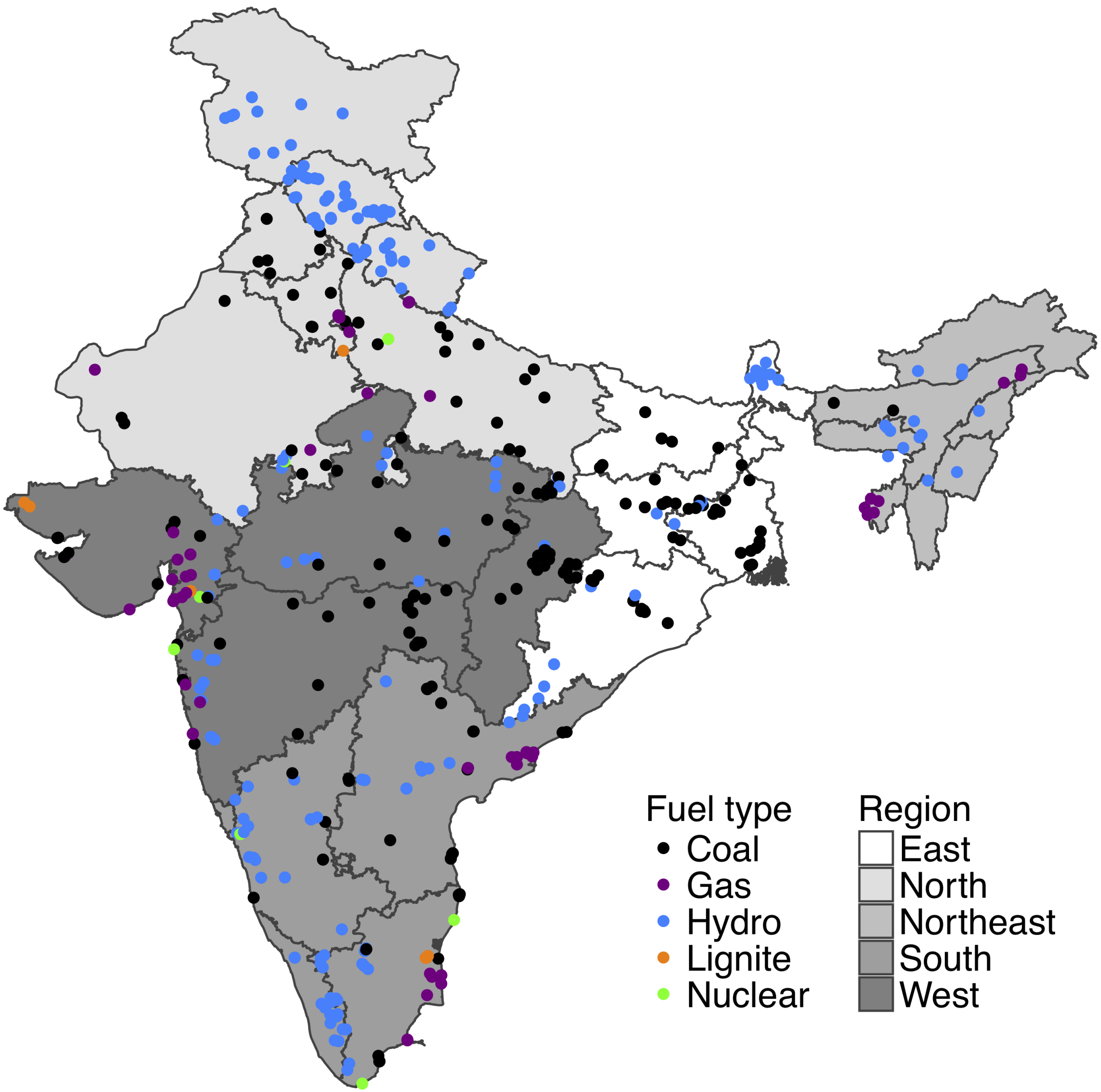
$$COST_t^{DIFF_R} = COST_t^{OBS} - COST_t^{LC_R} \tag{3.7}$$

Here, we make the extreme assumption that each region is a completely separate market, or that there is zero interregional transmission capacity. We repeat this exercise using the 13 subregions of the Indian electric grid, which also accounts for *intra*regional transmission constraints by assuming interregional *and* intraregional autarky. The resulting cost difference calculations represent a lower bound on the short-run misallocation wedge, as trade between and within regions should only lower the cost of the idealized least-cost benchmark.

3.4.5 Measurement Error

It is possible that our plant-specific marginal costs data are imperfect—particularly when we construct our own marginal costs, rather than relying on MERIT costs reported by the Ministry of Power. Measurement error in marginal costs will impact our total cost calculations: if our data understate (overstate) the marginal cost of some plants in our sample, we will compute total costs that are too low (high). However, we are primarily interested

Figure 3.4.3: Electricity Market Regions of India



Notes: This figure displays the locations of the 485 power plants that appear in the CEA's Daily Generation Reports. The plant colors indicate the fuel type—coal, gas, hydroelectric, lignite, and nuclear. States are shaded by their major transmission region—North, North East, East, West, and South.

in calculating the cost of *deviations* from a least-cost dispatch order. The relative cost *differences* between least-cost dispatch and observed dispatch will be largely insensitive to classical and non-classical measurement error in plant-specific marginal costs, for two reasons. First, we use the same cost numbers to calculate observed costs and least-cost costs, meaning that inaccuracies will impact both estimates similarly. Second, and more importantly, our calculations aggregate costs across many inframarginal plants up to a single marginal plant. For measurement error to significantly impact our findings, it would need to systematically overstate (understate) the operating costs of plants with low (high) costs.

3.5 Results

3.5.1 Marginal Costs

We begin by constructing the marginal-cost-based merit order for the Indian power sector. Using the Ministry of Power data, we simply rank plants from lowest to highest marginal cost. Figure 3.5.4 shows the results of this exercise, which effectively solves the cost minimization problem in Equation (3.2). As expected, we find that hydroelectric generation tends to have the lowest marginal cost. The Tarapur Atomic Power Station is the cheapest operating nuclear plant, with a marginal cost of $19 per MWh; the most expensive nuclear power plant, Kudankulam Atomic Power Station, has a reported cost of $68 per MWh.[18]

Coal plants make up the bulk of the merit order, and range from $13 per MWh to $74 per MWh, with an average of $41 per MWh; lignite plants also average $41 per MWh. Gas plants are more costly, ranging from $22 per MWh to $90 per MWh.[19] These marginal costs are broadly consistent with other estimates of the cost of generation in the United States (Davis and Hausman (2016)) and in India (Ryan (2017)).

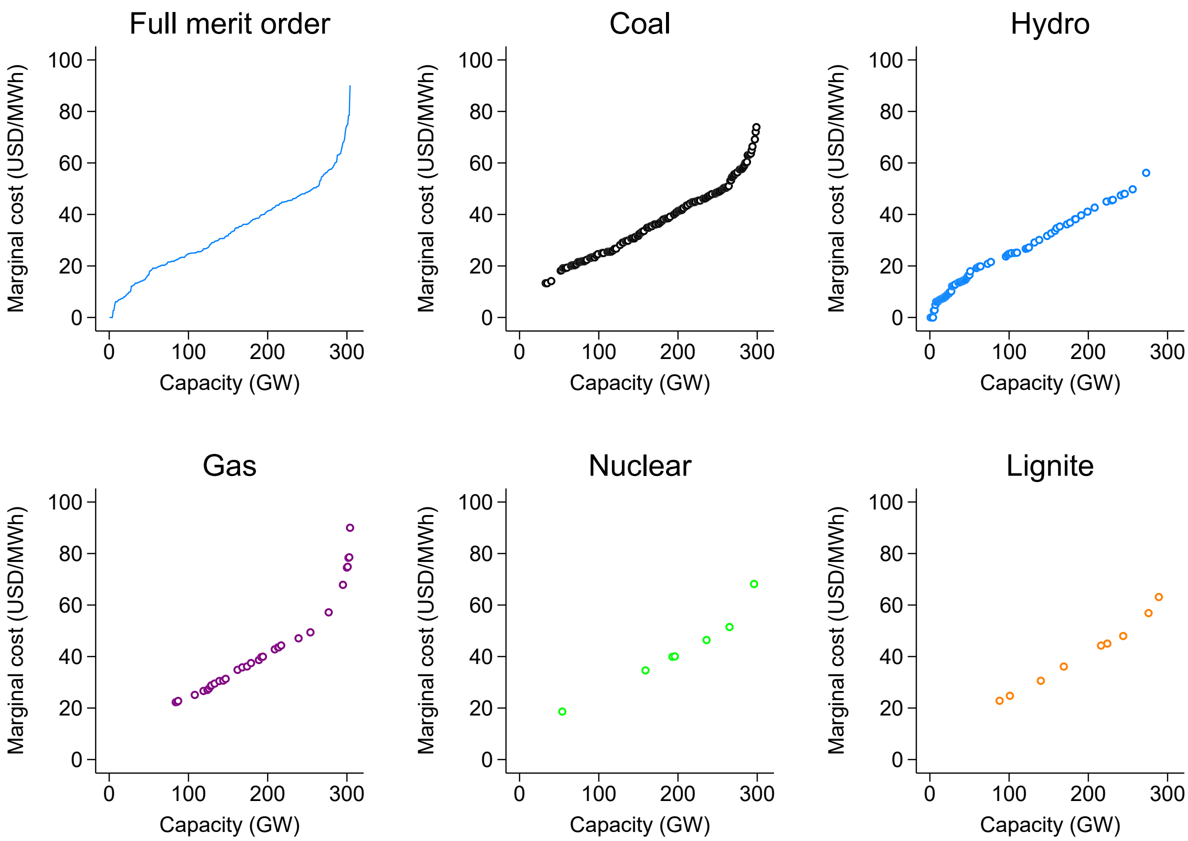
3.5.2 Least-cost Dispatch

Armed with these marginal costs, we can follow the procedure outlined in Section 3.4 to compute generating costs of least-cost dispatch for each day in the sample. We first calculate total load on day t as $\sum_i Q_{it}$—that is, observed generation for each plant in our sample, for which we have nonmissing data. As described above, we exclude hydroelectric

[18] Tarapur operates a light water reactor, while Kudankulam has a Russian heavy water reactor. It is somewhat surprising that most nuclear power plants have relatively high marginal costs. It is possible that for nuclear power, this "marginal cost" variable is more akin to a levelized cost, as marginal costs per MWh of nuclear generation are notoriously difficult to obtain. At the same time, anecdotal evidence suggest that nuclear power plants may indeed incur substantial marginal operating costs: `https://timesofindia.indiatimes.com/india/Cost-of-nuclear-power-proving-high-DAE-in-a-fix/articleshow/27920490.cms`.

[19] We exclude the single liquid-fuel-based power plant from Figure 3.5.4. This plant, Basin Bridge, has a marginal cost of $270 per MWh.

Figure 3.5.4: Marginal-Cost-Based Merit Order



Notes: This figure shows the Indian electricity supply merit order, where we rank plants by their marginal cost according to the Ministry of Power's MERIT cost data. The top-left panel shows the full merit order, combining all fuel types into one pseudo supply curve. The remaining five panels break this merit order into its various component plants. In each panel, one dot represents 1 GW of capacity; larger plants will be represented by multiple dots. Coal-fired plants make up the bulk of the merit order, followed by hydroelectric, gas, nuclear, and lignite plants. We omit the one liquid-fuel plant for graphical purposes.

plants from these results, because they face complex dynamic constraints. By dropping hydro generation from the sample prior to constructing our measure of total load, we essentially assume that they are always dispatched optimally.

To compute the cost of least-cost dispatch, we "dispatch" plants in order of least cost, at full capacity, until we reach total observed generation. This creates a vector $\mathbf{Q}^*_{it}$ where for a given marginal cost MC^*_t:

$$\begin{aligned} Q^*_{it} &= \overline{Q}_i && \text{if } MC_{it} < MC^*_t \\ Q^*_{it} &\in \left[0, \overline{Q}_i\right] && \text{if } MC_{it} = MC^*_t \\ Q^*_{it} &= 0 && \text{if } MC_{it} > MC^*_t \\ \textstyle\sum_i Q^*_{it} &= \textstyle\sum_i Q_{it} \end{aligned}$$

This vector $\mathbf{Q}^*_{it}$ solves the constrained minimization problem in Equation (3.2), and allows us to calculate the cost of least-cost dispatch on day t as $\sum_i MC_{it} \times Q^*_{it}$. Because we only observe MERIT data for 2017, these calculations use $MC_{it} = MC_i$ that is not time-varying.[20]

Figure 3.5.5 displays the results of this exercise. Panel A shows the total costs of least-cost dispatch over the course of our sample. The average daily cost is \$66.6 million, with a standard deviation of \$7.9 million. The lowest cost we observe in the sample is \$42.1 million, and the highest is \$87.9 million. Panel B displays the marginal cost of the marginal plant (or MC^*_t). As expected, the cost of the marginal plant tracks closely with the total cost, and with total daily load shown in Figure 3.3.1.[21] On average, the marginal plant's marginal cost per MWh is \$46.1, with a minimum value of \$38.4 and a maximum value of \$50.4.

3.5.3 Least-cost versus Observed Dispatch

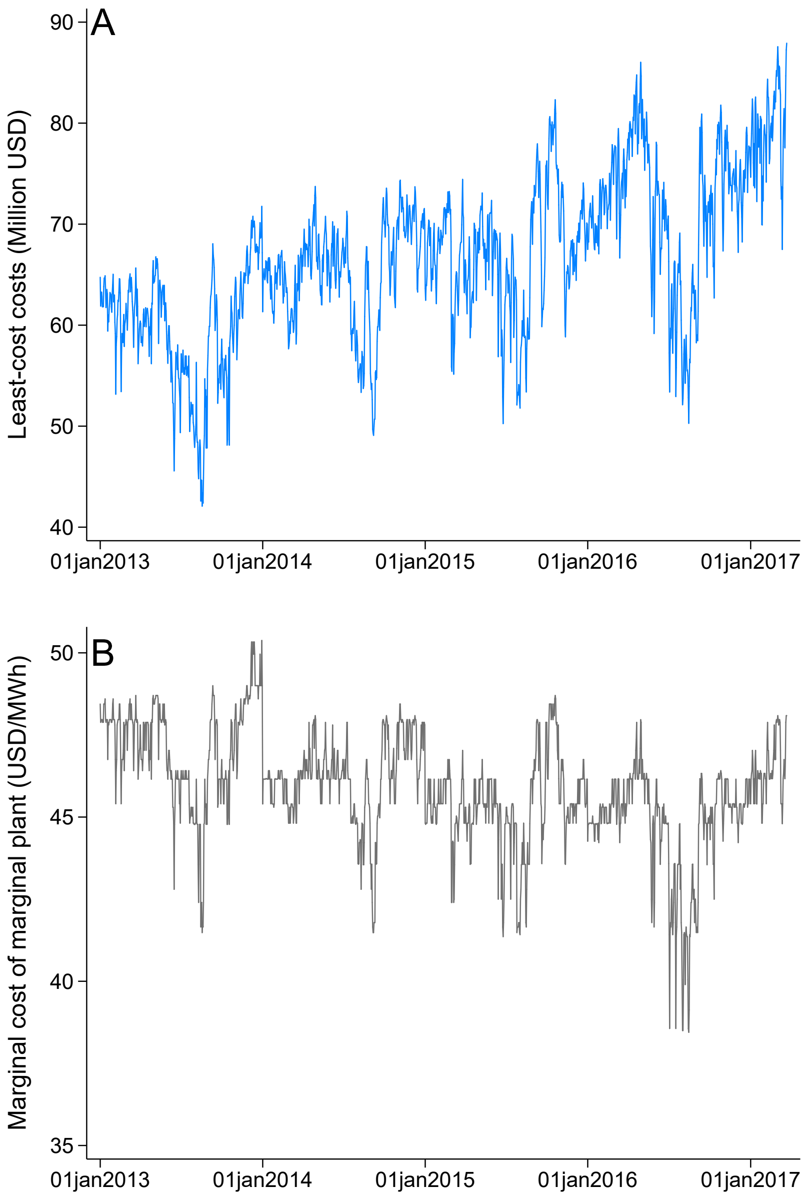
Next, we compute actual costs following Equation (3.3). We simply multiply each plant's observed generation by its marginal cost, and sum over all plants for each day in the sample. This allows us to calculate the difference between the costs of observed and least-cost dispatch, as in Equation (3.4). Figure 3.5.6 presents these cost differences. Panel A shows a kernel density of the daily cost gap between observed and least-cost dispatch, and Panel B plots the daily time series of cost difference per MWh.

Average daily observed costs of electricity generation in our sample are \$79.4 million per day, or around \$29 billion per year. This translates to an average cost of \$36.1 per MWh. Observed daily costs range from \$55.2 million per day to \$99.3 million per day. The gap between the cost of observed and least-cost dispatch is quite large: the average cost gap is \$12.8 million per day, or \$5.9 per MWh. The first row of Table 3.5.1 presents these results numerically.

[20] Costs in the MERIT data are reported in rupees. We convert to USD with an exchange rate of 60 rupees to 1 dollar.

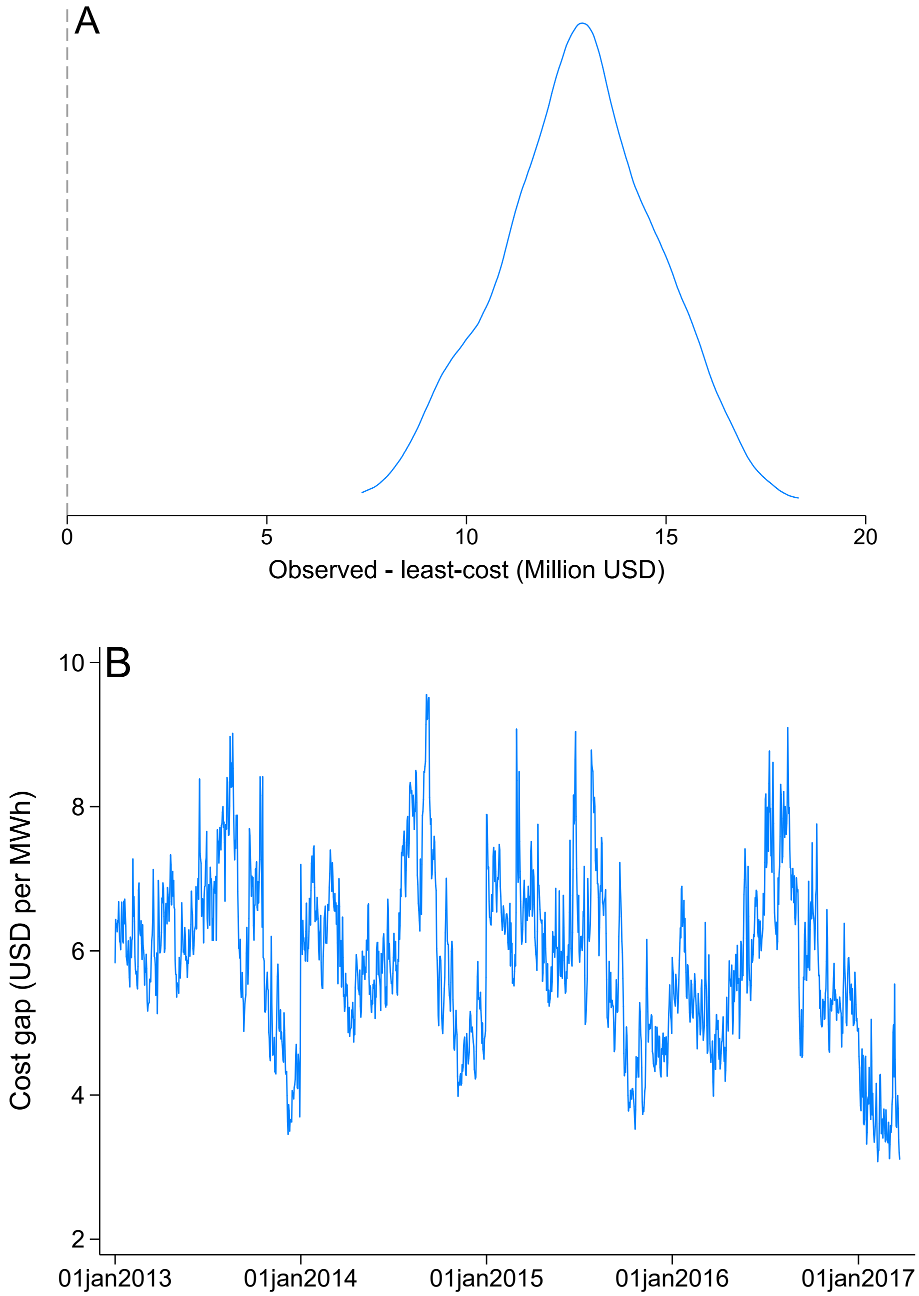
[21] Note that total daily load from Figure 3.3.1 is not identical to total daily observed generation in these calculations, as we only use plants with MERIT data (excluding hydroelectric generation) in the results shown in Figure 3.5.5.

Figure 3.5.5: Costs of Least-cost Dispatch



Notes: This figure displays our estimates of the cost of least-cost generation in India, using the Ministry of Power's MERIT data on marginal costs. As described in Sections 3.4 and 3.5, we construct the least-cost-dispatch-based cost estimates by ranking plants from lowest to highest marginal costs, "dispatching" each plant to generate up to its capacity, until we meet total daily load. Panel A displays the total short-run costs of least-cost generation for the plants in our sample, which averages \$66.6 million per day, with a minimum of \$42.1 million per day, and a maximum of \$87.9 million per day. Panel B shows the marginal cost of the marginal plant under least-cost dispatch. As expected, higher total cost days (which correspond to higher overall load days, as shown in Figure 3.3.1) must rely on higher-cost plants to be operating. The average marginal cost of the marginal plant in our sample is \$45.9 per MWh, and ranges from \$38.4 to \$50.4 across the sample period.

Figure 3.5.6: Cost Difference between Observed and Least-cost Dispatch



Notes: This figure shows the differences in cost between least-cost and observed dispatch in the Indian electricity market, computed using marginal costs from the Ministry of Power's MERIT data. Panel A shows the distribution of the cost difference between least-cost dispatch and observed dispatch across the sample. This cost difference ranges from \$7.4 million per day to \$18.3 million per day, with an average of \$12.8 million per day. Panel B shows the cost gap *per MWh* over the sample. On average, this cost difference is \$5.9 per MWh, compared to average observed costs of \$36 per MWh—around 17 percent of observed costs.

Table 3.5.1: Costs of Electricity Supply

Analysis	Least-cost (M USD / day)	Observed (M USD / day)	Difference (M USD / day)	$100 \cdot \left(\frac{\text{Diff}}{\text{Obs}}\right)$
A. Main results				
Unrestricted transmission	66.6 [42.1, 87.9]	79.4 [55.2, 99.3]	12.8 [7.4, 18.3]	16.3 [8.9, 25.7]
Interregional autarky	69.0 [45.8, 88.7]	79.4 [55.2, 99.3]	10.4 [6.6, 14.0]	13.2 [8.0, 19.4]
Intraregional autarky	69.4 [46.3, 88.7]	78.3 [54.3, 97.4]	8.8 [5.8, 12.0]	11.4 [6.9, 17.4]
B. Sensitivities				
Constructed cost data	23.3 [16.8, 28.8]	28.0 [20.7, 33.5]	4.7 [3.0, 6.4]	16.8 [10.3, 24.1]
Capacity 80 percent	70.5 [44.8, 91.6]	79.4 [55.2, 99.3]	8.8 [2.4, 15.3]	11.3 [3.1, 22.1]
No 90-percent OOM plants	50.5 [31.7, 73.4]	59.8 [40.6, 79.8]	9.2 [5.0, 13.5]	15.6 [8.1, 24.5]

Notes: This table reports our main results, calculated at the daily level across all 1,530 days in our sample. Mean estimates are listed first, with the minimum and maximum values in brackets. As described in Sections 3.4 and 3.5, we construct the cost of least-cost dispatch by ranking plants from lowest to highest marginal costs, dispatching each plant to produce up to its capacity, until aggregate supply meets total daily electricity demand. The total cost associated with this least-cost dispatch is the sum of each power plant's marginal cost times this plant's electricity generation across all power plants. Observed costs sum over plants of each plant's observed electricity generation times its marginal cost. The third column simply subtracts least-cost aggregate costs from observed aggregate costs. "Unrestricted transmission" allows for reallocation across all plants in the electricity market. "Interregional autarky" allows for reallocation across all plants within each grid region, but not between regions. "Intraregional autarky" allows for reallocation across all plants within each grid subregion, but not between subregions or regions. (Region A2 in the Northeast has too few plants to be effectively reallocated and is dropped from this analysis only.) "Constructed cost data" uses our own constructed measure of each plant's marginal costs over time. "Capacity 80 percent" defines capacity as the 80th percentile (rather than 98th percentile) of each plant's maximum observed generation. "No 90-percent OOM plants" removes the 25 plants which we find to be generating out-of-merit on more than 90 percent of sample days, to confirm they are not driving all of our results (this effectively exclude plants near cities). We present the average daily estimates in the first row for each analysis type, and the minimum and maximum across the sample in brackets. Sensitivities in Panel B assume unrestricted transmission both across and within regions.

Annually, this amounts to a cost gap of \$4.7 billion per year. Put differently, the cost difference between observed dispatch and least-cost dispatch amounts to over 16 percent of realized variable costs. This suggests that there may be substantial short-run gains from reallocation in the Indian electricity market.

We calculate a substantial cost gap between observed and least-cost dispatch, indicating that electricity generation departs markedly from the cost-based merit order. This cost gap could reflect systematic departures, wherein many plants are frequently dispatched by a rule of thumb that is imperfectly correlated with marginal cost. On the other hand, a few large high-cost plants could regularly generate out of merit and be responsible for most of the deviations we observe.

For each day in the sample, we classify a plant as "out of merit" if the plant generates despite a least-cost generation of zero, or vice versa. In other words, we assign plant i to be out of merit on day t if $Q_{it} > 0$ and $Q^*_{it} = 0$ *or* if $Q_{it} = 0$ and $Q^*_{it} > 0$. Panel A of Figure 5 reports a histogram for the percent of days that plants generate out-of-merit, for all plants in our sample. This reveals that the majority of plants rarely generate out of merit, while a specific subset of plants frequently generate out of merit. In fact, roughly 25 plants generate out of merit on at least 90 percent of the days in our sample. This suggests that reducing out-of-merit generation for these few plants could yield substantial short-run cost reductions.

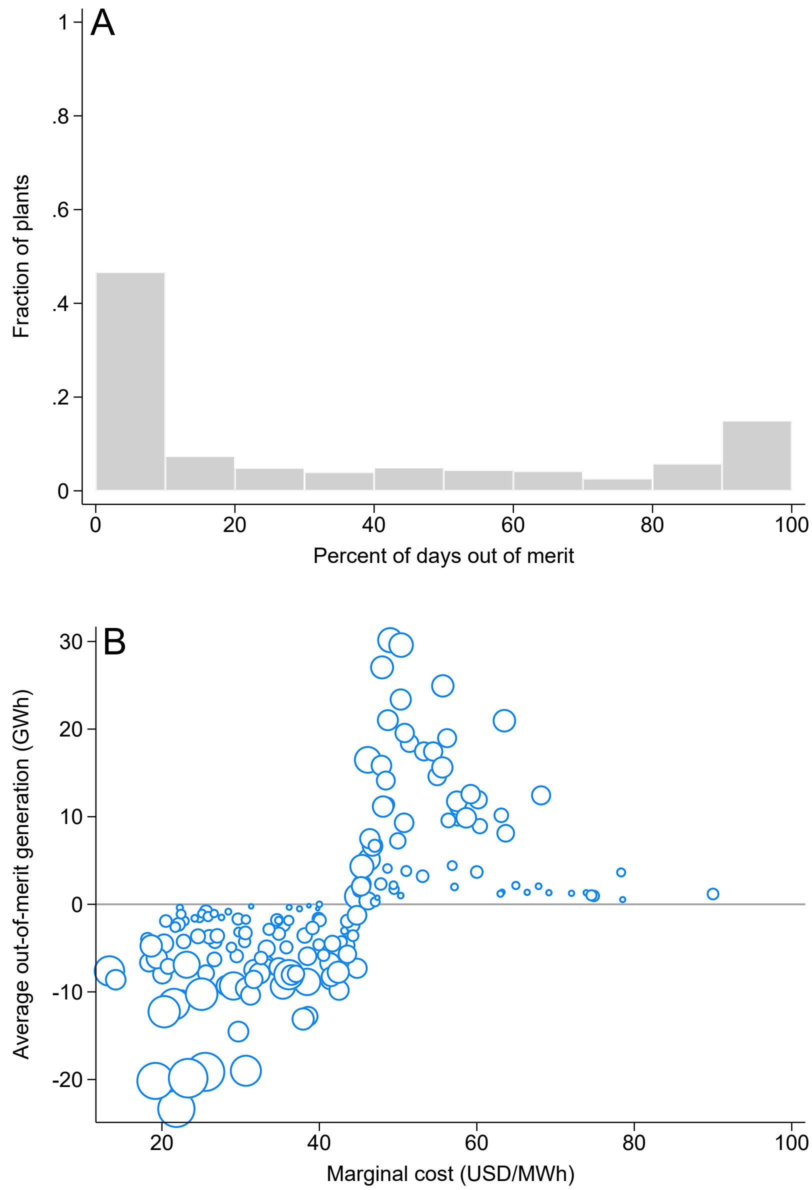
We then compute the level difference between observed generation and least-cost generation for each plant-day, or $Q_{it} - Q^*_{it}$. Panel B of Figure 3.5.7 plots each plant's difference between observed and least-cost generation against its marginal costs. Unsurprisingly, departures from least-cost generation are strongly correlated with marginal costs of generation: plants with higher marginal costs are more likely to generate out-of-merit. In particular, we find that while low-cost plants generate slightly too little on average, there are a few higher-cost plants that generate substantially more than our least-cost benchmark.

3.5.4 Transmission Constraints

Whereas our results have thus far imposed a strong assumption of no transmission constraints, we now consider the opposite extreme by assuming autarky for both regions and subregions of the Indian electric grid. Using Equation (3.5), we calculate the costs of least-cost dispatch separately for plants in each of 5 regions (or 13 subregions), and aggregate the total cost differences across regions (or subregions). Figure 3.5.8 reports two distributions of daily cost gaps, assuming respectively: (i) unconstrained transmission (reproduced from Panel A of Figure 3.5.6); and (ii) interregional autarky, with unconstrained transmission within subregions.

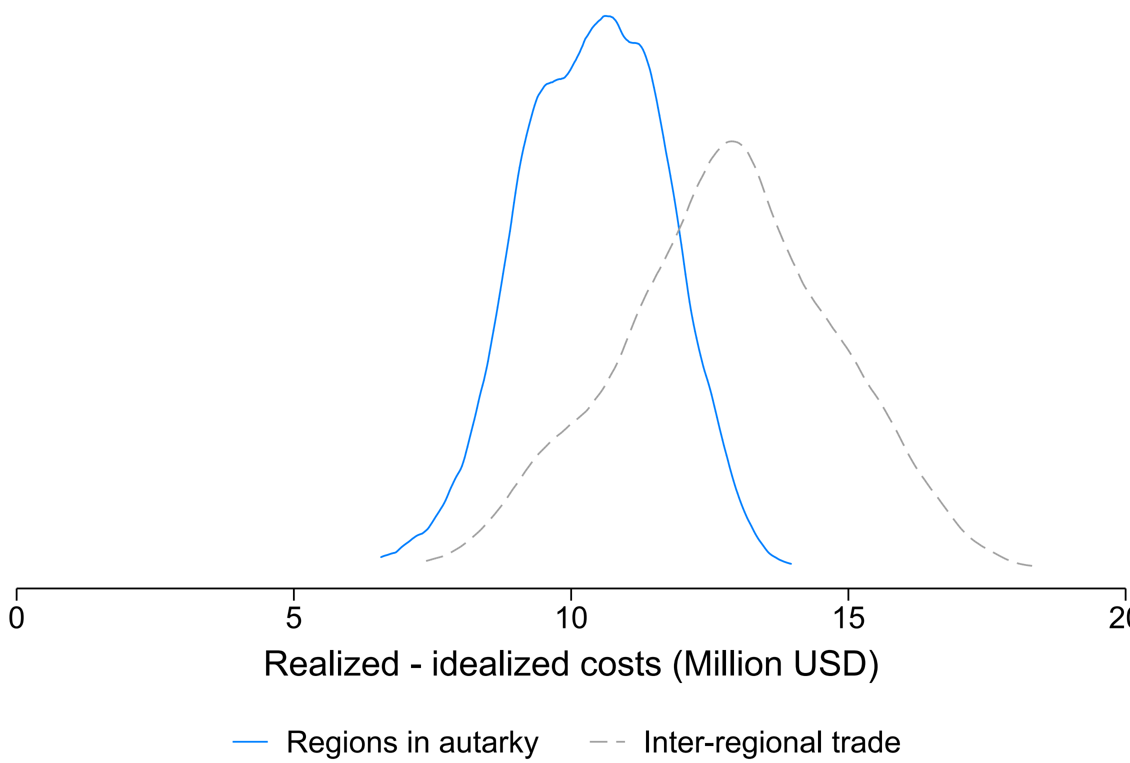
Under interregional autarky, we find that the average cost difference between observed and least-cost dispatch is \$10.4 million per day, or 13 percent (reported in the second row of Table 3.5.1). This is only \$2.4 million per day (or 19 percent) less than the average cost difference under fully unconstrained transmission. In other words, interregional transmission constraints appear to account for less than 20 percent of the cost difference between

Figure 3.5.7: Out-of-merit Generation



Notes: This figure explores out-of-merit generation in least-cost dispatch as compared with observed dispatch. Panel A shows the percent of days that each plant generates out-of-merit, or is observed to be operating in the CEA data when it should not be operating according to our least-cost dispatch order. While most plants rarely generate out-of-merit, around 20 percent of plants generate in the observed dispatch order but not in the least-cost dispatch order more than 90 percent of the time. Panel B shows the average difference between observed generation and generation under least-cost dispatch as compared with marginal costs per MWh. There is a strong correlation between marginal costs and out of merit generation: low-cost plants appear to be generating too little, while higher-cost plants are generating too much. These figures use the Ministry of Power's MERIT data on marginal costs of generation, and both panels are weighted by capacity.

Figure 3.5.8: Role of Interregional Transmission Constraints



Notes: This figure displays the cost difference (in millions of dollars per day) between least-cost dispatch and observed dispatch under two different scenarios. The dashed gray line shows the distribution of cost differences when we create the least-cost dispatch order by ranking *all* plants by their marginal cost. This essentially assumes that electricity from any plant can meet load anywhere in the system. By contrast, the solid blue line displays cost differences from an exercise in which we treat India's five transmission regions—as shown in Figure 3.4.3—as five closed economies. That is, all demand in a given region can be met by plants in that region only. This estimates an upper bound on the impact of interregional cost differences on the total cost gap between least-cost and observed dispatch, under the extreme assumption that there is no transmission capacity connecting regions. Even under this strong assumption, we find that less than 20 percent of the cost gap is driven by interregional transmission constraints, suggesting that more than 80 percent of the difference in cost between least-cost and observed dispatch remains unaccounted for. This figure was constructed using marginal cost data from the Ministry of Power's MERIT database; the unit of observation in this figure is a day.

observed and least-cost generation allocations. We see a similar result if we rescale cost differences by total generation. The average cost gap is \$4.8 per MWh under interregional autarky, only 19 percent less than the \$5.9/MWh average cost gap under unconstrained transmission. This comparison is striking: even after accounting for interregional transmission constraints, widely considered to be the most important physical limitation in electricity markets, the vast majority of misallocation in electricity generation remains.

Under both interregional *and* intraregional autarky, we calculate an average cost gap of \$8.8 million per day, or 11 percent (reported in the third row of Table 3.5.1). This implies that transmissions constraints between regions/subregions explain less than 31 percent of the cost difference between observed and least-cost allocations. Having accounted for the physical constraints of electricity transmission both across and within regions, the remaining 11 percent cost gap—equivalent to \$3.2 billion per year—represents a conservative estimate of the short-run misallocation wedge in Indian electricity supply.

3.5.5 Robustness Checks

3.5.5.1 Constructed Marginal Cost Data

As an alternative to the MERIT data, we construct our own measures of marginal cost. Here, we define marginal costs based on time-varying fuel prices and plant-specific heat rates alone, as we lack data on other components of marginal cost:

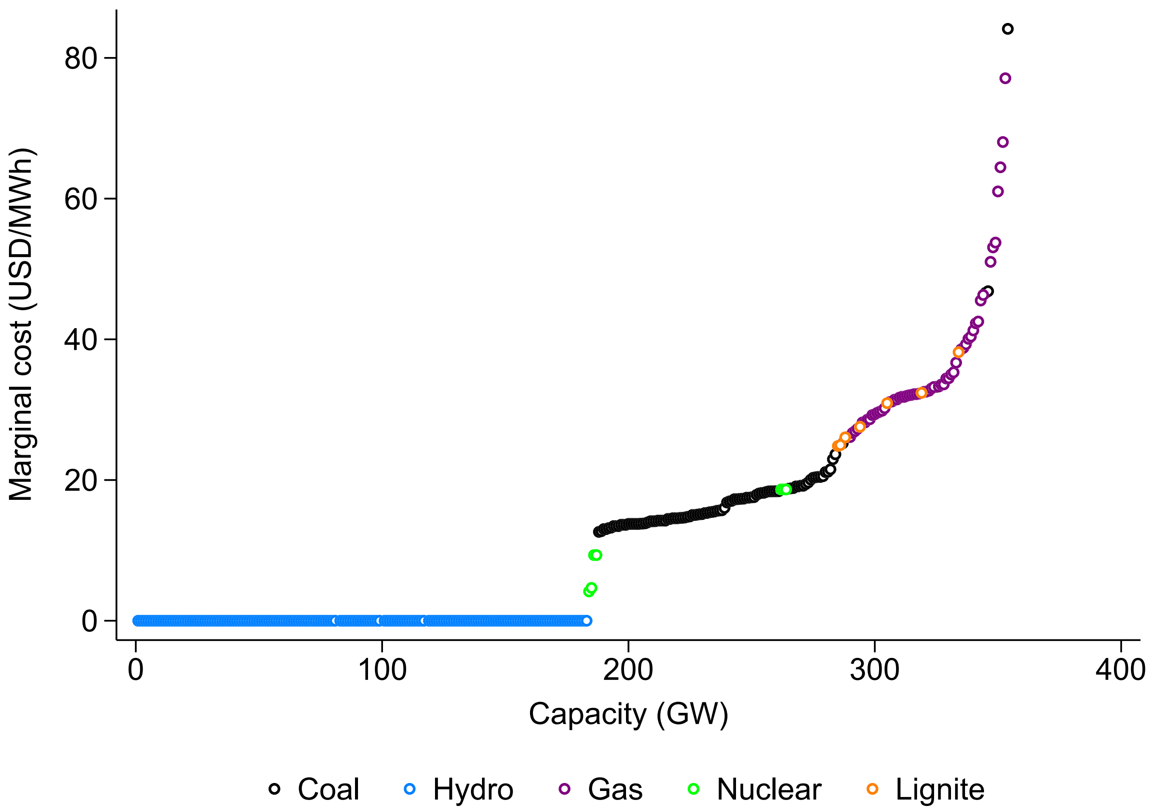
$$MC_{it} = P_{it}^{fuel} \cdot HR_{it} \tag{3.8}$$

As a result, these constructed costs are less comprehensive than those reported by the Ministry of Power. However, our constructed marginal cost variable should assuage concerns about misreporting bias on the part of plants or State Load Despatch Centres. Figure 3.5.9 shows our constructed marginal costs.

The resulting merit order is quite different from that based on the Ministry of Power data. First, our calculations assume that hydroelectric generation has a marginal cost of \$0 per MWh, and we use fuel-only costs for nuclear power.[22] Furthermore, our merit order is sharply differentiated by fuel type: hydroelectric plants are cheapest, then low-cost nuclear, then coal (with some nuclear mixed in), and then a mix of gas and lignite. This is perhaps unsurprising, as coal is cheaper than gas, and lignite plants tend to be less efficient. Given that Equation (3.8) includes only heat rates and fuel prices, most of the variation in marginal costs comes from fuel prices alone (rather than a combination of fuel prices and variation in operations and maintenance costs). Our constructed marginal cost also tend to be significantly less than cost reported by the Ministry of Power. In our constructed cost sample, the average coal plant has a marginal cost of \$15.6 per MWh, the average gas plant has a marginal cost of \$36.5 per MWh, the average lignite plant's

[22] Our nuclear costs are based on information from the Chairman of the Department of Atomic Energy's expert committee on nuclear tariffs, described in *The Hindu*: `http://www.thehindu.com/todays-paper/tp-opinion/Why-India-should-opt-for-nuclear-power/article14850892.ece`.

Figure 3.5.9: Marginal-Cost-Based Merit Order, Constructed Costs



Notes: This figure shows the Indian electricity supply merit order, where we rank plants by their marginal cost according to our own constructed marginal cost estimates. This figure shows all fuel types together, with colors denoting fuel types. Each dot represents 1 GW of capacity; larger plants will be represented by multiple dots. India has a large amount of hydroelectric generating capacity - though this is hamstrung in reality by dynamic considerations, which we assume has a marginal cost of zero. The next entries in the merit order are low-cost nuclear facilities; followed by coal-fired plants, interspersed with several other nuclear power plants; more coal; and a mix of coal-, natural gas-, and lignite-based plants.

marginal cost is \$29.3 per MWh, and the nuclear plants range in cost from \$4.2 per MWh to \$18.7 per MWh.

As with the Ministry of Power data, we can use these marginal costs measures to compute the total variable costs of least-cost dispatch for each day in the sample. We find an average cost of \$23.3 million per day—about one third of costs of least-cost dispatch that we compute using the Ministry of Power data—with a standard deviation of \$1.9 million. This translates to an average variable cost of approximately \$14.8 per MWh, from dispatching in order of our constructed marginal costs.

We compute average observed costs of \$28.0 million per day, or \$17.6 per MWh, using our constructed costs estimates. Combining these results, we find an average cost difference between observed and least-cost-dispatch-based costs of \$4.7 million per day (ranging from \$3.0–\$6.4 million per day). This is a cost difference of \$3.0 per MWh on average, ranging from \$1.7 per MWh to \$4.5 per MWh. While these cost numbers are substantially smaller in levels than those computed with the Ministry of Power data, the relative sizes of the cost difference as a share of total observed costs are quite similar. Using constructed cost data, the cost gap between observed vs. least-cost dispatch represents approximately 17 percent of total observed costs, compared to 16 percent when using Ministry of Power cost data. We report these results in the fourth row of Table 3.5.1, which corroborate our main findings using MERIT cost data.

3.5.5.2 Plant Capacities

Thus far, we have set each plant's annual capacity equal to the 98th percentile of its distribution of daily electricity generation in that year. However, this could potentially overstate the capacity that plants can realistically supply on a typical day. As a robustness check, we instead assign capacity equal to the 80th percentile of this distribution. We report these results in the fifth row of Table 3.5.1, and they are quite similar to our preferred method. This implies that our results are not sensitive to our method of assigning plants' generating capacities.

3.5.5.3 Plants Always Out-of-merit

Imposing intraregional autarky ignores any local transmission constraints that are common in high-population, high-demand regions of the grid. Transmission bottlenecks into large urban areas likely necessitate that certain power plants always receive prioritized dispatch, regardless of their marginal costs. We conduct a sensitivity analysis that recalculates the cost gap excluding the 25 plants that we find to be generating out of merit on over 90 percent of days in our sample. We report these results in the last row of Table 3.5.1, and they are quite close (in percentage terms) to our preferred estimates.

3.6 Discussion

This paper's goal is to quantify departures from least-cost dispatch in India's electricity supply. We find that the observed dispatch of power plants in India results in annual total variable costs that are \$4.7 billion higher than the costs implied by least-cost dispatch. As a point of comparison, Cicala (2017) estimates that a shift from traditional command-and-control dispatch to market-based dispatch in the U.S. reduced electricity generating costs by \$3 billion per year. This suggests that the costs gaps we calculated for the Indian electricity sector are neither unreasonably large nor economically insignificant.

In Section 3.5, we find that after accounting for physical transmission constraints, a conservative estimate of the short-run misallocation wedge is \$3.2 billion per year. Below, we investigate three potential drivers that might explain the misallocation we observe: (1) the exercise of market power by electricity suppliers; (2) political factors such as favoritism or changes in political representation; and (3) long-term electricity supply contracts which could lock in electricity production at certain plants while market conditions change.

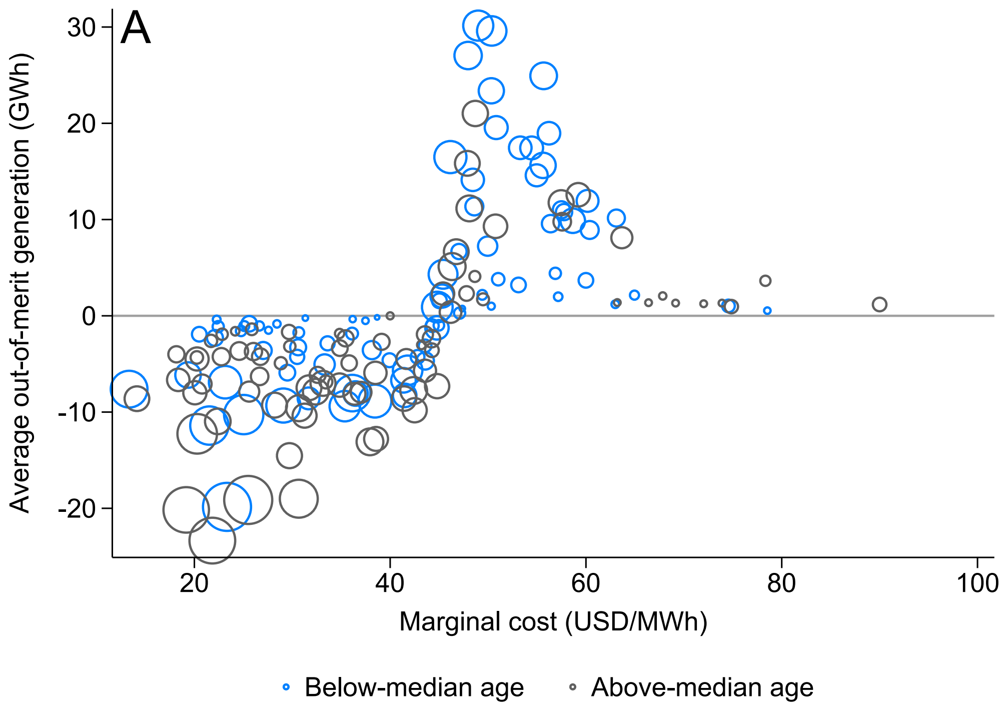
3.6.1 Market Power

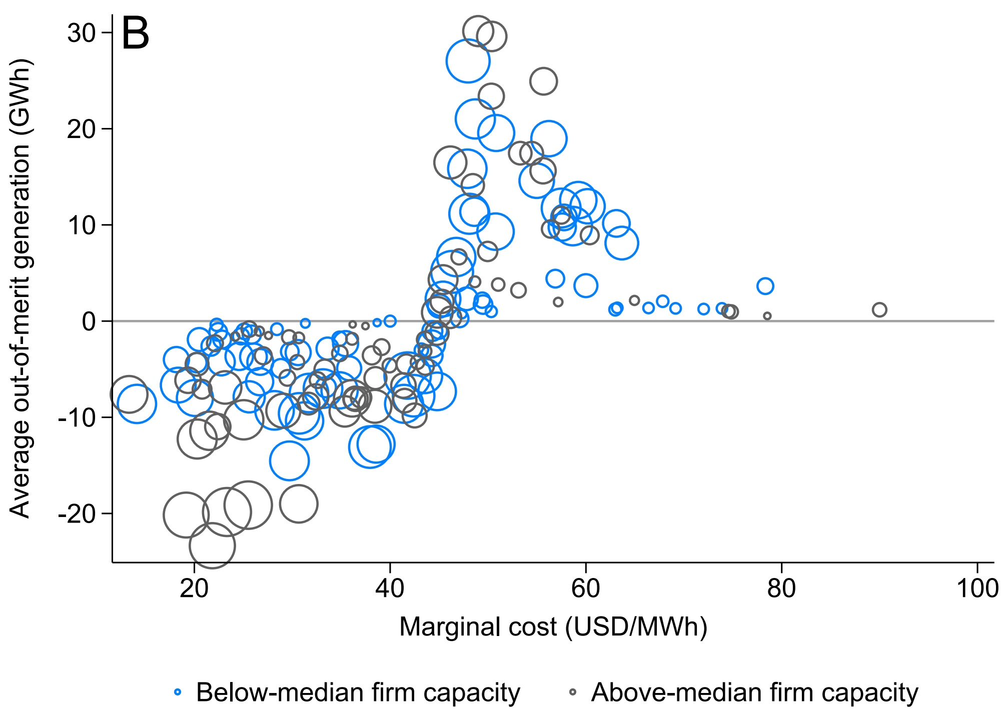
First, we consider the extent to which firms might be exercising market power, which is common in wholesale electricity markets even when supply is not heavily concentrated among a few firms (Borenstein (2002)). In the Indian context, Ryan (2017) estimates average quantity-weighted markups of roughly 20 percent above marginal costs, for electricity sold on the short-run day-ahead market during 2009–2010.[23] However, market power is unlikely to explain a significant share of the short-run cost gap, for two reasons. First, roughly 90 percent of electricity is sold on medium- and long-term contracts with fixed tariffs, which preclude the exercise of short-run market power. Second, recent transmission capacity expansions have likely decreased firms' abilities to exercise market power, by effectively increasing the number of plants able to supply electricity to a given location of demand (Borenstein, Bushnell, and Stoft (2000); Ryan (2017)).

Figure 3.6.10 plots deviations from least-cost generation against plants' marginal cost, splitting plants by vintage (Panel A), total capacity of their owning entities (Panel B), and ownership type (Panel C). If market power were a major driver of the cost gaps we calculate, we would expect to see plants with above-median firm capacity to withhold output from their relatively high-cost plants in order to drive up market prices and earn greater inframarginal rents form their relatively low-cost plants. However, Panel B reveals that out-of-merit generation for high-cost plants is comparable across large vs. small firms; if anything, larger firms appear to operate their low-cost (inframarginal) plants relatively *less* than small firms. This suggests that market power is unlikely to be driving the bulk of the cost gap between observed and least-cost dispatch.

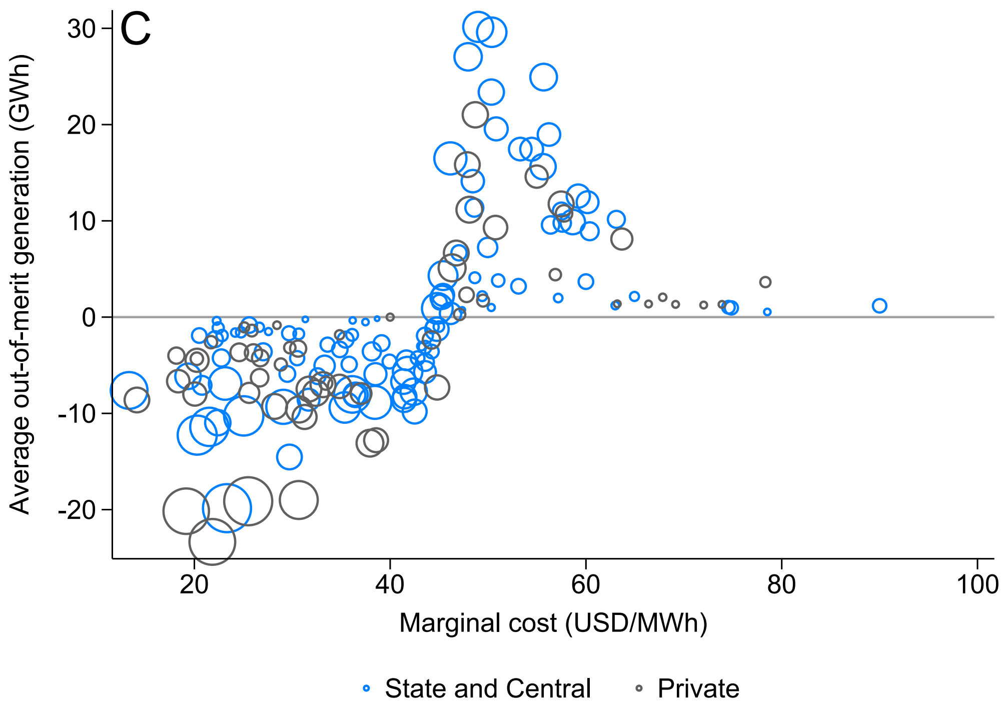
[23]In levels, Ryan (2017) estimates an average short-run price-cost markup of roughly \$8.5/MWh.

Figure 3.6.10: Out-of-merit Generation – Heterogeneity Analysis





(Figured continued on next page.)



Notes: This figure shows the relationship between out-of-merit generation and marginal costs in dollars per MWh, weighed by capacity (as in Figure 3.5.7). Here, we explore heterogeneity in out-of-merit generation. Panel A shows heterogeneity by plant age. Blue dots show plants below the median age, built after 2007; and gray dots show older plants, built in or before 2007. We use the age of the average unit in the plant to define plant age. We see no systematic difference in the relationship between marginal cost and out-of-merit generation between older and newer plants. Panel B demonstrates a preliminary test for market power: are (high-cost) plants owned by larger firms (gray dots) more likely to generate out of merit than plants owned by smaller firms (blue dots)? We again see no discernible patterns. Finally, Panel C shows examines out-of-merit generation by firm type: government run (central and state, blue dots) versus privately owned (gray dots). Ownership does not appear to predict out-of-merit generation. The same is true if we split the state and central sector into their own categories.

3.6.2 Political Economy Factors

Next, we consider political factors, which may induce departures from least-cost dispatch. Political influence might take many forms. For example, power plants with close connections to politicians or political parties may benefit from preferential treatment in electricity dispatch or receive lower-priced fuel supply agreements. This has the potential to disproportionately advantage government-owned plants, compared to privately owned plants. Alternatively, privately owned plants might be able to more easily purchase political influence (e.g. with bribes). Panel C of Figure 3.6.10 reveals no systematic differences between government-owned (state and central) and privately owned plants. Based on this crude heuristic, it appears unlikely that out-of-merit generation persists due to any structural advantages enjoyed by government-owned plants.

However, we may expect electoral outcomes to induce changes in out-of-merit dispatch, as electricity provision is a channel through which politicians may garner support or impart economic benefits. Baskaran, Min, and Uppal (2015) find that Indian state governments appear to manipulate the flow of electricity to key constituencies around the time of close elections. Using a close election regression discontinuity (RD) design, Asher and Novosad (2017) demonstrate that Indian constituencies represented by a politician in the ruling party benefit from political favoritism, as politicians weaken the implementation of regulations faced by firms. We employ a similar RD strategy to test whether out-of-merit generation increases for plants located in constituencies where the ruling party barely wins an election, compared to plants located in constituencies where the ruling party barely loses an election.

Following Asher and Novosad (2017), we construct the RD running variable as the electoral margin from the most recent election, for constituency c in month m, as:

$$\text{margin}_{cm} = \frac{\text{votes}_{cm}^{RULING} - \text{votes}_{cm}^{OPPOSITION}}{\text{votes}_{cm}^{TOTAL}} \tag{3.9}$$

Because Indian elections often comprise multiple parties, $\text{margin}_{cm} > 0$ does not guarantee that the ruling party (at the state level) is also the winning party (at the local level). Hence, we implement a fuzzy RD design:

$$Y_{icm} = \beta_0 + \beta_1 \,\text{margin}_{cm} + \beta_2 \,\widehat{D}_{icm} + \beta_3 \,\text{margin}_{cm} \cdot \mathbf{1}[\text{margin}_{cm} > 0] + \varepsilon_{icm} \tag{3.10}$$

$$\begin{aligned} D_{icm} = \gamma_0 + \gamma_1 \,\text{margin}_{cm} + \gamma_2 \,\mathbf{1}[\text{margin}_{cm} > 0] + \dots \\ + \gamma_3 \,\text{margin}_{cm} \cdot \mathbf{1}[\text{margin}_{cm} > 0] + \omega_{icm} \end{aligned} \tag{3.11}$$

The treatment indicator D_{icm} equals 1 if the ruling party won the most recent election in constituency c, which is home to power plant i. We restrict our estimation sample to the three months m after an election in constituency c, in order to isolate effects immediately after a close election victory. The outcome variable Y_{icm} is plant i's cumulative deviation from least-cost dispatch across all days in month m, residualized to control for plant fixed effects and month-of-sample fixed effects.[24] We estimate local linear polynomials in the running variable on each side of the RD threshold (i.e. $\text{margin}_{cm} = 0$), using the robust RD estimation procedure developed by Calonico, Cattaneo, and Titiunik (2014).

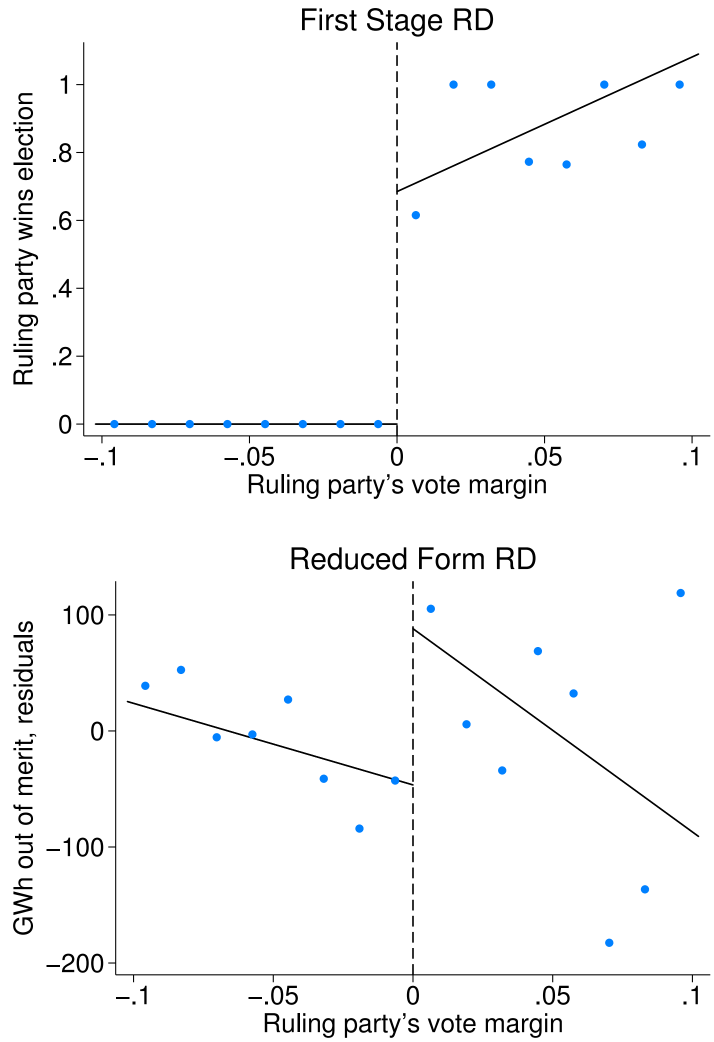
Figure 3.6.11 reports these RD results, with an optimal bandwidth of $[-0.10, 0.10]$. The left panel illustrates how $D_{icm} = 0$ when the running variable is negative, as the ruling party cannot win without a plurality of votes. When the running variable is positive, the ruling party typically (but not always) wins the local election. The right panel plots the reduced form results, where (residualized) out-of-merit generation appears to increase discontinuously at the RD threshold. Table 3.6.2 reports these results numerically. The first-stage point estimate is highly statistically significant, and the rescaled second-stage point estimate is statistically significant with a p-value of 0.02. The magnitude of this effect is quite large—248 GWh is equal to 1.71 (0.83) standard deviations of residualized (raw) out-of-merit generation at the plant-month level. While these results do no persist beyond 3–6 months after close elections, they provide suggestive evidence that political economy factors such as favoritism or corruption might explain a substantial share of India's out-of-merit generation.

3.6.3 Long-term Contracts

Finally, we consider medium- and long-term contracts, which represent roughly 90 percent of India's electricity production during 2016–2017. If high-cost units are generating out of merit in order to satisfy long-term contract obligations, then observed departures from least-cost dispatch may not necessarily reflect allocative inefficiencies. Seemingly inefficient long-term contracts may in fact be economically efficient—reflecting rational, market-based preferences or hedging against market risk. Contracts resulting in out-of-merit generation may also reflect regulatory inefficiencies in the dispatch or contract approval processes. As we discuss in Section 3.2, the National Load Despatch Centre (NLDC) favors long- and medium-term contracts over short-term electricity supply when choosing the dispatch order. The Central Electricity Regulatory Commission (CERC) must also approve electricity prices negotiated between state-owned generation companies and power purchasers. If plants were granted greater flexibility to arbitrage between their long-term contract positions and the day-ahead electricity spot market, this would likely yield substantial economic benefits by improving the short-run allocative efficiency of physical generation outcomes.

[24] We residualize Y_{icm} prior to estimating Equations (3.10)–(3.11) in order to utilize the full time series of data for each plant (including months that do not immediately follow a close election in plant i's constituency), and the full cross-section of plants in each month (including plants that did not experience a close election immediately prior to month m). This greatly improves the statistical power of our RD estimates. We cluster standard errors by constituency, with corrects the degrees of freedom for the purposes of inference.

Figure 3.6.11: Out-of-merit Generation after Close Elections



Notes: These figure reports RD results from estimating Equations (3.10)–(3.11), using the running variable margin$_{cm}$ defined in Equation (3.9), and including the 3 months m immediately following a close election in each constituency c. The left panel plots the first stage RD, where the outcome variable is an indicator equal to 1 if the ruling party (at the state level) won the most recent local election. The right panel plots the reduced-form RD, where the outcome variable is the cumulative GWh of out-of-merit generation for plant i in constituency c in month m, residualized to control for plant fixed effects and month-of-sample fixed effects. We residualize these fixed effects prior to RD estimation in order to increase statistical power. This allows us to utilize the full time series of data for each plant (including months that do not immediately follow a close election in plant i's constituency), and the full cross-section of plants in each month (including plants that did not experience a close election immediately prior to month m). We use the robust RD estimation procedure developed by Calonico, Cattaneo, and Titiunik (2014) with a local linear polynomial for point estimation, a local quadratic polynomial for bias correction, an optimal RD bandwidth of $[-0.10, 0.10]$, weighting based on distance from the RD threshold using a triangular kernel, and standard errors clustered by legislative constituency. Table 3.6.2 reports these same results numerically.

Table 3.6.2: Regression Discontinuity for Close Elections

	GWh out of merit, residuals
Fuzzy RD estimate	247.68^{**}
	(106.64)
	[0.02]
First-stage RD estimate	0.63^{***}
	(0.17)
	[0.00]
Optimal RD bandwidth	0.10
Plant-month observations in bandwidth	666
RD scaled by within std dev	1.71
RD scaled by composite std dev	0.83

Notes: These table reports RD results from estimating Equations (3.10)–(3.11), using the running variable margin_{cm} defined in Equation (3.9), and including the 3 months m immediately following a close election in each constituency c. The treatment indicator is equal to 1 if the ruling party (at the state level) won the most recent local election. The outcome variable is the cumulative GWh of out-of-merit generation for plant i in constituency c in month m, residualized to control for plant fixed effects and month-of-sample fixed effects. We residualize these fixed effects prior to RD estimation in order to increase statistical power. This allows us to utilize the full time series of data for each plant (including months that do not immediately follow a close election in plant i's constituency), and the full cross-section of plants in each month (including plants that did not experience a close election immediately prior to month m). We use the robust RD estimation procedure developed by Calonico, Cattaneo, and Titiunik (2014) with a local linear polynomial for point estimation, a local quadratic polynomial for bias correction, an optimal RD bandwidth of $[-0.10, 0.10]$, weighting based on distance from the RD threshold using a triangular kernel, and standard errors clustered by legislative constituency (with p-values in brackets). The 247.68 point estimate is scaled by the first stage, while the bottom two rows compare this point estimate to the standard deviations of the outcome variable and the unresidualized outcome variable (respectively). Figure 3.6.11 reports these same results graphically. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

3.7 Conclusion

Economic development is strongly linked to increased energy consumption. Inefficient electricity market operations, particularly in low- and middle-income countries, may therefore hamper economic growth and impose undue burdens on poor consumers. In this paper, we assemble a novel dataset on daily electricity market operations in India, which we use to quantify the cost of electricity supply in a major developing-country power market. We calculate the total short-run variable costs of Indian electricity generation to be approximately 29 billion U.S. dollars per year, or $36 per MWh, on average.

We compare these observed costs of electricity supply to a counterfactual where power plants generate electricity in order of lowest-to-highest cost, and we find that this "least-cost" dispatch would decrease variable costs of electricity generation by roughly 4.7 billion U.S. dollars per year, or nearly 17 percent. After conservatively accounting for both interregional and intraregional transmission constraints, the remaining cost gap between observed vs. least-cost dispatch is 3.2 billion U.S. dollars per year, or over 11 percent. This result is striking, as it suggests that contrary to both popular wisdom and existing public policy, expanding interregional transmission capacity in India may not deliver large reductions in the cost of electricity supply. It also represents a substantial short-run misallocation wedge in Indian electricity supply.

We investigate three potential drivers of this misallocation: (1) market power; (2) political economy factors; and (3) long-term contracts. We find little evidence of market power, which is unsurprising given that rate-of-return regulated power plants cannot earn inframarginal rents. We find evidence that power plants are more likely to generate out of merit after the ruling political party wins a local election. Given that roughly 90 percent of electricity generation occurs on long-term contracts, allowing power plants to more flexibly arbitrage the day-ahead spot market would likely alleviate a substantial share of the short-run misallocation we are finding.

In future work, we plan to utilize plant-level data on planned versus unplanned outages in order to control for extended maintenance periods. This will facilitate more nuanced least-cost counterfactuals based on the subset of plants that *could have* generated on a given day. We also plan to incorporate data on upstream fuel markets to investigate whether upstream inefficiencies are contributing to misallocation in the downstream electricity market. Finally, we hope to study power plant emissions and air pollution concentrations, in order to quantify how misallocation impacts pollution concentrations and their associated damages on human health and the environment.

References

Abeberese, Ama Baafra. Forthcoming. "Electricity Cost and Firm Performance: Evidence from India." *Review of Economics and Statistics.*

Alam, Muneeza M. 2013. "Coping with Blackouts: Power Outages and Firm Choices." Working Paper.

Alexandrov, Alexei, Russell Pittman, and Olga Ukhaneva. 2017. "Royalty Stacking in the U.S. Freight Railroads: Cornout vs. Coase." Working Paper.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell. 2016. "How Do Electricity Shortages Affect Industry? Evidence from India." *American Economic Review* 106 (3): 587–624.

Asher, Sam, and Paul Novosad. 2017. "Politics and Local Economic Growth: Evidence from India." *American Economic Journal: Applied Economics* 9 (1): 229–273.

———. 2018. "Rural Roads and Local Economic Development." Working Paper.

Atkin, David, and Dave Donaldson. 2015. "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs." NBER Working Paper 21439.

Atkinson, Scott E., and Joe Kerkvliet. 1986. "Measuring the Multilateral Allocation of Rents: Wyoming Low-Sulfur Coal." *RAND Journal of Economics* 17 (3): 416–430.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." NBER Working Paper 17897.

Banerjee, Sudeshna Ghosh, Douglas Barnes, Bipul Singh, Kristy Mayer, and Hussain Samad. 2014. *Power for All: Electricity Access Challenge in India.* Washington, DC: The World Bank.

Barnett, A. H. 1980. "The Pigouvian Tax Rule under Monopoly." *American Economic Review* 70 (5): 1037–1041.

Barron, Manuel, and Maximo Torero. 2017. "Household Electrification and Indoor Air Pollution." *Journal of Environmental Economics and Management* 86:81–92.

Baskaran, Thushyanthan, Brian Min, and Yogesh Uppal. 2015. "Election Cycles and Electricity Provision: Evidence from a Quasi-experiment with Indian Special Elections." *Journal of Public Economics* 126 (1): 64–73.

Bergquist, Lauren Falcao. 2017. "Pass-Through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya." Working Paper.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127 (2): 587–644.

Borenstein, Severin. 2002. "The Trouble with Electricity Markets: Understanding California's Restructuring Disaster." *Journal of Economic Perspectives* 16 (1): 191–211.

Borenstein, Severin, James B. Bushnell, and Frank A. Wolak. 2002. "Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market." *American Economic Review* 92 (5): 1376–1405.

Borenstein, Severin, James Bushnell, and Steven Stoft. 2000. "The Competitiveness Effects of Transmission Capacity in a Deregulated Electricity Industry." *RAND Journal of Economics* 31 (2): 294–325.

Borenstein, Severin, and Ryan Kellogg. 2014. "The Incidence of an Oil Glut: Who Benefits from Cheap Crude Oil in the Midwest?" *Energy Journal* 35 (1): 15–33.

Brehm, Paul. 2017. "Natural Gas Prices, Electric Generation Investment, and Greenhouse Gas Emissions." Working Paper.

Buchanan, James M. 1969. "External Diseconomies, Corrective Taxes, and Market Structure." *American Economic Review* 59 (1): 174–177.

Burlig, Fiona, and Louis Preonas. 2016. "Out of the Darkness and Into the Light? Development Impacts of Rural Electrification." Energy Institute at Haas Working Paper No. 268.

Bushnell, James B., Erin T. Mansur, and Celeste Saravia. 2008. "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured U.S. Electricity Markets." *American Economic Review* 98 (1): 237–266.

Busse, Meghan R., and Nathaniel O. Keohane. 2007. "Market Effects of Environmental Regulation: Coal, Railroads, and the 1990 Clean Air Act." *RAND Journal of Economics* 38 (4): 1159–1179.

Callaway, Duncan S., Meredith Fowlie, and Gavin McCormick. 2018. "Location, Location, Location: The Variable Value of Renewable Energy and Demand-Side Efficiency Resources." *Journal of the Association of Environmental and Resource Economists* 5 (1): 39–75.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Design.” *Econometrica* 82 (6): 2295–2326.

Cameron, A. Colin, and Douglas L. Miller. 2015. “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources* 50 (2): 317–372.

Carl, Jeremy, and David Fedor. 2016. “Tracking Global Carbon Revenues: A Survey of Carbon Taxes Versus Cap-and-Trade in the Real World.” *Energy Policy* 96:50–77.

Census of India. 2011. “Mapping Activities.” `http://censusindia.gov.in/2011-common/map.html`.

Central Electricity Authority. 2011. *Load Generation Balance Report: 2011–12.*

Chakravorty, Ujjayant, Kyle Emerick, and Majah-Leah Ravago. 2016. “Lighting Up the Last Mile: The Benefits and Costs of Extending Electricity to the Rural Poor.” Working Paper.

Chan, H. Ron, Harrison Fell, Ian Lange, and Shanjun Li. 2017. “Efficiency and Environmental Impacts of Electricity Restructuring on Coal-fired Power Plants.” *Journal of Environmental Economics and Management* 81:1–18.

Chan, Hei Sing Ron, Maureen L. Cropper, and Kabir Malik. 2014. “Why Are Power Plants in India Less Efficient than Power Plants in the United States?” *American Economic Review* 104 (5): 586–590.

Chand, T. R. Kiran, K. V. S. Badarinath, Christopher D. Elvidge, and B. T. Tuttle. 2009. “Spatial Characterization of Electrical Power Consumption Patterns over India Using Temporal DMSP–OLS Night-time Satellite Data.” *International Journal of Remote Sensing* 30 (3): 647–661.

Chen, Xi, and William D. Nordhaus. 2011. “Using Luminosity Data as a Proxy for Economic Statistics.” *Proceedings of the National Academy of Sciences* 108 (21): 8589–8594.

Christian, Paul, and Christopher B. Barrett. 2017. “Revisiting the Effect of Food Aid on Conflict: A Methodological Caution.” World Bank Policy Research Working Paper WPS8171.

Chu, Yin, J. Scott Holladay, and Jacob LaRiviere. Forthcoming. “Pass-Through from Fossil Fuel Market Prices to Procurement Costs of the U.S. Power Producers.” *Jounal of Industrial Economics.*

Cicala, Steve. 2015. “When Does Regulation Distort Costs? Lessons from Fuel Procurement in U.S. Electricity Generation.” *American Economic Review* 105 (1): 411–444.

———. 2017. “Imperfect Markets Versus Imperfect Regulation in U.S. Electricity Generation.” NBER Working Paper 23053.

Conley, T. G. 1999. “GMM Estimation with Cross Sectional Dependence.” *Journal of Econometrics* 92 (1): 1–45.

Conley, Timothy G. 2008. “Spatial Econometrics.” In *The New Palgrave Dictionary of Economics,* 2nd ed., edited by Steven N. Durlauf and Lawrence E. Blume, 741–747. Basingstoke: Nature Publishing Group.

Corts, Kenneth S. 1999. “Conduct Parameters and the Measurement of Market Power.” *Journal of Econometrics* 88 (2): 227–250.

Cournot, Augustin. 1838. “Of the Influence of Taxation on Commodities Produced under a Monopoly.” In *Researches into the Mathematical Principles of the Theory of Wealth,* translated by Nathaniel T. Bacon, 67–78. London: Macmillan, 1897.

Covert, Thomas R., and Ryan Kellogg. 2017. “Crude by Rail, Option Value, and Pipeline Investment.” NBER Working Paper 23855.

Cullen, Joseph A. 2015. “Dynamic Response to Environmental Regulation in the Electricity Industry.” Working Paper.

Cullen, Joseph A., and Erin T. Mansur. 2017. “Inferring Carbon Abatement Costs in Electricity Markets: A Revealed Preference Approach Using the Shale Revolution.” *American Economic Journal: Economic Policy* 9 (3): 106–133.

Cullen, Joseph A., and Stanley S. Reynolds. 2016. “The Long Run Impact of Environmental Policies on Wholesale Electricity Markets: A Dynamic Competitive Analysis.” Working Paper.

Davis, Lucas W., and Erich Muehlegger. 2010. “Do Americans Consume Too Little Natural Gas? An Empirical Test of Marginal Cost Pricing.” *RAND Journal of Economics* 41 (4): 791–810.

Davis, Lucas W., and Catherine Wolfram. 2012. “Deregulation, Consolidation, and Efficiency: Evidence from U.S. Nuclear Power.” *American Economic Journal: Applied Economics* 4 (4): 194–225.

Davis, Lucas, and Catherine Hausman. 2016. “Market Impacts of a Nuclear Power Plant Closure.” *American Economic Journal: Applied Economics* 8 (2): 92–122.

Dinkelman, Taryn. 2011. “The Effects of Rural Electrification on Employment: New Evidence from South Africa.” *American Economic Review* 101 (7): 3078–3108.

Doll, Christopher N. H. 2008. *CIESIN Thematic Guide to Night-time Light Remote Sensing and its Applications.* Center for Earth Science Information Network (CIESIN).

Doll, Christopher N. H., Jan-Peter Muller, and Jeremy G. Morley. 2006. “Mapping Regional Economic Activity from Night-time Light Satellite Imagery.” *Ecological Economics* 57 (1): 75–92.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108 (4–5): 899–934.

Ebener, Steeve, Christopher Murray, Ajay Tandon, and Christopher D. Elvidge. 2005. "From Wealth to Health: Modeling the Distribution of Income per Capita at the Sub-national Level Using Night-time Light Imagery." *International Journal of Health Geographics* 4 (1): 5.

Elvidge, Christopher D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis. 1997. "Relation between Satellite Observed Visible-near Infrared Emissions, Population, Economic Activity and Electric Power Consumption." *International Journal of Remote Sensing* 18 (6): 1373–1379.

Elvidge, Christopher D., Marc L Imhoff, Kimberly E Baugh, Vinita Ruth Hobson, Ingrid Nelson, Jeff Safran, John B Dietz, and Benjamin T Tuttle. 2001. "Night-time lights of the world: 1994–1995." *ISPRS Journal of Photogrammetry and Remote Sensing* 56 (2): 81–99.

Faber, Benjamin. 2014. "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System." *Review of Economic Studies* 83 (2): 1046–1070.

Fabra, Natalia, and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *American Economic Review* 104 (9): 2872–2899.

Fabrizio, Kira R., Nancy L. Rose, and Catherine D. Wolfram. 2007. "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on U.S. Electric Generation Efficiency." *American Economic Review* 97 (4): 1250–1277.

Fell, Harrison, and Daniel T. Kaffine. Forthcoming. "The Fall of Coal: Joint Effects of Fuel Prices and Renewables on Generation and Emissions." *American Economic Journal: Economic Policy.*

Fetzer, Thiemo. 2014. "Social Insurance and Conflict: Evidence from India." Working Paper.

Fitzgerald, Timothy. 2013. "Frackonomics: Some Economics of Hydraulic Fracturing." *Case Western Reserve Law Review* 63 (4): 1337–1362.

Fowlie, Meredith. 2010. "Emissions Trading, Electricity Restructuring, and Investment in Pollution Abatement." *American Economic Review* 100 (3): 837–869.

Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan. 2016. "Market-Based Emissions Regulation and Industry Dynamics." *Journal of Political Economy* 124 (1): 249–302.

Ganapati, Sharat. 2017. "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies." Working Paper.

Ganapati, Sharat, Joseph S. Shapiro, and Reed Walker. 2016. "The Incidence of Carbon Taxes in U.S. Manufacturing: Lessons from Energy Cost Pass-Through." NBER Working Paper 22281.

Gelman, Andrew, and Guido Imbens. 2017. "Why High-order Polynomials Should not be Used in Regression Discontinuity Designs." *Journal of Business & Economics Statistics.*

Gertler, Paul, Orie Shelef, Catherine Wolfram, and Alan Fuchs. 2016. "The Demand for Energy-Using Assets among the World's Rising Middle Classes." *American Economic Review* 106 (6): 1366–1401.

Gillingham, Kenneth, James Bushnell, Meredith Fowlie, Michael Greenstone, Charles Kolstad, Alan Krupnick, Adele Morris, Richard Schmalensee, and James Stock. 2016. "Reforming the U.S. Coal Leasing Program." *Science* 354 (6316): 1096–1098.

Hastings, Justine S. 2004. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California." *American Economic Review* 94 (1): 317–328.

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas." *Brooking Papers on Economic Activity* Spring:71–125.

He, Qingxin, and Jonathan M. Lee. 2016. "The Effect of Coal Combustion Byproducts on Price Discrimination by Upstream Industries." *International Journal of Industrial Organization* 44:11–24.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64 (4): 605–654.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028.

Holladay, J. Scott, and Jacob LaRiviere. 2017. "The Impact of Cheap Natural Gas on Marginal Emissions from Electricity Generation and Implications for Energy Policy." *Journal of Environmental Economics and Management* 85:205–227.

Hortaçsu, Ali, and Chad Syverson. 2007. "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices." *Journal of Political Economy* 115 (2): 250–301.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *Proceedings of the National Academy of Sciences* 107 (35): 15367–15372.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–1448.

Hughes, Jonathan E. 2011. "The Higher Price of Cleaner Fuels: Market Power in the Rail Transport of Fuel Ethanol." *Journal of Environmental Economics and Management* 62 (2): 123–139.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615–635.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79 (3): 933–959.

Interagency Working Group on Social Cost of Greenhouse Gases. 2016. "Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis." United States Government, Technical Support Document.

International Energy Agency. 2011. *World Energy Outlook 2011: Energy for All.*

Jamasb, Tooraj, Rabindra Nepal, and Govinda R. Timilsina. 2015. "A Quarter Century Effort Yet to Come of Age: A Survey of Power Sector Reforms in Developing Countries." World Bank Policy Research Working Paper 7330.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas. 2017. "Cash for Carbon: A Randomized Controlled Trial of Payments for Ecosystem Services to Reduce Deforestation." *Science* 357 (6348): 267–273.

Jha, Akshaya. 2015. "Dynamic Regulatory Distortions: Coal Procurement at U.S. Power Plants." Working Paper.

———. 2017. "Regulatory Induced Risk Aversion: Coal Procurement at U.S. Power Plants." Working Paper.

Joskow, Paul L. 1985. "Vertical Integration and Long-term Contracts: The Case of Coal-burning Electric Generating Plants." *Journal of Law, Economics, & Organization* 1 (1): 33–80.

———. 1987. "Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets." *American Economic Review* 77 (1): 168–185.

———. 1988. "Price Adjustment in Long-term Contracts: The Case of Coal." *Journal of Law and Economics* 31 (1): 47–83.

———. 2006. "Competitive Electricity Markets and Investment in New Generating Capacity." Mimeo.

———. 2008. "Capacity Payments in Imperfect Electricity Markets: Need and Design." *Utilities Policy* 16 (3): 159–170.

Katz, Michael L. 1987. "The Welfare Effects of Third-degree Price Discrimination in Intermediate Good Markets." *American Economic Review* 77 (1): 154–167.

Kerkvliet, Joe, and Jason F. Shogren. 1992. "The Impacts of Environmental Regulation on Coal Procurement Strategies: Design Coal and Multi-attributed Quality." *Journal of Environmental Management* 35 (2): 83–91.

Kim, Wook, Deb Chattopadhyay, and Jong-bae Park. 2010. "Impact of Carbon Cost on Wholesale Electricity Price: A Note on Price Pass-through Issues." *Energy* 35 (8): 3441–3448.

Knittel, Christopher R., Ben S. Meiselman, and James H. Stock. 2017. "The Pass-Through of RIN Prices to Wholesale and Retail Fuels under the Renewable Fuel Standard." *Journal of the Association of Environmental and Resource Economists* 4 (4): 1081–1119.

Knittel, Christopher R., Konstantinos Metaxoglou, and Andre Trindade. 2015. "Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets." NBER Working Paper 21627.

Knittel, Christopher, Konstantinos Metaxoglou, and Andre Trindade. 2016. "Are We Fracked? The Impact of Falling Gas Prices and the Implications for Coal-to-Gas Switching and Carbon Emissions." *Oxford Review of Economic Policy* 32 (2): 241–259.

Kolstad, Charles D., and Frank A. Wolak. 1983. "Competition in Interregional Taxation: The Case of Western Coal." *Journal of Political Economy* 91 (3): 443–460.

Kosnik, Lea, and Ian Lange. 2011. "Contract Renegotiation and Rent Re-distribution: Who Gets Raked over the Coals?" *Journal of Environmental Economics and Management* 62 (2): 155–165.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2016. "Experimental Evidence on the Demands for and Costs of Rural Electrification." NBER Working Paper 22292.

Levy, Jonathan I., Lisa K. Baxter, and Joel Schwartz. 2009. "Uncertainty and Variability in Health-Related Damages from Coal-Fired Power Plants in the United States." *Risk Analysis* 29 (7): 1000–1014.

Li, Jing, and James H. Stock. 2017. "Cost Pass-Through to Higher Ethanol Blends at the Pump: Evidence from Minnesota Gas Station Data." Working Paper.

Li, Xi, Linlin Ge, and Xiaoling Chen. 2013. "Detecting Zimbabwe's Decadal Economic Decline Using Nighttime Light Imagery." *Remote Sensing* 5 (9): 4551–4570.

Linn, Joshua, Erin Mastrangelo, and Dallas Burtraw. 2014. "Regulating Greenhouse Gases from Coal Power Plants under the Clean Air Act." *Journal of the Association of Environmental and Resource Economists* 1 (1/2): 97–134.

Linn, Joshua, and Lucija Muehlenbachs. Forthcoming. "The Heterogeneous Impacts of Low Natural Gas Prices on Consumers and the Environment." *Journal of Environmental Economics and Management.*

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5 (2): 200–231.

Lobao, Linda, Minyu Zhou, Mark Partridge, and Michael Betz. 2016. "Poverty, Place, and Coal Employment across Appalachia and the United States in a New Economic Era: Poverty, Place, and Coal Employment." *Rural Sociology* 81 (3): 343–386.

MacDonald, James M. 1989. "Railroad Deregulation, Innovation, and Competition: Effects of the Staggers Act on Grain Yransportation." *Journal of Law and Economics* 32 (1): 63–95.

———. 2013. "Railroads and Price Discrimination: The Roles of Competition, Information, and Regulation." *Review of Industrial Organization* 43 (1–2): 85–101.

Mansur, Erin T. 2007. "Do Oligopolists Pollute Less? Evidence from a Restructured Electricity Market." *Journal of Industrial Economics* 55 (4): 661–689.

———. 2008. "Measuring Welfare in Restructured Electricity Markets." *Review of Economics and Statistics* 90 (2): 369–386.

———. 2013. "Prices versus Quantities: Environmental Regulation and Imperfect Competition." *Journal of Regulatory Economics* 44 (1): 80–102.

Marion, Justin, and Erich Muehlegger. 2011. "Fuel Tax Incidence and Supply Conditions." *Journal of Public Economics* 95 (9–10): 1202–1212.

McRae, Shaun. 2015. "Infrastructure Quality and the Subsidy Trap." *American Economic Review* 105 (1): 35–66.

Meng, Kyle C. 2016. "Estimating Path Dependence in Energy Transitions." NBER Working Paper 22536.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81 (1): 113–152.

———. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics* 129 (1): 151–213.

Min, Brian. 2011. "Electrifying the Poor: Distributing Power in India." Working Paper.

Min, Brian, and Kwawu Mensan Gaba. 2014. "Tracking Electrification in Vietnam Using Nighttime Lights." *Remote Sensing* 6 (10): 9511–9529.

Min, Brian, Kwawu Mensan Gaba, Ousmane Fall Sarr, and Alassane Agalassou. 2013. "Detection of Rural Electrification in Africa Using DMSP-OLS Night Lights Imagery." *International Journal of Remote Sensing* 34 (22): 8118–8141.

Ministry of Law and Justice. 2003. *The Electricity Act, 2003.*

Ministry of Power. 2005a. *National Electricity Policy.*

———. 2005b. *Rajiv Gandhi Grameen Vidyutikaran Yojna Scheme of Rural Electricity Infrastructure and Household Electrification.*

———. 2008. *Continuation of Rajiv Gandhi Grameen Vidyutikaran Yojana in the XI Plan-Scheme of Rural Electricity* File No. 44/37/07-D(RE). New Delhi.

———. 2012. *Rajiv Gandhi Grameen Vidyutikaran Yojana.*

———. 2013. *Continuation of Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) in 12th & 13th Plan - Scheme of Rural Electricity Infrastructure and Household Electrification.*

———. 2014a. *Deendayal Upadhyaya Gram Jyoti Yojana (DDUGJY).*

———. 2014b. *Guidelines for Preparation of DPRs under XII Plan of RGGVY.*

———. 2015. *Rajya Sabha Unstarred Question No. 1622.*

Montgomery, W. David. 1972. "Markets in Licenses and Efficient Pollution Control Programs." *Journal of Economic Theory* 5 (3): 395–418.

Muehlegger, Erich, and Richard L. Sweeney. 2017. "Competition and Pass-Through of Input Cost Shocks: Evidence from the U.S. Fracking Boom." NBER Working Paper 24025.

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

Nazifi, Fatemeh. 2016. "The Pass-Through Rates of Carbon Costs on to Electricity Prices within the Australian National Electricity Market." *Environmental Economics and Policy Studies* 18 (1): 41–62.

Noor, Abdisalan M., Victor A. Alegana, Peter W. Gething, Andrew J. Tatem, and Robert W. Snow. 2008. "Using Remotely Sensed Night-time Light as a Proxy for Poverty in Africa." *Population Health Metrics* 6 (1): 5.

Nordhaus, William D. 1993. "Reflections on the Economics of Climate Change." *Journal of Economic Perspectives* 7 (4): 11–25.

Oates, Wallace E., and Diana L. Strassmann. 1984. "Effluent Fees and Market Structure." *Journal of Public Economics* 24 (1): 29–46.

Olley, G. Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry.” *Econometrica* 64 (6): 1263–1297.

Pigou, Arthur C. 1932. *The Economics of Welfare, 1920.* London: Macmillan.

Pouliot, Sebastien, Aaron Smith, and James H. Stock. 2017. “RIN Pass-Through at Gasoline Terminals.” Working Paper.

Prater, Marvin, Adam Sparger, and Daniel O’Neil Jr. 2014. *Railroad Concentration, Market Shares, and Rates.* Technical report. U.S. Department of Agriculture, Agricultural Marketing Service.

Reguant, Mar. 2014. “Complementary Bidding Mechanisms and Startup Costs in Electricity Markets.” *Review of Economic Studies* 81 (4): 1708–1742.

Revesz, R., M. Greenstone, M. Hanemann, M. Livermore, T. Sterner, D. Grab, P. Howard, and J. Schwartz. 2017. “Best Cost Estimate of Greenhouse Gases.” *Science* 357 (6352): 655.

Rud, Juan Pablo. 2012. “Electricity Provision and Industrial Development: Evidence from India.” *Journal of Development Economics* 97 (2): 352–367.

Rural Electrification Corporation. 2010. *Rural Electrification Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY).*

Ryan, Nicholas. 2017. “The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market.” NBER Working Paper 23106.

Ryan, Stephen P. 2012. “The Costs of Environmental Regulation in a Concentrated Industry.” *Econometrica* 80 (3): 1019–1061.

Schmalensee, Richard. 1981. “Output and Welfare Implications of Monopolistic Third-degree Price Discrimination.” *American Economic Review* 71 (1): 242–247.

Schmalensee, Richard, and Robert N Stavins. 2013. “The SO_2 Allowance Trading System: The Ironic History of a Grand Policy Experiment.” *Journal of Economic Perspectives* 27 (1): 103–122.

Schmidt, Stephen. 2001. “Market Structure and Market Outcomes in Deregulated Rail Freight Markets.” *International Journal of Industrial Organization* 19 (1): 99–131.

Startz, Meredith. 2016. “The Value of Face-to-face: Search and Contracting Problems in Nigerian Trade.” Working Paper.

Stolper, Samuel. 2016. “Competition and Incidence: Automotive Fuel Tax Pass-Through at State Borders.” Working Paper.

———. 2017. “Who Bears the Burdern of Energy Taxes? The Role of Local Pass-Through.” Working Paper.

Townsend, Alexander C., and David A. Bruce. 2010. "The Use of Night-time Lights Satellite Imagery as a Measure of Australia's Regional Electricity Consumption and Population Distribution." *International Journal of Remote Sensing* 31 (16): 4459–4480.

Tsujita, Yuko, ed. 2014. *Inclusive Growth and Development in India: Challenges for Underdeveloped Regions and the Underclass.* United Kingdom: Palgrave Macmillan.

UNDP. 2015. "Energy access." `http://www.undp.org/content/undp/en/home/ourwork/climate-and-disaster-resilience/sustainable-energy/energy-access.html`.

U.S. Government Accountability Office. 2006. *Freight Railroads: Industry Health Has Improved, but Concerns about Competition and Capacity Should Be Addressed.* Report to Congressional Requesters GAO-07-94.

Varian, Hal R. 1985. "Price Discrimination and Social Welfare." *American Economic Review* 75 (4): 870–875.

Weyl, E. Glen, and Michal Fabinger. 2013. "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition." *Journal of Political Economy* 121 (3): 528–583.

Wilson, Wesley W. 1996. "Legislated Market Dominance in Railroad Markets." *Research in Transportation Economics* 4:49–67.

Wolak, Frank A. 1996. "Why Do Firms Simultaneously Purchase in Spot and Contract Markets? Evidence from the United States Steam Coal Market." In *Agricultural Markets: Mechanisms, Failures, and Regulations,* edited by David Martimort, 109–168. Elsevier Science.

———. 2016. "Assessing the Impact of the Diffusion of Shale Oil and Gas Technology on the Global Coal Market." Working Paper.

Wolak, Frank A., and Charles D. Kolstad. 1988. "Measuring Relative Market Power in the Western U.S. Coal Market Using Shapley Values." *Resources and Energy* 10 (4): 293–314.

Wolfram, Catherine, Orie Shelef, and Paul Gertler. 2012. "How Will Energy Demand Develop in the Developing World?" *Journal of Economic Perspectives* 26 (1): 119–138.

Wooldridge, Jeffrey M. 2007. "Inverse Probability Weighted Estimation for General Missing Data Problems." *Journal of Econometrics* 141 (2): 1281–1301.

World Bank. 2015. "Energy Overview." `http://www.worldbank.org/en/topic/energy/overview`.

Appendix A

Appendix: Market Power in Coal Shipping and Implications for U.S. Climate Policy

A.1 Theory

A.1.1 Derivation of Comparative Statics

This section provides a full derivation of the comparative static $\frac{d\mu_{oj}}{dZ_j}$, reported in Equation (1.3) of the main text. I start with rail carrier i's profit function, reproduced here from Equation (1.1):

$$\pi_{ioj}(q_{ioj}) = q_{ioj}\Big[P_{oj}(Q_{oj};\mathbf{Z}_{oj}) - C_o - S(\mathbf{T}_{oj})\Big] - F_{oj} \tag{A.1}$$

The oligopolist rail carrier earns revenue $q_{ioj}P_{oj}$ from selling type-o coal to plant j, while incurring commodity costs $C_o q_{ioj}$, o-to-j shipping costs $S(\mathbf{T}_{oj})q_{ioj}$, and a fixed cost of entry F_{oj}. Plant j's inverse demand is a function of $Q_{oj} = N_{oj}q_{ioj}$, the total quantity of coal purchased across all N_{oj} symmetric oligopolists. It also depends on the parameter vector $\mathbf{Z}_{oj}$, which includes Z_j, the natural gas price of plant-j's competitors in electricity supply.

Firm i's first-order condition is:

$$\frac{\partial \pi_{ioj}}{\partial q_{ioj}} = P_{oj}(Q_{oj};\mathbf{Z}_{oj}) + q_{ioj}\frac{\partial P_{oj}}{\partial Q_{oj}}\frac{\partial Q_{oj}}{\partial q_{ioj}} - C_o - S(\mathbf{T}_{oj}) \tag{A.2}$$

For simplicity, I assume that $S(\mathbf{T}_{oj})$ does not depend on q_{ioj}, which abstracts from rail capacity constraints and increasing returns to scale in shipping.[1] Totally differentiating Equation (A.2) by q_{ioj} and Z_j, and rearranging:[2]

$$\text{(A.3)} \qquad \frac{dq_{ioj}}{dZ_j} = \frac{\dfrac{\partial P_{oj}}{\partial Z_j} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j}\dfrac{\partial Q_{oj}}{\partial q_{ioj}} q_{ioj}}{-\left(2\dfrac{\partial P_{oj}}{\partial Q_{oj}}\dfrac{\partial Q_{oj}}{\partial q_{ioj}} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}^2}\left(\dfrac{\partial Q_{oj}}{\partial q_{ioj}}\right)^2 q_{ioj} + \dfrac{\partial P_{oj}}{\partial Q_{oj}}\dfrac{\partial^2 Q_{oj}}{\partial q_{ioj}^2} q_{ioj}\right)}$$

Invoking symmetry across all N_{oj} oligopolists, I can substitute $q_{ioj} = \frac{Q_{oj}}{N_{oj}}$. I also substitute the "conduct parameter" $\theta_{oj} \equiv \frac{\partial Q_{oj}}{\partial q_{ioj}}$, which simplifies notation and serves as a heuristic for distance from perfect competition.[3] Rewriting (A.3):

$$\text{(A.4)} \qquad \frac{dq_{ioj}}{dZ_j} = \frac{\dfrac{\partial P_{oj}}{\partial Z_j} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j}\dfrac{Q_{oj}\theta_{oj}}{N_{oj}}}{-\left(2\dfrac{\partial P_{oj}}{\partial Q_{oj}}\theta_{oj} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}^2}\dfrac{Q_{oj}\theta_{oj}^2}{N_{oj}} + \dfrac{\partial P_{oj}}{\partial Q_{oj}}\underbrace{\dfrac{\partial \theta_{oj}}{\partial q_{ioj}}}_{=0}\dfrac{Q_{oj}}{N_{oj}}\right)}$$

I assume $\frac{\partial \theta_{oj}}{\partial q_{ioj}} = 0$, because small changes in q_{ioj} are unlikely to change the relationship between carrier i's quantity q_{ioj} and total demand Q_{oj}.[4]

[1] My empirical analysis relaxes this assumption, by allowing rail transport costs to vary with shipment size. Because power plants are small relative to the coal producing sector, I also assume that C_o is independent of q_{ioj}. Even if plant j consumes a large share of the coal produced by county o, coal is substitutable across counties and C_o depends more on coal mining costs and global commodity markets.

[2] This is a direct application of the Implicit Function Theorem, where $\frac{\partial \pi_{ioj}}{\partial q_{ioj}}$ is a function of q_{ioj} and Z_j, and the level set is $\frac{\partial \pi_{ioj}}{\partial q_{ioj}} = 0$.

[3] These derivations assume $\theta_{oj} > 0$. In my empirical approximation for $\frac{d\mu_j}{dZ_j}$, I assign $\hat{\theta}_j = 1 - W_j$, where $W_j = 1$ if plant j has access to a more competitive waterborne option. I do not intend Equation (1.8) to be an exact empirical analogue, as the comparative statics capture infinitesimal changes in Z_j, rather than discrete changes in Z_j. However, these derivations hold for extremely small values of θ_{oj}.

[4] In other words, a small change in q_{ioj} should not change whether firm i behaves as a Cournot oligopolist ($\theta_{oj} \approx 1$) or as a dominant firm facing a competitive fringe of waterborne coal shipments ($\theta_{oj} \approx 0$).

Next, I derive the comparative static for $Q_{oj} = \sum_i q_{ioj}$. Totally differentiating Q_{oj} by Z_j, and invoking symmetry across all N_{oj} rail carriers:

$$\frac{dQ_{oj}}{dZ_j} = \sum_i \frac{\partial Q_{oj}}{\partial q_{ioj}} \frac{dq_{ioj}}{dZ_j} \tag{A.5}$$

$$\frac{dQ_{oj}}{dZ_j} = N_{oj}\theta_{oj} \frac{dq_{ioj}}{dZ_j} \tag{A.6}$$

$$\Rightarrow \quad \frac{dQ_{oj}}{dZ_j} = \frac{\dfrac{\partial P_{oj}}{\partial Z_j} N_{oj} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}\theta_{oj}}{-\left(2\dfrac{\partial P_{oj}}{\partial Q_{oj}} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}^2} \dfrac{Q_{oj}\theta_{oj}}{N_{oj}}\right)} \tag{A.7}$$

The first term in the numerator, $\frac{\partial P_{oj}}{\partial Z_j} N_{oj}$, captures the level-effect of the demand shift, which should be weakly positive. Intuitively, an inward demand shift (i.e. $dZ_j < 0$) should reduce plant j's level of coal consumption Q_{oj}. The second term in the numerator, $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}\theta_{oj}$, captures the extent to which the demand shift dZ_j changes the slope of inverse demand. If demand becomes more elastic as gas prices fall (i.e. $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} < 0$), rail carriers should increase their best-response quantities. This term is scaled by the conduct parameter θ_{oj}, and it converges to zero as route oj becomes more competitive. The denominator of Equation (A.7) must be positive by the second-order condition.

The final step is to convert this comparative static into the total derivative of price P_{oj} with respect to Z_j:

$$\frac{dP_{oj}}{dZ_j} = \frac{\partial P_{oj}}{\partial Q_{oj}} \frac{dQ_{oj}}{dZ_j} + \frac{\partial P_{oj}}{\partial Z_j} \tag{A.8}$$

$$\frac{dP_{oj}}{dZ_j} = \frac{\partial P_{oj}}{\partial Q_{oj}} \left[\frac{\dfrac{\partial P_{oj}}{\partial Z_j} N_{oj} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}\theta_{oj}}{-\left(2\dfrac{\partial P_{oj}}{\partial Q_{oj}} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}^2} \dfrac{Q_{oj}\theta_{oj}}{N_{oj}}\right)} \right] + \frac{\partial P_{oj}}{\partial Z_j} \tag{A.9}$$

$$\frac{dP_{oj}}{dZ_j} = \left[\frac{\dfrac{\partial P_{oj}}{\partial Z_j} N_{oj} + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}\theta_{oj}}{-\left(2 + \dfrac{\partial^2 P_{oj}}{\partial Q_{oj}^2} \left(\dfrac{\partial P_{oj}}{\partial Q_{oj}}\right)^{-1} \dfrac{Q_{oj}\theta_{oj}}{N_{oj}}\right)} \right] + \frac{\partial P_{oj}}{\partial Z_j} \tag{A.10}$$

Let $E_{D_{oj}} \equiv \left(\frac{\partial^2 P_{oj}}{\partial Q_{oj}^2}\right)\left(\frac{\partial P_{oj}}{\partial Q_{oj}}\right)^{-1} Q_{oj}\left(\frac{\theta_{oj}}{N_{oj}}\right)$, or the elasticity of the slope of inverse demand scaled by $\frac{\theta_{oj}}{N_{oj}}$ (a heuristic for the degree of competitiveness). Making this substitution, and simplifying:

$$\text{(A.11)} \qquad \frac{dP_{oj}}{dZ_j} = \frac{\frac{\partial P_{oj}}{\partial Z_j} N_{oj} + \frac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j} Q_{oj}\theta_{oj}}{-\left(2 + E_{D_{oj}}\right)} + \frac{\partial P_{oj}}{\partial Z_j}\left(\frac{2 + E_{D_{oj}}}{2 + E_{D_{oj}}}\right)$$

$$\text{(A.12)} \qquad \Rightarrow \quad \frac{dP_{oj}}{dZ_j} = \frac{\frac{\partial P_{oj}}{\partial Z_j}\left(2 + E_{D_{oj}} - N_{oj}\right) - \frac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j} Q_{oj}\theta_{oj}}{2 + E_{D_{oj}}}$$

I define markups as $\mu_{oj} \equiv P_{oj} - C_o - S(\mathbf{T}_{oj})$, assuming that C_o and $S(\mathbf{T}_{oj})$ are each independent of q_{ioj} and Z_j.[5] Hence, Equation (1.3) follows directly from (A.12):

$$\Rightarrow \quad \frac{d\mu_{oj}}{dZ_j} = \frac{\frac{\partial P_{oj}}{\partial Z_j}\left(2 + E_{D_{oj}} - N_{oj}\right) - \frac{\partial^2 P_{oj}}{\partial Q_{oj}\partial Z_j} Q_{oj}\theta_{oj}}{2 + E_{D_{oj}}}$$

$\square$

I make several simplifying assumptions to arrive at this comparative static. I assume that plant j cannot resell purchased coal to other plants, which effectively enables rail carrier i to optimize each oj market independently. A less restrictive formulation of the rail carrier's problem would include $[O \times J]$ arbitrage constraints, which would bind if the P_{oj} were sufficiently high/low to make coal resale cost-effective. This is the classic representation of 3rd-degree price discrimination, as presented by Schmalensee (1981) and others.

I also ignore dependencies in coal demand across plants and coal types. In reality, plant j's demand for type-o coal certainly depends on *both* the prices it faces for coal from other counties *and* the prices faced by other plants (Varian (1985); Katz (1987)). I can incorporate this vector of non-oj prices into the parameter vector $\mathbf{Z}_{oj}$, or use more straightforward notation and explicitly include this vector of prices as an argument entering inverse demand: $P_{oj}(Q_{oj}; \mathbf{P}_{-(oj)}, \mathbf{Z}_{oj})$.

My model ignores the potential for binding rail price regulation, which is obviously unrealistic. Since 1996, the Surface Transportation Board has adjudicated 33 rate cases disputing the "reasonableness" of coal-by-rail shipping rates, which likely represents only a small fraction of rail rates constrained by the threat (or even the *perceived* threat) of a regulatory challenge.

[5] During my sample period, U.S. natural gas prices were virtually uncorrelated with diesel prices, the main component of $\mathbf{T}_{oj}$. Figure 1.2.2 shows that delivered coal prices do not respond to short-to-medium run changes in the Henry Hub spot price.

I can modify Equations (1.1) and (A.1) to explicitly account for each of these assumptions, in order to illustrate what the above derivation assumes away:

$$\Pi_i(\mathbf{q}_i) = \mathbf{q}_i \cdot \Big[\mathbf{P}(\mathbf{Q};\mathbf{Z}) - \mathbf{C}_i - \mathbf{S}_i(\mathbf{T})\Big] - \mathbf{F}_i \cdot \mathbf{1}\Big[\mathbf{q}_i > \mathbf{0}\Big] \tag{A.13}$$

s.t.

$$\Big|P_{oj}(Q_{oj};\mathbf{Q}_{-(oj)},\mathbf{Z}) - \mathbf{P}(\mathbf{Q};\mathbf{Z})\Big| \leq \mathbf{A}_j \quad \forall\ o,\ \forall\ j \tag{A.14}$$

$$\mathbf{P}(\mathbf{Q};\mathbf{Z}) \leq \mathbf{R}_i \tag{A.15}$$

Here, rail carrier i jointly optimizes across all $[O \times J]$ pairs, and $\mathbf{q}_i$ is a vector of length $[O \times J]$ coal quantities on each route oj. $\mathbf{P}(\mathbf{Q};\mathbf{Z})$ is the $[O \times J]$-dimensional inverse demand function, which depends on the $[O \times J]$ vector of coal quantities $\mathbf{Q}$ and a matrix of demand parameters $\mathbf{Z}$ with $[O \times J]$ rows. $\mathbf{C}_i$ is an $[O \times J]$ vector of mine-mouth coal costs, as faced by carrier i. $\mathbf{S}_i(\mathbf{T})$ is an $[O \times J]$ vector of carrier i's coal shipping costs. $\mathbf{F}_i$ is an $[O \times J]$ vector of carrier i's fixed costs of maintaining each oj route, which is multiplied by an $[O \times J]$ vector of indicators for carrier i's entry decision along each route.

The $[O \times J]$ constraints in (A.14) prevent arbitrage across coal plants, and $\mathbf{A}_j$ is an $[O \times J]$ vector of the price wedges at which arbitrage becomes feasible (i.e. reflecting plant j's costs of shipping coal to/from all other plants).[6] The constraint (A.15) introduces the threat of regulatory intervention, where $\mathbf{R}_i$ is an $[O \times J]$ vector of the maximum price carrier i is willing to set on each oj route given the risk of a rate challenge. Assuming that (A.14) and (A.15) never bind, and assuming that the first-derivative matrix of $\mathbf{P}(\mathbf{Q};\mathbf{Z})$ is diagonal,[7] this more general profit function collapses to Equations (1.1) and (A.1).

A.1.2 Derivation of Carbon Tax Pass-Through

In this section, I derive expressions for the carbon tax pass-through (ρ) implied by a change in natural gas prices (ΔZ) and a reoptimization of coal markups ($\Delta\mu$). Consider a single coal plant j in a market with many natural gas plants. The ratio of coal-to-gas marginal costs governs the relative ordering of plants on the electricity supply curve, which is the primary factor influencing plant j's operating decisions (see Figure 1.3.3 in the main text). Plant j's cost ratio is (suppressing subscripts):

$$CR = \frac{MC^{coal}}{MC^{gas}} = \frac{HR_{coal} \cdot (P + MC^{env}_{coal})}{HR_{gas} \cdot (Z + MC^{env}_{gas})} \tag{A.16}$$

[6] $\mathbf{A}_j$ is of length $[O \times J]$ in order to conform with the dimension of $\mathbf{P}(\mathbf{Q};\mathbf{Z})$, meaning that there are O null j-to-j no-arbitrage constraints within plant j's set of arbitrage constraints. Vertical lines denote the absolute value operator, applied element-wise to the difference between the scalar $P_{oj}(Q_{oj};\mathbf{Q}_{-(oj)},\mathbf{Z})$ and the vector $\mathbf{P}(\mathbf{Q};\mathbf{Z})$. Busse and Keohane (2007) note that in order to arbitrage around the railroads, plants would need to transfer coal from their on-site storage piles onto trucks, a more costly mode of transportation. While coal resale is quite rare in practice, the *threat* of arbitrage may limit railroads' willingness to price discriminate.

[7] That is, $\frac{\partial P_{oj}}{\partial Q_{nk}} = 0$ for $k \neq j$ or $n \neq o$; and $\frac{\partial P_{oj}}{\partial Z_{nk}} = 0$ for $k \neq j$ or $n \neq o$, for any element of $\mathbf{Z}$ Z.

HR_{fuel} is the heat rate in units of MMBTU/MWh, or the rate at which each plant converts fuel into electricity. P denotes the coal price, while Z denote the gas price, both in \$/MMBTU. Finally, MC^{env}_{fuel} represents the marginal costs of environmental compliance per MMBTU of fuel, which is positive for plants that operate pollution control devices or participate in allowance trading programs for SO_2, NO_x, or CO_2.[8]

Let ΔZ denote a change in the gas price Z, which implies a new cost ratio:

$$CR' = \frac{HR_{coal} \cdot (P + MC^{env}_{coal})}{HR_{gas} \cdot (Z + \Delta Z + MC^{env}_{gas})} \tag{A.17}$$

Note that a hypothetical carbon tax t would yield an identical change in the cost ratio, holding gas prices constant:

$$CR' = \frac{HR_{coal} \cdot (P + MC^{env}_{coal})}{HR_{gas} \cdot (Z + \Delta Z + MC^{env}_{gas})} = \frac{HR_{coal} \cdot (P + MC^{env}_{coal} + tE_{coal})}{HR_{gas} \cdot (Z + MC^{env}_{gas} + tE_{gas})} \tag{A.18}$$

The tax t is denominated in \$ per metric ton CO_2, and E_{fuel} is the CO_2 emissions rate for each fuel, in metric tons CO_2 per MMBTU. Because $E_{coal} > E_{gas}$, t exists for any feasible change in gas prices ΔZ.[9]

Solving Equation (A.18) for t:

$$\begin{aligned}
\frac{P + MC^{env}_{coal}}{Z + \Delta Z + MC^{env}_{gas}} &= \frac{P + MC^{env}_{coal} + tE_{coal}}{Z + MC^{env}_{gas} + tE_{gas}} \\
(P + MC^{env}_{coal})(Z + MC^{env}_{gas} + t \cdot E_{gas}) &= (Z + \Delta Z + MC^{env}_{gas})(P + MC^{env}_{coal} + tE_{coal}) \\
(P + MC^{env}_{coal})tE_{gas} &= (Z + \Delta Z + MC^{env}_{gas})tE_{coal} + \Delta Z(P + MC^{env}_{coal}) \\
t\Big[(P + MC^{env}_{coal})E_{gas} - (Z + \Delta Z + MC^{env}_{gas})E_{coal}\Big] &= \Delta Z(P + MC^{env}_{coal})
\end{aligned}$$

$$\Rightarrow \quad t(\Delta Z) = \frac{\Delta Z(P + MC^{env}_{coal})}{(P + MC^{env}_{coal})E_{gas} - (Z + \Delta Z + MC^{env}_{gas})E_{coal}} \tag{A.19}$$

For a change in gas prices ΔZ, $t(\Delta Z)$ represents the equivalent carbon tax implied by this gas price change.

[8] HR_{gas} and MC^{env}_{gas} are weight-averaged across all gas plants that compete with plant j in electricity dispatch. This abstracts from non-fuel variable costs, such as labor, plant maintenance, and other inputs. I make this simplifying assumption for two reasons. First, electricity production is Leontieff in fuel inputs (Fabrizio, Rose, and Wolfram (2007)), and non-fuel inputs are of second-order importance to marginal operating decisions of fossil generators (Cicala (2017)). Second, reliable data on non-fuel variable costs are unavailable, as I discuss in Appendix A.2.2.3 below. Hence, I omit non-fuel inputs from marginal costs to be consistent with my coal demand estimation algorithm in Appendix A.4.

[9] I adapt this framework from Cullen and Mansur (2017), who illustrate this conceptual mapping between relative fuel prices and carbon tax very nicely in Figure 4 of their paper. Natural gas combustion has a homogeneous emssions rate of 0.053 metric tons CO_2 per MMBTU (`https://www.eia.gov/tools/faqs/faq.php?id=73&t=1`). While coal's carbon content does vary slightly by grade, its average CO_2 emissions rate of 0.095 metric tons CO_2 per MMBTU is relatively homogeneous, compared to its SO_2, NO_x, or Hg emissions rates.

However, rail carriers may reoptimize coal markups in response to ΔZ. If markups change by $\Delta\mu$, I can rewrite Equation (A.17):

$$CR' = \frac{HR_{coal} \cdot (P + \Delta\mu + MC^{env}_{coal})}{HR_{gas} \cdot (Z + \Delta Z + MC^{env}_{coal})} \tag{A.20}$$

Pass-through of the implicit tax $t(\Delta Z)$ is a function of $\Delta\mu$. If $\Delta\mu = 0$, then the pass-through rate is $\rho = 1$: the coal plant faces the full implicit tax $t(\Delta Z)$, without any changes in markups that weaken or strengthen this price signal. If $\text{sign}(\Delta\mu) = \text{sign}(\Delta Z)$, then $\Delta\mu$ weakens the effect of ΔZ on CR', translating to a incomplete pass-through ($\rho < 1$). If $\text{sign}(\Delta\mu) = -\text{sign}(\Delta Z)$, then $\Delta\mu$ strengthens the effect of ΔZ on CR', translating to overshifting ($\rho > 1$).

Modifying Equation (A.18) to allow for changes in markups ($\Delta\mu$) and incomplete pass-through ($\rho \neq 1$) for coal plants only:

$$\frac{HR_{coal} \cdot (P + \Delta\mu + MC^{env}_{coal})}{HR_{gas} \cdot (Z + \Delta Z + MC^{env}_{gas})} = \frac{HR_{coal} \cdot (P + MC^{env}_{coal} + \rho\, t(\Delta Z) E_{coal})}{HR_{gas} \cdot (Z + MC^{env}_{gas} + t(\Delta Z) E_{gas})} \tag{A.21}$$

Here, $t(\Delta Z)$ is the carbon tax implied by ΔZ under full pass-through (i.e. $\Delta\mu = 0$, $\rho = 1$). If markups adjust (i.e. $\Delta\mu \neq 0$), this causes coal plants to face a different proportion (i.e. $\rho \neq 1$) of this implicit tax. Note that ρ appears only in the numerator, as I assume that full tax pass-through of the tax for natural gas prices.

Solving Equation (A.21) for ρ:

$$\frac{P + \Delta\mu + MC^{env}_{coal}}{Z + \Delta Z + MC^{env}_{gas}} = \frac{P + MC^{env}_{coal} + \rho\, t(\Delta Z) E_{coal}}{Z + MC^{env}_{gas} + t(\Delta Z) E_{gas}}$$

$$\Big[P + MC^{env}_{coal} + \rho t(\Delta Z) E_{coal}\Big]\Big[Z + \Delta Z + MC^{env}_{gas}\Big] = \Big[P + \Delta\mu + MC^{env}_{coal}\Big]\Big[Z + MC^{env}_{gas} + t(\Delta Z) E_{gas}\Big]$$

$$\rho = \frac{\Big[P + \Delta\mu + MC^{env}_{coal}\Big]\Big[Z + MC^{env}_{gas} + t(\Delta Z) E_{gas}\Big] - \Big[P + MC^{env}_{coal}\Big]\Big[Z + \Delta Z + MC^{env}_{gas}\Big]}{t(\Delta Z)\Big[Z + \Delta Z + MC^{env}_{gas}\Big] E_{coal}}$$

$$\rho = \frac{\Delta\mu\Big[Z + MC^{env}_{gas}\Big] + t(\Delta Z)\Big[P + \Delta\mu + MC^{env}_{coal}\Big] E_{gas} - \Delta Z\Big[P + MC^{env}_{coal}\Big]}{t(\Delta Z)\Big[Z + \Delta Z + MC^{env}_{gas}\Big] E_{coal}}$$

$$\rho = \frac{(P + \Delta\mu + MC^{env}_{coal}) E_{gas}}{(Z + \Delta Z + MC^{env}_{gas}) E_{coal}} + \frac{\Delta\mu (Z + MC^{env}_{gas}) - \Delta Z (P + MC^{env}_{coal})}{(Z + \Delta Z + MC^{env}_{gas}) E_{coal}} \left[\frac{1}{t(\Delta Z)}\right]$$

Substituting $t(\Delta Z)$ from Equation (A.19), and rearranging:

$$\rho = \frac{(P + \Delta\mu + MC^{env}_{coal}) E_{gas}}{(Z + \Delta Z + MC^{env}_{gas}) E_{coal}} + \\ \frac{\Delta\mu (Z + MC^{env}_{gas}) - \Delta Z (P + MC^{env}_{coal})}{(Z + \Delta Z + MC^{env}_{gas}) E_{coal}} \left[\frac{(P + MC^{env}_{coal}) E_{gas} - (Z + \Delta Z + MC^{env}_{gas}) E_{coal}}{\Delta Z (P + MC^{env}_{coal})}\right]$$

$$\rho = \frac{(P + \Delta\mu + MC^{env}_{coal})E_{gas}}{(Z + \Delta Z + MC^{env}_{gas})E_{coal}} + \frac{\Delta\mu(Z + MC^{env}_{gas})E_{gas}}{\Delta Z(Z + \Delta Z + MC^{env}_{gas})E_{coal}} - \frac{\Delta\mu(Z + MC^{env}_{gas})}{\Delta Z(P + MC^{env}_{coal})} + 1 - \frac{(P + MC^{env}_{coal})E_{gas}}{(Z + \Delta Z + MC^{env}_{gas})E_{coal}}$$

$$\rho = 1 + \frac{\Delta\mu\, E_{gas}}{(Z + \Delta Z + MC^{env}_{gas})E_{coal}} + \frac{\Delta\mu(Z + MC^{env}_{gas})E_{gas}}{\Delta Z(Z + \Delta Z + MC^{env}_{gas})E_{coal}} - \frac{\Delta\mu(Z + MC^{env}_{gas})}{\Delta Z(P + MC^{env}_{coal})}$$

$$\Rightarrow \quad \rho(\Delta\mu, \Delta Z) = 1 + \frac{\Delta\mu}{\Delta Z}\left(\frac{E_{gas}}{E_{coal}} - \frac{Z + MC^{env}_{gas}}{P + MC^{env}_{coal}}\right) \tag{A.22}$$

This expression shows that $\Delta\mu$ leads to incomplete pass-through via two channels. The first fraction in parentheses adjusts for the wedge in emissions factors, while the second fraction rescales for the baseline difference in marginal costs.

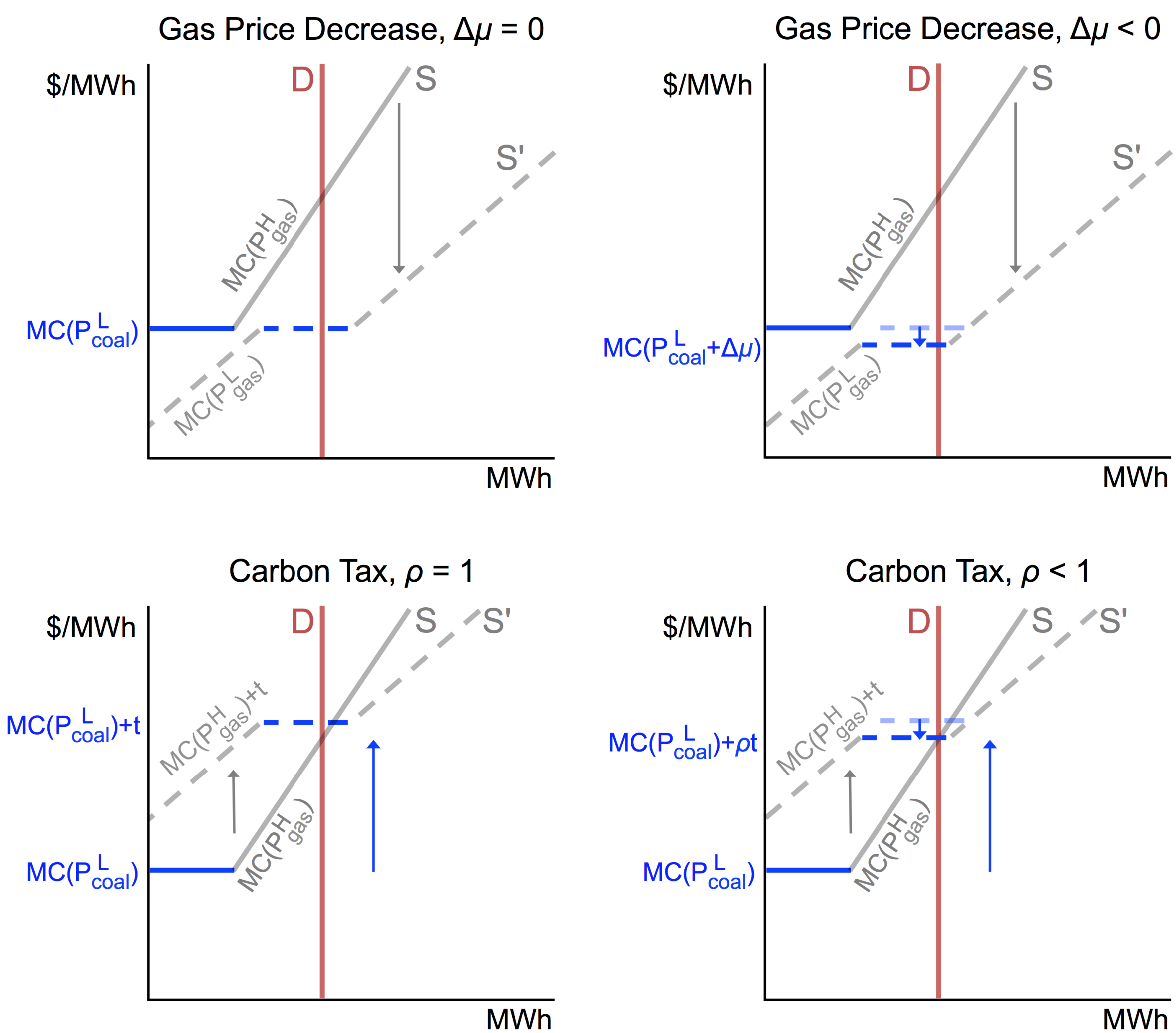
For expositional clarity, I simplify this derivation in the main text by removing heat rates and setting non-fuel costs to zero. My pass-through estimates in Table 1.7.7 directly apply Equation (A.22), setting $\Delta Z = 1$ and plugging in coefficient estimates for $\Delta\mu$ (setting $MC^{env}_{coal} = MC^{env}_{gas} = 0$ in Panel A). Note that because $\Delta\mu$ represents a change in plant j's markup, $\rho(\Delta\mu, \Delta Z)$ should really be $\rho_j(\Delta\mu_j, \Delta Z_j)$, where P_j and $MC^{env}_{coal,j}$ denote plant-specific costs, and Z_j and $MC^{env}_{gas,j}$ denote the costs of plant j's gas-fired competitors. Variation in delivered coal price P_j drives most of the heterogeneity in pass-through rates across plants *within* each column of Table 1.7.7.

Cullen and Mansur (2017) note that the mapping between the cost ratio (CR) and an implicit carbon tax (t) relies on several key assumptions. Most importantly, electricity demand must be perfectly inelastic; in reality, demand in wholesale electricity markets is close to perfectly inelastic, as there is extremely limited technical capacity for demand response. It also assumes that only coal and natural gas may be marginal in electricity markets; in reality, other technologies (e.g. diesel or hydroelectric generators) are rarely on the margin in regions of the U.S. that also feature non-trivial coal generating capacity. Given these two reasonable assumptions, electricity generation depends only on the *ordering* of plants along the supply curve. The "equivalent" carbon tax would produce the same ordering of marginal costs, but the tax would shift all plants' marginal costs upward, yielding the same generation outcomes at higher electricity prices. Figure A.1.1 illustrates how a gas price decrease can yield the same *allocation* of generation as a carbon tax, comparing the top-left vs. bottom-left panels. Comparing the top-right vs. bottom-right panels, a decrease in coal markups ($\Delta\mu < 0$) can increase the allocation of coal-fired electricity generation in a manner that mimics incomplete pass-through of a carbon tax to coal generators ($\rho < 1$).

Mapping the fuel price ratio to a carbon tax also assumes that higher electricity prices would not alter plants' bidding strategies, due to either dynamic operating constraints (which Cullen (2015) finds to be second-order), exercise of market power (Mansur (2013)), or differential pass-through of shocks to fuel vs. carbon prices (Fabra and Reguant (2014)).

Finally, this mapping only holds in the short-run, where both the stock of generators and their CO_2 emissions rates are fixed.

Figure A.1.1: Electricity Supply with Gas Price Decrease or Carbon Tax



Notes: This figure illustrates how a gas price decrease mimics a carbon tax on the electricity sector in the short-run, using the same stylized electricity market as Figure 1.3.3 in the main text. The top-left panel reproduces the top-left panel of Figure 1.3.3. My empirical results find that coal markups decrease due to a decrease in gas prices, and I illustrate this decrease in markups ($\Delta\mu < 0$) in the top-right panel. In the absence of a gas price decrease, there exists a carbon tax (t) that would have yielded the same supply curve as the top-left panel, except vertically shifted upwards (i.e. bottom-left panel). Incomplete pass-through of that carbon tax (i.e. $\rho < 1$ in the bottom-right panel) can result in the same generation allocation under a (counterfactual) carbon tax as decreasing coal markups after a (factual) gas price drop.

A.1.3 Incidence

Weyl and Fabinger (2013) derive the following expression for the incidence (I) of a tax (t) — or the ratio of changes in consumer surplus (CS) vs. producer surplus (PS) — in a symmetric oligopoly:

$$I = \frac{dCS/dt}{dPS/dt} = \frac{\rho}{1 - (1 - \theta/N)\,\rho} \tag{A.23}$$

Here, ρ is the pass-through rate of the tax, θ is the conduct parameter, and N is the number of symmetric firms in the market.[10] As pass-through becomes more incomplete (i.e., as ρ decreases from 1), incidence decreases and consumers pay proportionately less of the tax burden. For a given pass-through rate ρ, a less competitive market structure (i.e., greater θ/N) also implies lower incidence, because producers stand to lose more under a tax if they are already extracting more oligopoly rents.[11]

In my setting, I assume a symmetric oligopoly where all plants face either an effective rail monopoly (i.e. captive, $N_j = 1$) or an effective rail duopoly (i.e. non-captive, $N_j = 2$). I also assume that plants' proximity to a navigable waterway governs their rail market structure, where coal shipping is Cournot (i.e. $\theta_j = 1$) for plants without a water delivery option, and competitive (i.e. $\theta_j = 0$) for plants with water as an outside option. These assumptions reduce Equation (A.23) to four possible mappings between pass-through an incidence:

$$I_j = \begin{cases} \rho_j & \text{if } \theta_j = 1\,,\ N_j = 1 \\ \dfrac{\rho_j}{1 - \rho_j/2} & \text{if } \theta_j = 1\,,\ N_j = 2 \\ \dfrac{\rho_j}{1 - \rho_j} & \text{if } \theta_j = 0\,,\ \rho_j < 1 \\ \infty & \text{if } \theta_j = 0\,,\ \rho_j = 1 \end{cases} \tag{A.24}$$

If pass-through is incomplete (i.e. $\rho_j < 1$), then for a given pass-through rate, a less competitive market (i.e. higher θ_j or lower N_j) reduces the share of the tax borne by coal plants (i.e. consumers) relative to rail carriers (i.e. producers). For example, suppose the pass-through rate is $\rho_j = 0.8$ for three plants: a captive plants with no water option ($\theta_j = 1$, $N_j = 1$), a non-captive plant with no water option ($\theta_j = 1$, $N_j = 2$), and a plant with a water option ($\theta_j = 0$). This would imply incidences of 0.8, 1.3, and 4, respectively. Infinite incidence occurs under perfect competition if supply is perfectly elastic, causing $\rho_j = 1$ and $dPS/dt = 0$.

[10] See Weyl and Fabinger (2013), p. 547. Note that using my notation, θ/N corresponds to θ in the notation of Weyl and Fabinger.

[11] That is, $\partial I/\partial\rho > 0$; and $\partial I/\partial(\theta/N) < 0$, unless $I < 0$, which can only occur if $\rho > 1$.

I can reformulate incidence to summarize the proportion of the tax burden in the coal-by-rail market that is borne by coal plants (i.e. consumers):

$$\text{(A.25)}\quad \frac{I_j}{1+I_j} = \frac{\left(\frac{dCS}{dt}\right)_j}{\left(\frac{dCS}{dt}\right)_j + \left(\frac{dPS}{dt}\right)_j} = \frac{\rho_j}{1+(\theta_j/N_j)\,\rho_j} = \begin{cases} \dfrac{\rho_j}{1+\rho_j} & \text{if } \theta_j = 1\,,\ N_j = 1 \\ \dfrac{\rho_j}{1+\rho_j/2} & \text{if } \theta_j = 1\,,\ N_j = 2 \\ \rho_j & \text{if } \theta_j = 0 \end{cases}$$

For a pass-through rate of $\rho_j = 0.8$, a plant could bear 44 percent (if $\theta_j = 1$, $N_j = 1$), 57 percent (if $\theta_j = 1$, $N_j = 2$), or 80 percent (if $\theta_j = 0$) of the lost surplus in coal markets. Importantly, the full tax burden would depend on the extent to which coal plants could pass on marginal emissions costs via higher wholesale electricity prices. If emissions tax pass-through in wholesale electricity markets is 1 (consistent with Fabra and Reguant (2014)), then a carbon tax *could* increase profits for coal plants that are relatively clean/efficient and have low pass-through rates in coal markets. While the fracking boom simulated the effect of a carbon tax on the *relative* costs of coal vs. gas plants, it had the opposite effect on electricity prices, as low gas prices caused electricity prices to fall (Linn and Muehlenbachs (forthcoming)). Lower electricity prices meant that coal plants were very unlikely to be "winners" in the fracking boom, though incomplete pass-through in coal shipping likely caused certain coal plants to be "smaller losers".

My theoretical framework assumes that rail carriers purchase coal from a perfectly competitive mining sector. While this greatly simplifies the mathematical derivations in Appendix A.1.1, this assumption is not crucial for evaluating markups for coal deliveries.[12] However, the market structure at the mine-mouth does impact how mines and railroads share the tax burden. Given that coal is both spatially and physically heterogeneous, the fundamental assumptions of perfect competition likely do not hold.

I assume that C_o in Equations (1.1) and (A.1) is exogenous, which implies that rail carriers face perfectly elastic coal supply (a standard assumption for *homogeneous* commodity markets). This assumes that coal mines earn zero economic rents, and hence incur no lost profits due to a downstream carbon tax. However coal mines would face non-zero tax burden under any of three alternative market structures.

First, suppose that coal mines and rail carriers coordinate, behaving as vertically integrated monopolists. This may occur without formal profit-sharing, as geographically isolated mines in Wyoming's Powder River Basin depend heavily on rail carriers to transport their coal to market, while Western rail carriers with sunk investments in railroad tracks stand to gain considerable profits by cooperating with mines in a repeated game. In this case, I could rewrite the rail carrier's profit function replacing coal quantity (q_{ioj}) with the mine's production function, and replacing constant marginal cost (C_o) with total mining input costs. This would cause coal mines and rail carriers to jointly share the burden of a carbon tax.

[12] My regression specifications control for the average *equilibrium* coal price by county-year, without restricting the market structure of the mining sector.

Second, suppose that coal mines can exert market power at the mine-mouth when selling to rail carriers. This may occur if a few large firms dominate mining operations (as in the Powder River Basin), if multiple rail carriers compete to purchase coal from a single mining firm, or if mines can elect to sell to non-rail intermediaries (e.g. river barges). In this case, double marginalization would shift delivered coal prices even further from the competitive benchmark (ignoring externalities).[13] This would create a second opportunity for incomplete pass-through of a cost shock (or a carbon tax), but adjustments in markups would not be coordinated along the coal supply chain. If mines responded by reducing mine-mouth markups, this would reduce the tax burden borne by rail carriers. Mines earning market power rents would also stand to lose under a carbon tax.

Third, suppose that rail carriers can exert monopsony power at the mine-mouth. This may occur if coal mines are captive to a single rail carrier with strong bargaining power: whereas these mines depend on revenue from selling a single product, diversified rail carriers may divert resources (e.g. locomotives, labor) to other profitable shipping opportunities. In this case, rail carriers could adjust prices both at the mine-mouth and at the power plant, and they would likely bear an even greater tax burden for having extracted rents on both sides of the market. If a downstream carbon tax caused rail carriers to raise mine-mouth prices (i.e. incomplete pass-through at the mine-mouth), the effects on profits in the mining sector would be theoretically ambiguous, and would depend on whether mine-mouth prices increased by enough to offset the reduction in coal quantity.

A.2 Data

A.2.1 Coal Transaction and Production Data

A.2.1.1 Coal Shipment Data

The core dataset in my analysis is the Energy Information Administration (EIA) database of monthly fossil fuel deliveries to power plants. These survey data are at the monthly "order" level, and power plants are required to report each purchase order or supplier contract separately. According to EIA's official documentation, "aggregation of coal receipt data into a single line item is allowed if the coal is received under the same purchase order or contract and the purchase type, fuel, mine type [i.e., surface vs. underground], state of origin, county of origin, and supplier are identical for each delivery." Throughout my analysis, I refer to observations in this dataset as "shipments" (indexed by s); this short-hand aggregates multiple physical shipments into a single "shipment" observation (e.g., daily train deliveries on the same long-term contract).[14]

[13] Alexandrov, Pittman, and Ukhaneva (2017) find no empirical evidence of double marginalization between railroads, in cases where multiple rail carriers own segments along the same shipping route.

[14] See `https://www.eia.gov/electricity/monthly/pdf/technotes.pdf` for detailed descriptions of EIA's fossil fuel delivery data.

Since 2008, EIA has collected fossil fuel delivery data on Form 923, with monthly reporting required for all fossil plants larger than 50 MW in total generating capacity. Form 923 also reports monthly data on a sample of smaller plants (1–50 MW capacity), and all plants larger than 1 MW are required to report fossil fuel receipts annually. Prior to 2008, monthly data were collected on two separate forms, each with a 50 MW minimum reporting requirement: the Federal Energy Regulatory Commission's (FERC) Form 423 (for utility-owned plants between 1983–2007); and EIA Form 423 (for non-utility-owned plants between 2002–2007).[15] Given that my analysis period spans the 2008 changeover, I restrict my sample to include coal plants larger than 50 MW; this represents over 99 percent of U.S. coal-fired electricity generation. Prior to 2002, non-utility-owned plants were not required to report fuel deliveries, and the vast majority of such plants were divested by utilities between 1997 and 2002.

My sample period starts in 2002 and extends through 2015, the most recent year with available data across all key data sources. Beginning my analysis in 2002 affords five years of data prior to the beginning of the fracking boom in 2007 (Hausman and Kellogg (2015)). This also minimizes the potential for any confounding effects from electricity market deregulation and coal plant divestments, as most of these changes occurred prior to 2002, due to the 1998–2000 California Electricity Crisis (Fabrizio, Rose, and Wolfram (2007); Borenstein (2002)). Linn and Muehlenbachs (forthcoming) also document substantial data irregularities in coal deliveries prior to 2001, which is another reason I start my sample period in 2002.

The 423/923 data report average prices, total quantities, and average attributes for each fuel "shipment". Average prices are reported in dollars per ton (for coal), dollars per barrel (for oil), and dollars per thousand cubic feet (for gas). They are inclusive of commodity costs, shipping costs, and markups, equivalent to P_{ojms} from my theoretical framework. EIA withholds price data for deliveries to non-utility plants; hence, my analysis focuses on utility plants only, as I do not observe these prices. The data also identify the fuel type of each shipment—for example, bituminous vs. sub-bituminous coal; fuel oil vs. kerosene for oil; natural gas vs. liquefied petroleum gas for gas. Total quantities are reported in tons, barrels, and thousand cubic feet, and I can covert these physical quantities into units of energy content by multiplying by the average MMBTU content for each shipment (i.e. MMBTU per physical unit). The data report two additional physical attributes for each shipment: average sulfur content and average ash content. Each of these physical attributes (BTU, sulfur, ash) affects the price of coal deliveries, as plants tend to value coal with relatively high BTU content, and relatively low sulfur and ash content. Each of my markup regressions includes BTU, sulfur, and ash content as coal commodity controls in the matrix $\mathbf{C}_{ojms}$.

Each fossil fuel shipment is classified by purchase type, as either spot market or long-term contract. Most contract shipments have reported expiration dates, however this

[15] EIA Form 923 data are available at `https://www.eia.gov/electricity/data/eia923/`, which contain annual/monthly fossil fuel receipts data beginning in 2008. FERC Form 423 and EIA Form 423 are each available at `https://www.eia.gov/electricity/data/eia423/`.

variable is not consistently coded across years. Hence, I control for a coarser (but more consistently coded) measure of contract length as a part of $\mathbf{C}_{ojms}$: an indicator for contracts expiring within 2 years. Longer coal contracts (i.e. for which this indicator is zero) are likely to have higher coal prices, because plants trade off higher costs for lower variance in coal prices and more reliable deliveries (Jha (2017)). For the same reason, I also include a spot market indicator as a part of $\mathbf{C}_{ojms}$, as spot shipments tend to have lower delivered prices than contract shipments. Finally, contract shipments tend to be relatively less flexible, and my analysis of markup changes treats these two transaction types both pooled and separately.

The 423/923 data report the originating county of each coal shipment (reported since 1990), coal supplier names (reported since 2002 with markedly incomplete coverage in earlier years), and originating mine names and identifiers (reported since 2008). Given that supplier names and mine names are not (precisely) reported through my 2002–2015 sample period, I elect to treat the originating county as the unit of origin for each coal shipment. Two last variables are key for my analysis: each delivery's primary and secondary mode of transportation. Given that I estimate a model of *rail* shipping costs, mistakenly including barge shipments or truck shipments would induce misspecification. These variables are only reported for 2008–2015 sample years, however I am able to extrapolate backwards to assign transportation modes for 2002–2007 shipments using observed modes within each origin-destination pair.[16] I exclude all non-rail shipments from my main regression analysis on coal markups.

Besides using the 423/923 dataset as the backbone of my markup regressions, I use monthly average delivered fuel prices to construct the cost ratio CR_{ud} in my coal demand estimation. For coal prices, I use the BTU-weighted average monthly price received by each utility-owned coal plant, linearly interpolating prices for any missing months (P_{jm} in Equation (A.26)). For gas prices, I similarly calculate BTU-weighted average monthly delivered prices for utility-owned gas plants. However, these prices obscure day-to-day variation in gas prices. I leverage time series from natural gas trading hubs to assign daily average gas prices, comparing monthly average hub prices with monthly average 423/923 prices to add retail distribution costs into Z_{gd} in Equation (A.27). (See Appendix A.2.5 below.)

A.2.1.2 Aggregated Coal Prices

I supplement EIA coal delivery data with two data sources of aggregate coal prices. First, I use average mine-mouth sales prices at the county-year level, which EIA publishes in its *Annual Coal Report* (ACR).[17] EIA discloses the average annual price for open market coal

[16]For example, supposed that 99 percent of tons shipped from origin o to destination j between 2008–2015 were rail shipments. I can assign the missing 2002–2007 transportation modes for the same oj pair as "rail" with a high degree of confidence. This backwards extrapolation is reasonably unambiguous for the vast majority of oj pairs.

[17]I extract average coal sales prices from the data tables published in EIA *Annual Coal Report* PDFs, which are available for download at `https://www.eia.gov/coal/annual/`.

sales for counties with a sufficient number of mining firms to disclose aggregate prices, which corresponds to 62 percent of coal production. The ACR also reports average prices at the state-year level, separately for surface and underground mines, which I combine with coal production data (described below) to algebraically infer average prices for withheld counties. These prices correspond to C_o in my theoretical framework, or the average coal commodity cost per ton. I control for these county-by-year average prices in $\mathbf{C}_{ojms}$ in Equations (1.5) and (1.9), allowing me to better isolate delivered coal markups by capturing within-county changes in coal price.[18]

I use monthly average prices for coal delivered to electric power plants, as published in EIA's *Electric Power Monthly* (EPM). These prices are state-by-month aggregates, and EIA reports separate average prices for utility owned vs. non-utility plants (a.k.a. independent power producers).[19] This allows me to assign average delivered coal prices for non-utility plants, for which I do not observe prices at the purchase-order level. EPM withholds prices for state-months with too few firms, and I assign average prices for withheld cells by algebaically inferring missing prices from aggregate quantities (where possible), or by assigning region-by-month average prices. These prices only enter my analysis through the fuel cost ratio in my counterfactual estimation (CR_{ud} in Equation (A.45)), as I allow utility plants' generation to depend on the average monthly coal price across *all* (utility and non-utility) generators in each plant's PCA.[20] I also use average EPM prices for natural gas deliveries (constructed analogously) for a sensitivity analysis in Table A.5.26 and Figure A.5.30.

A.2.1.3 Coal Production and Mine Characteristics

I use several publicly available datasets on coal mining and production, published by the Mine Safety and Health Administration (MSHA), an agency with the U.S. Department of Labor.[21] The "Mines" dataset reports numerous mine-specific characteristics, all linked to a unique longitudinal mine identifier. For each mine, these data report its name;

[18]Month fixed effects control for changes in the global commodity price, while coal county fixed effects control for each county's average mine-mouth price. $\mathbf{C}_{ojms}$ also controls for the BTU content, sulfur content, and ash content of each coal shipment. Hence, the main reason to include county-year average prices as an additional control is to capture (otherwise unobserved) changes in cross-county differences in coal price.

[19]Electric Power Monthly data tables are available at `https://www.eia.gov/electricity/monthly/backissues.html` for 2003–2016, or by modifying the following link for each monthYYYY combination from 2011 to the present: `https://www.eia.gov/electricity/monthly/current_year/march2017.zip`. Tables 4.10.A and 4.10.B report average delivered coal prices, while Tables 4.6.A and 4.6.B report average delivered coal quantities.

[20]In both my demand estimates and counterfactual estimates, I estimate generation for utility plants only (i.e. those plants with publicly disclosed prices in EIA coal delivery data). While my demand estimates condition on the average coal price for each plant-month, my counterfactual estimates condition on the average coal price for each PCA-month—weight averaging across utility plants (using EIA 423/923 prices) and non-utility plants (using EIA EPM prices).

[21]These data are available for download at `http://www.msha.gov/OpenGovernmentData/OGIMSHA.asp`.

type (i.e., coal, metal); status (e.g., active, abandoned); county of location; latitude and longitude; and the average thickness (or height) of mine seams.[22] The "Quarterly Mine Employment and Coal Production Report" dataset (MSHA Form 7000-2) reports coal production, average number of employees, and total employee-hours worked, as reported by each mine operator for each quarter in a calendar year, and disaggregated by mine subunit (i.e. underground operations, surface operations, office work).

Using MSHA mine identifiers, I merge these two datasets with EIA's Form 7A, or "Annual Survey of Coal Production and Preparation".[23] These data allow me to cross-validate mine-specific production and employment for each year. EIA also classifies each mine as either "surface" or "underground"—categories that are consistent with the definitions used to report aggregate prices in the Annual Coal Report. EIA splits mines into 8 distinct mining regions, or basins: Appalachia Central, Appalachia Northern, Appalachia Southern, Illinois Basin, Powder River Basin, Uinta Region (in Utah, Colorado, and southern Wyoming), Interior (i.e., mines west of Illinois and East of Wyoming/Colorado), and Western (i.e., non-Powder River Basin, non-Uinta).

Following Cicala (2015), I use stratigraphic data from the U.S. Geological Survey (USGS) to calculate the depth of mine seams, as coal closer to the surface tends to be less expensive to extract. I use both the USTRAT and COALQUAL databases, which together include over 200,000 geocoded core sample collected by federal and state authorities to map U.S. coal deposits.[24] I convert these data into a raster, in order to assign each coal mine a time-invariant depth, basic on its geographic coordinates.[25]

I use mine coordinates and production data to determine the time-invariant production-weighted latitude and longitude for each coal producing county. This serves as an input into my graph algorithm that calculates shortest rail shipping distance and defines rail captiveness. I also use annual coal production by county to algebraically infer withheld ACR coal prices. Finally, I use the above datasets to construct several time-varying controls, in order to conduct sensitivity analysis on the components of $\mathbf{C}_{ojms}$ (in Table A.5.11 and A.5.29 below). These controls, all production-weighted averages at the quarter-year level, include: mine age, seam thickness, seam depth, the share of coal produced from (more expensive) underground mines, the share of mine employees working underground (which increases labor costs), hours worked per ton of coal produced.

[22] Thicker mine seams, or coal strata, are associated with less expensive coal extraction costs.

[23] EIA Form 7A datasets are available at `http://www.eia.gov/coal/data.cfm#production`.

[24] The USTRAT data are located at `https://ncrdspublic.er.usgs.gov/ncrds_data/`, while the COALQUAL data are located at `https://ncrdspublic.er.usgs.gov/coalqual/`.

[25] I use `R`'s `sp` package to convert USGS coordinates into a gridded spatial dataset. Then, I apply inverse distance weights to spatially interpolate between coordinates, using the function `idw` in `R`'s `gstat` package. Finally, using `R`'s `raster` package, I rasterize the interpolated grid and extract the seam depth corresponding to each pair of coal mine coordinates.

A.2.2 Power Plant Data

A.2.2.1 Plant Characteristics and Operations

EIA's Form 860, also known as the "Annual Electric Generator Report," is an annual survey of all U.S. electric power plants with greater than 1 MW in total generating capacity. I use these data to establish the universe of electric power plants for each year between 1990 and 2015.[26] For each plant-year, these data report the plant name and identifier; county of location; parent utility name and identifiers; regulatory status; an indicator for cogeneration; and the plant's primary purpose (i.e., to sell electricity to the electric power sector, for all plants in my sample). In addition, each annual data file reports characteristics for each plants constituent generating units, including each generator's type (e.g., steam turbine, combined-cycle, combustion turbine); status (e.g., operating, retired); year constructed; nameplate capacity (i.e., the maximum megawatts of electricity the generator is built to produce at any point in time); and primary fuel consumed (e.g., natural gas, bituminous coal).

I use data on power plant operations, collected by the following EIA forms, in reverse chronological order: 923, 906/920, 906, and 759.[27] My primary variables of interest are monthly heat input by fuel (i.e. fuel consumption in MMBTUs), and net electricity generated from each fuel (i.e. MWh sold to the electricity grid), which are both reported at the plant-month level. The data also report monthly fuel consumption disaggregated to the boiler level, and monthly net generation disaggregated to the generator level. These sub-plant units often map 1-to-1, where a single unit boils water and generates electricity; however, in many cases, multiple boilers map to a single generator or vice versa.[28] These operations data allow me to calculate plants' utilization rates and heat rates, or inverse thermal efficiency in units of MMBTU/MWh.

EIA Forms 767, 860, and 923 collect detailed data on plants' pollution abatement costs and pollution control devices.[29] At the plant level, these forms report annual capital expenditures on pollution abatement, and annual operation and maintenance costs of pollution control devices such as scrubbers (flue gas desulfurization units) and flue gas particulate collectors. They also report revenues from selling plant byproducts, most notably gypsum byproduct from flue gas desulfurization. At the boiler level, they report detailed data on scrubber characteristics and operations, while providing a crosswalk to match boiler identifiers with generator identifiers. I use these environmental compliance

[26]EIA Form 860 data are available for download at `http://www.eia.gov/electricity/data/eia860/`.

[27]Since 2008, Form 923 has collected data on both fuel receipts and plant operations. For 2001–2007, EIA collected plant-specific generation data on Forms 906 and 920. Prior to 2001, these data were collected separately for utility plants (Form 759) and non-utility plants (Form 906). All years of data are available at `http://www.eia.gov/electricity/data/eia423/`, `https://www.eia.gov/electricity/data/eia923/`, or `https://www.eia.gov/electricity/data/eia923/eia906u.html`.

[28]Form 906/920 boiler and generator data are both missing for 2006–2007.

[29]Prior to 2005, EIA collected these data on Form 767 (available at `http://www.eia.gov/electricity/data/eia767/`). Data collection appears to have lapsed in 2006 and resumed in 2007, with Forms 860 and 923 collecting most of the information formerly collected on Form 767.

data to control for the presence of a scrubber, the main pollution control technology that influences plants' coal purchases.[30] I also use variable environmental costs (O&M net of byproduct revenues) in constructing unit-specific marginal costs (i.e. MC^{env} in Equations (A.26) and (A.27)).

Finally, I use the Environmental Protection Agency's (EPA) Emissions & Generation Resource Integrated Database (eGRID) to assign plant geographic coordinates and electricity markets regions.[31] EPA assigns plants to regions of the electricity grid based at three distinct hierarchies. First, North American Electric Reliability Corporation (NERC) regions define 8 contiguous reliability regions of the transmission grid; there is substantial trade of electricity within but not across NERC regions. Second, many plants participate in wholesale electricity markets, and eGRID data assign these plants to a particular market, or Independent System Operator (ISO). I combine these two definitions for my demand estimation, giving preference to the ISO market regions while using NERC regions for plants that do not sell to an ISO.[32]

Third, eGRID assigns the majority of plants to a power control area (PCA), or the region on the electricity transmission grid in which the plant resides and over which a single Balancing Authority dispatches plants to instantaneously meet electricity demand. Applying a consistent PCA definition across plants is not trivial, as PCA boundaries evolve over time. Cicala (2017) identifies 98 major PCAs in the U.S., excluding Alaska and Hawaii. I perform a multistep PCA matching procedure to ensure that PCA definitions are both consistent across plants and across sample years, similar to the procedure used by Linn and Muehlenbachs (forthcoming). I cross-verify these definitions with two additional data sources: (i) Federal Energy Regulatory Commission (FERC) Form 714, "Annual Electric Balancing Authority and Planning Area Report"; and (ii) electricity market supply curves from SNL Financial.[33] I construct marginal cost ratios by averaging marginal costs across plants within each PCA.

---

[30] Plants with scrubbers tend to purchase high-sulfur bituminous coal, while plants who have not invested in this capital-intensive SO_2 abatement technology tend to purchase low-sulfur sub-bituminous coal.

[31] EPA's eGRID data are available at `http://www.epa.gov/energy/egrid/`, however these data only exist for the following years: 1996–2000, 2004, 2005, 2007, 2009, 2010, 2012, and 2014.

[32] The 7 ISOs are California ISO (CAISO); Electric Reliability Council of Texas (ERCOT); ISO New England (ISONE); Midcontentent ISO (MISO, formerly Midwest ISO); New York ISO (NYISO); PJM (formerly Pennsylvania-New Jersey-Maryland); and Southwest Power Pool (SPP). I define two market regions as North American Electric Reliability Corporation's (NERC) regions: Florida Reliability Coordinating Council (FRCC); and SERC Reliability Corporation (formerly Southeast Electric Reliability Council). The remaining two non-ISO market regions are subsets of NERC's Western region that exclude California: Northwest Power Pool (NWPP); and the Southwest Reserve Sharing Group (SRSG). Following Callaway, Fowlie, and McCormick (2018), I include smaller California PCAs in the CAISO region. I define market regions to be time-invariant. Following Linn and Muehlenbachs (forthcoming), I make minor adjustments to ISO/NERC boundaries such that all PCAs lie within a single market region.

[33] FERC Form 714 data are available at `https://www.ferc.gov/docs-filing/forms/form-714/data.asp` and report the name of all power plants residing in each Balancing Authority (in many cases, Balancing Authorities are close to isomorphic with PCAs). SNL data are proprietary, and available at `http://www.snl.com` under the tab "Generation Supply Curve".

A.2.2.2 EPA Air Markets Program Data

My demand estimation leverages high-frequency Continuous Emissions Monitoring Systems (CEMS) data on power plant operations, which are available for download through the EPA's Air Markets Program Data portal.[34] CEMS data include all fossil generating units that are either larger than 25 MW in capacity, or whose emissions are regulated under an EPA program. For each unit, for every hour since 2000, the data report fuel input (in MMBTUs), gross generation (in MWh), and emissions of SO_2, NO_x, and CO_2 (in tons of each pollutant). I use hourly CEMS generation for each unit in my demand estimation (Equation (A.29)) and total daily CEMS generation for each unit in my counterfactual regressions (Equation (A.45)).

CEMS data report the primary fuel for each unit (i.e. coal, natural gas, oil), and I use this primary fuel variable to determine which units are coal vs. gas units in my demand estimation.[35] The data also classify units into (primarily) three broad categories: boilers, combined cycle plants, and combustion turbines. I take these definitions as given and treat the CEMS unit identifier as the main sub-plant unit throughout my analysis. For most coal plants, CEMS units correspond to boilers rather than generators, and I can match over 96 percent of CEMS units to an EIA boiler identifier. I restrict my main estimation sample to the subset of plants that meet all of the following criteria: (1) plants that appear in CEMS data; (2) plants with at least one constituent CEMS unit with coal as a primary fuel; (3) plants categorized by CEMS as an "electric utility"; (4) plants that also appear as coal-consuming plants in EIA's 860 data; (5) plants that receive coal deliveries in EIA's 423/923 data; (6) plants larger than 50 MW in total generating capacity.

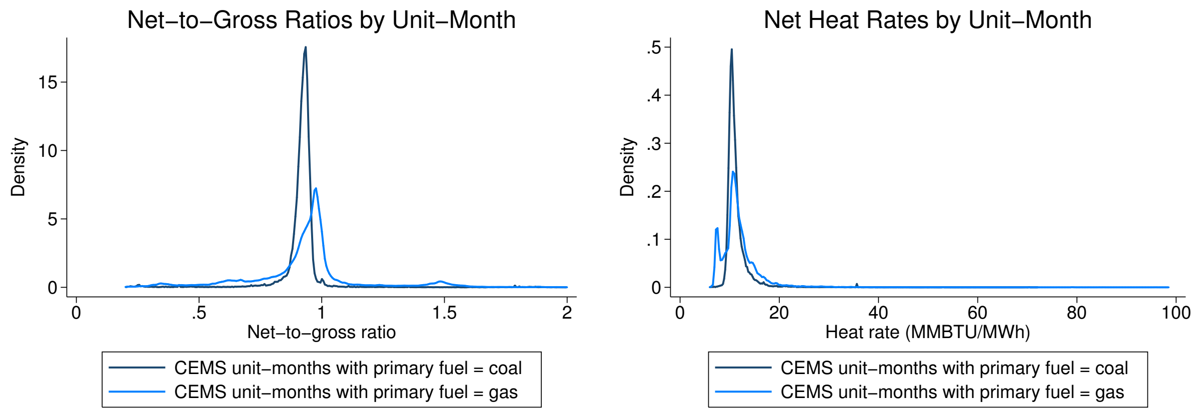
CEMS data report *gross* electricity generation, or the total amount of electricity generated in each unit-hour. However, a power plant's relevant unit of economic output is *net* electricity generation, which subtracts electricity that is generated but not sold to the electric grid.[36] Fortunately, EIA data report net generation at the unit-month level, and EIA's boiler-to-generator crosswalk enables me to compare net generation and gross generation for most unit-months. I calculate average net-to-gross ratios, which allow me to rescale hourly CEMS generation to more precisely measure electricity sold onto the

[34] See `http://ampd.epa.gov/ampd/`. These data files are available for bulk download at both daily and hourly temporal resolutions, at `ftp://ftp.epa.gov/dmdnload/emissions/daily/` and `ftp://ftp.epa.gov/dmdnload/emissions/hourly/`. For a detailed (archived) factsheet on CEMS data protocols, see `https://web.archive.org/web/20090211082920/http://epa.gov/airmarkets/emissions/continuous-factsheet.html`.

[35] For units whose primary fuel that alternates between coal and gas, I assume that they are 100 percent coal units when their primary fuel is listed as coal, and 0 percent coal units when their primary fuel is listed as natural gas. My demand estimation omit the portions of these units' time series during which their primary fuel is gas.

[36] Power plants use a small percentage of their electricity generation on-site, both as an input to electricity production (e.g., to power coal conveyor belts and the circulation of cooling water) and as a means of operating pollution control devices (e.g., to power scrubbers). Net generation is the unit of the electricity supply curve. Because the purpose of CEMS is to monitor pollution, CEMS data report gross generation inclusive of generation that is not sold to the grid (as this generation still contributes pollution).

Figure A.2.2: Kernel Densities: Net-to-gross Ratios and Heat Rates



Notes: This figure plots kernel densities for net-to-gross ratios and heat rates, where each CEMS generating unit has a separate monthly observation. I plot separate densities for unit-months where coal vs. gas is the primary fuel. The left panel reveals greater dispersion in net-to-gross ratios for gas plants, and that coal plants tend to expend more electricity on site (i.e. a lower average net-to-gross ratio). The right panel show how virtually all heat rates are between 6–20 MMBTU/MWh, with natural gas combined-cycle plants surpassing coal plants in terms of efficiency (i.e. lower heat rates, which is the inverse of thermal efficiency).

grid. In cases where I cannot merge across EIA-CEMS data, I use a linear projection to fill in missing unit-months. I drop extreme outliers with net-to-gross ratios less than 0.2 or greater than 2, following Cicala (2017).[37] The distribution of my assigned net-to-gross ratios is almost entirely concentrated between 0.91 and 0.94, meaning that the median plant sells (on average) 93 percent of its total generated electricity onto the grid. The left panel of Figure A.2.2 plots kernel densities of the net-to-gross ratio, separately for coal- vs. gas-fired units.

I also assign heat rates at the month-unit level, by dividing each unit's MMBTU of fuel consumed per month (from EIA boiler-level data) by the unit's monthly net MWh of electricity generated (from EIA generator-level data).[38] As with net-to-gross ratios, I follow Cicala (2017) and remove outliers with heat rates less than 6 MMBTU/MWh and

[37]Natural gas combined-cycle plants frequently have net-to-gross ratios as large as 1.4, meaning that they sell 140 percent of the electricity that they produce (according to CEMS) onto the grid. This reflects the fact the only the steam portion of the combined-cycle unit is included in CEMS, while the (combined) turbine cycle, which both generates electricity and contributes residual residual heat to the steam boiler, does not report to CEMS. My demand estimation strategy abstracts from power plants' startup and shutdown periods (as do Davis and Hausman (2016); and Cicala (2017)), and this net-to-gross calculation smooths generation expended (but not sold) during startup and shutdown periods across each unit-month.

[38]Heat rates summarize the thermal efficiency of power plants, allowing me to convert MMBTUs of fuel in to MWh of electricity out. I use EIA's reported fuel consumption rather than CEMS reported fuel input, because the latter appear to be less precisely measured than CEMS generation and emissions data.

Table A.2.1: Allowance Trading Programs for SO_2, NO_x, and CO_2

Program	Years in place	Geographic coverage	Pollutants traded
Acid Rain Program (ARP)	1995–present	48 states + DC	SO_2
Ozone Transport Commission (OTC) NO_x Budget Program	1999–2002	10 eastern states	NO_x (May–Sept)
State Implementation Plan (SIP) NO_x Budget Trading Program (NBP)	2003–2008	23 eastern states	NO_x (May–Sept)
Clean Air Interstate Rule (CAIR)	2009–2015	26 eastern states 26 eastern states 26 eastern states	SO_2 NO_x (May–Sept) NO_x (annual)
Cross-State Air Pollution Rule (CSAPR)	2015–present	23 eastern states 25 eastern states 28 eastern states	SO_2 NO_x (May–Sept) NO_x (annual)
Regional Greenhouse Gas Initiative (RGGI)	2009–present	10 eastern states	CO_2
California Cap and Trade Program	2013–present	California	CO_2

Notes: This table includes the five major allowance trading programs implemented by the EPA during my 2002–2015 sample period (ARP, OTC, NBP, CAIR, CSAPR). It also includes one regional program (RGGI) and one state-level program (CA cap and trade). My demand estimation incorporates participation of each program into marginal costs (i.e. through MC^{env} in Equations (A.26) and (A.27)), multiplying plant-specific program participation indicators by prevailing allowance prices and by unit-specific emissions factors. Many plants must purchase two separate allowances for NO_x emissions: one allowance to comply with the annual NO_x requirement; and one allowance to comply with the more stringent ozone season NO_x requirement (for NO_x emitted between May an September).

greater than 100 MMBTU/MWh. I also use linear projection to populate missing monthly heat rates within each plant. For both net-to-gross ratios and heat rates, a substantial share of CEMS units do not match to EIA generator-level data (due to incompleteness in the boiler-to-generator crosswalk). For these units, I assign net-to-gross ratios and heat rates first by plant-unit-type-month and then by plant-month, using linear projections to predict missing values between each step. The right panel of Figure A.2.2 plots kernel densities of heat rates, separately for coal- vs. gas-fired units; virtually all heat rates fall between 6–20 MMBTU/MWh.

The dependent variable in my demand estimation is CEMS hourly generation divided by each unit's capacity, or its capacity factor (CF_{uh} in Equations (1.6) and (A.29)). Rather than use each unit's nameplate capacity, which is inconsistently reported in EIA unit-level data, I assign monthly capacity equal to the maximum observed hourly generation for each unit in each month. This approach accommodates seasonal differences in plant capacity due to temperature variation, along with additional operational constraints. It also ensures that the capacity factor is between 0 and 1, by construction.[39]

I construct marginal costs for each fossil generating unit by multiplying unit-specific heat rates by monthly plant-specific fuel prices—from monthly EIA coal delivery data for coal units (as described above in Appendix A.2.1.1), and from daily hub-specific prices for gas units (as described below in Appendix A.2.5). I add marginal environmental costs using EIA data on variable environmental costs, net of revenues from selling byproducts (merged at the unit-level where possible, see Appendix A.2.2.1). I also add the implied marginal costs of SO_2, NO_x, and CO_2 emissions, for units that are covered by allowance trading programs. I use EPA Air Markets Program Data to assign monthly participation dummies for the 7 major allowance trading programs listed in Table A.2.1.[40] To monetize the implied costs of emissions under each of these programs, I multiply these unit-month-specific participation dummies by unit-month-specific emissions rates for each relevant pollutant (in units of tons per net MWh, as calculated from CEMS SO_2, NO_x, and CO_2 emissions), and by the average monthly allowance prices for each program (see Appendix A.2.6 for a discussion of allowance price data).

A.2.2.3 Non-Fuel Variable Costs

Both coal- and gas-fired electricity production are Leontief in fuel inputs. However, power plants use a variety of additional inputs that vary in the short and medium run, including coolants, maintenance, and repairs (Fabrizio, Rose, and Wolfram (2007)). I do not account for these additional variable costs in my coal demand estimation, mainly because reliable data on non-fuel, non-environmental operating costs are not widely available. In my application, I am less concerned with characterizing plants' production functions than with predicting generation conditional on the cost of one input to that production function (i.e. fuel). If I were able to credibly control for variation in these non-fuel, non-environmental costs (as a part of CR_{ud} in Equation (1.6)), this would potentially improve the precision of my coal demand estimates. At the same time, ignoring this variation is unlikely to induce bias, as it is unlikely to be correlated with a price-taking plant's fuel costs.

Cicala (2017) also omits non-fuel, non-environmental costs from his electricity dispatch estimation, arguing that these costs are second-order. By contrast, Davis and Hausman (2016) include time-invariant, technology-specific estimates of plants' variable operations

[39]Note this this ratio is identical for both gross and net generation, as applying a monthly net-to-gross conversion factor rescales both the numerator and denominator of CF_{uh}.

[40]For additional background information on these programs, see (in the order presented in Table A.2.1):
`https://www.epa.gov/airmarkets/acid-rain-program`
`https://www.epa.gov/airmarkets/ozone-transport-commission-nox-budget-program`
`https://www.epa.gov/airmarkets/nox-budget-trading-program`
`https://archive.epa.gov/airmarkets/programs/cair/web/html/index.html`
`https://www.epa.gov/csapr`
`https://www.rggi.org/`
`https://www.arb.ca.gov/cc/capandtrade/capandtrade.htm`

and maintenance costs. I conduct sensitivity analysis in Appendix A.4.2 that similarly incorporates technology-specific non-fuel cost estimates. I use SNL default cost assumptions to assign a time-invariant cost adder for each of four CEMS unit types: coal boilers (\$2.67/MWh), gas boilers (\$3.04/MWh), gas combined cycle plants (\$1.26/MWh), and gas turbines (\$6.81/MWh).[41]

A.2.3 Rail Data

A.2.3.1 GIS Rail Data

My analysis leverages detailed GIS data on the U.S. rail network published by the Bureau of Transportation Statistics (BTS), an agency within the U.S. Department of Transportation. I use four BTS shapefiles of the rail network, published in 2014, 2012, 2011, and 2006.[42] Each shapefile includes an accompanying dataset of rail line-specific attributes, including: identifiers for each line's two terminal nodes; the length of each line in miles; each line's type (e.g. mainline, non-mainline, abandoned); an indicator for passenger (as opposed to freight) lines; each line's primary, secondary, and tertiary owners (where applicable); and a list of rail carriers with trackage rights on each line.

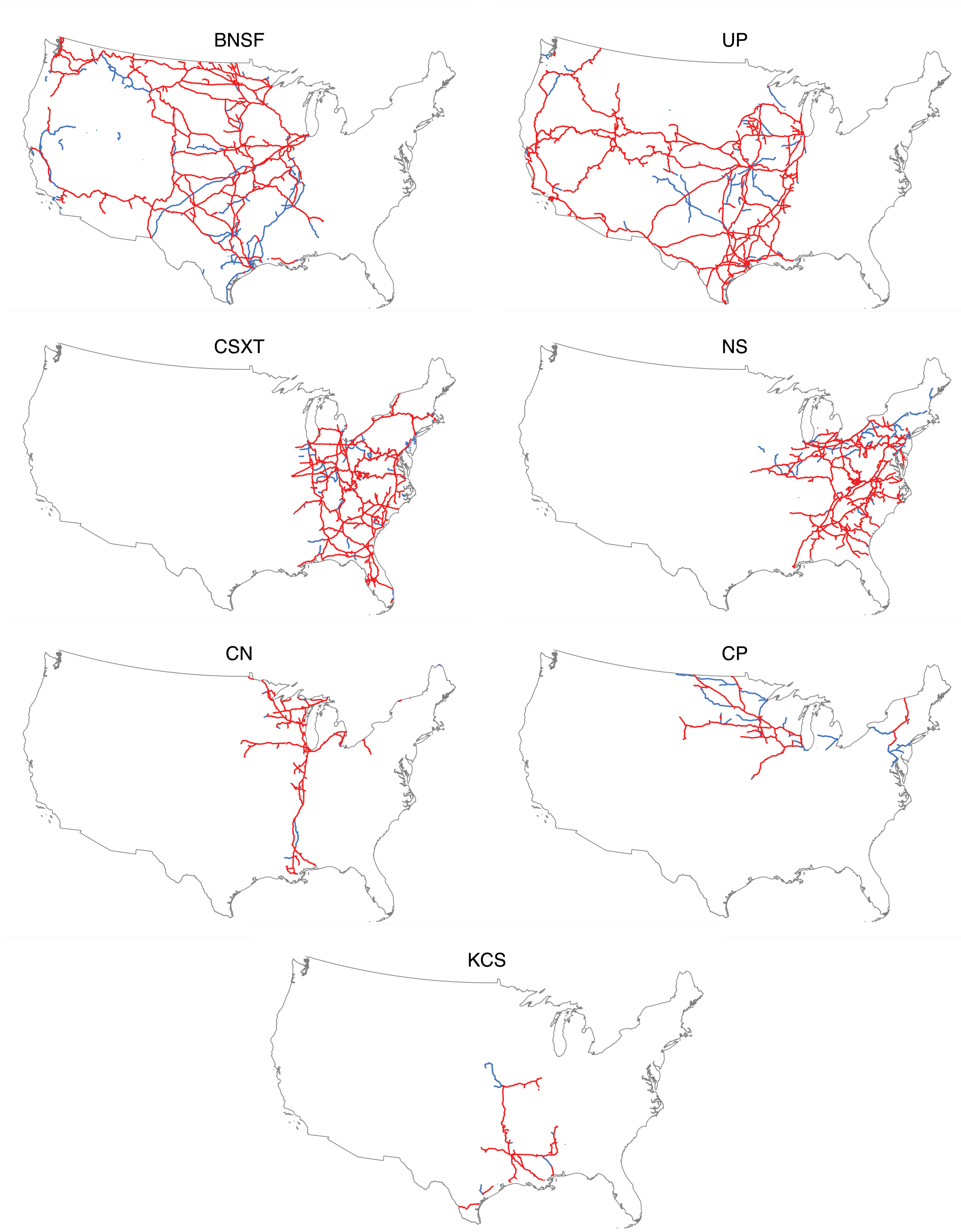
I focus exclusively on Class I rail carriers, or rail carriers with annual operating revenues greater than \$453 million. This revenue threshold associated with the Class I designation has increased gradually over time, and is defined by the Surface Transportation Board (STB). Currently there are 7 Class I rail carriers operating in the U.S.: BNSF and Union Pacific (UP) in the West; CSX (CSXT) and Norfolk Southern (NS) in the East; and Canadian National (CN), Canadian Pacific (CP), Kansas City Southern (KCS), three carriers with smaller geographic footprints. Figure A.2.3 maps the rail networks of each of these carriers separately, as a companion to the bottom panel of Figure 1.2.1 in the main text. Red lines are owned by their respective carriers, and are represented in Figure 1.2.1. Figure A.2.3 also includes rail lines for which each Class I carrier has primary trackage rights, but does not own (in blue).

I merge node and line identifiers across annual shapefiles, which reveals that the rail network was almost completely static between 2006 and 2014. 98 percent of mainline rail

[41] These defaults are averaged across 2009–2015 from SNL supply curves, available at `http://www.snl.com` ("Power" Menu → "Generation Supply Curve"; subscription required). SNL reports yearly non-fuel, non-environmental cost data for a subset of unit-years, reportedly taken from FERC Form 1. However, SNL data are only moderately correlated with FERC Form 1 data (available at `https://www.ferc.gov/docs-filing/forms/form-1/data.asp`), and each dataset covers only a fraction of CEMS units. Hence, I choose to assign default values consistently across all units.

[42] BTS GIS datasets for 2014, 2012, and 2011 are available at `https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/index.html`; and the 2006 shapefile is available at `https://catalog.data.gov/dataset/national-rail-network-1-100000-line-geographic-wgs84-bts-2006-us-rail-network-100k-lin-bts-2006`. The U.S. Geological Survey (USGS) has also published shapefiles of the U.S. rail network (available at `https://nationalmap.gov/small_scale/mld/1rails.html`); however, these datasets have less complete coverage and USGS node identifiers are not consistent across years. Shapefiles prior to 2005 are not publicly available online.

Figure A.2.3: Seven Class I Rail Networks



Notes: This figure maps the rail networks of the 7 Class I rail carriers. Red rail lines are owned and operated by each respective carrier, while carriers have primary trackage rights on blue lines (but do not own blue lines).

lines merge across all four shapefile years, with identical latitudes and longitudes of terminal nodes. Moreover, 99 percent of Class I track mileage maintained constant ownership between 2006 and 2014. This allows me to treat the rail network as time-invariant. A similar analysis in the 1980s or 1990s would need to account for the consolidation of Class I (and non-Class I) carriers; fortunately for my application, the last major Class I merger was in 1999.[43]

I use these GIS data to calculate the shortest rail distance between each originating coal county and each coal power plant. I also apply a remove-one-carrier algorithm to define rail captiveness, as detailed in Appendix A.3 below. As an additional shipping cost control, I proxy for rail network congestion using the reported rail traffic density of each line. BTS measures traffic density in million gross tons (MGT), and the 2014 (2011) shapefile reports densities for 2011 (2009) in discretized categories. I construct an indicator variable for high-density lines equal to 1 if a rail line has density greater than 50 MGT in 2011 *or* greater than 40 MGT in 2009, and equal to 0 otherwise. This classifies 11 percent of total track-miles as "high-density", and 18 percent of mainline track-miles as "high-density". I integrate this indicator variable across all track-miles on each oj shortest route to control for the fraction of each shipping route on high-density lines (as a component of $\mathbf{T}_{ojms}$ in my main regression specifications).

A.2.3.2 Rail Shipping Costs

I use the Association of American Railroads (AAR) fuel price index to control for changes in the cost of coal-by-rail shipping along each oj route. This index summarizes the change in the average price per gallon of No. 2 diesel fuel paid by the largest rail carriers, based on data from monthly surveys of rail operators. It is a single industry-wide number published monthly, which is inclusive of federal excise taxes, transportation, and handling expenses. The AAR constructs the fuel price index as a sub-component of the Surface Transportation Board's (STB) Rail Cost Adjustment Factor (RCAF). The RCAF is published quarterly (not monthly), and combines 7 different prices indices into a single number summarizing changes to rail freight costs: diesel fuel; labor; materials and supplies; equipment rents; depreciation; interest; and other expenses. The STB uses the RCAF as a basis for adjudicating rail rate cases.[44]

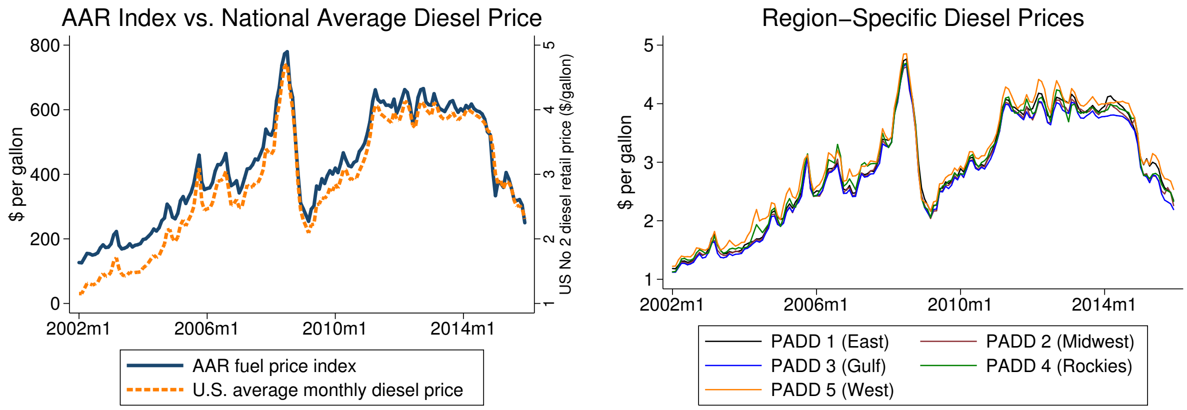
I construct a monthly fuel price index time series by from publicly available data from the U.S. Department of Transportation and from the AAR website.[45] I similarly

[43]For a comprehensive timeline of mergers between Class I railroads, see `https://en.wikipedia.org/wiki/Timeline_of_Class_I_railroads_(1977-present)`.

[44]Details on the RCAF and fuel price index are available at `https://www.aar.org/Documents/Rail%20Cost%20Indexes/Rail%20Cost%20Adjustment%20Factor%20RCAF/Index_RCAFDescription.pdf`.

[45]I use a spreadsheet of historic fuel price index values (2003–2012) from the Department of Transportation, available at `https://www.rita.dot.gov/bts/publications/multimodal_transportation_indicators/2013_02/fuel_prices/railroad_fuel`. For 2013–2015 months, I digitize PDFs published by AAR (e.g., `https://www.aar.org/Documents/Rail%20Cost%20Indexes/Index%20of%20Monthly%20Railroad%20Fuel%20Prices/MRF201409.pdf`). For 2002, I uses average monthly U.S. diesel prices to help extrapolate backwards (see next paragraph).

Figure A.2.4: AAR Fuel Price Index and Diesel Prices



Notes: The left panel compares the time series of the monthly AAR fuel price index to the U.S. national average No. 2 diesel prices. The correlation between the two price series is 0.99. The right panel plots region-specific diesel prices for the 5 PADD regions, revealing virtually no cross-sectional variation in diesel prices.

construct a quarterly RCAF time series from several historic documents published on both the STB and AAR websites.[46] In each case, I harmonize indices across years to construct a single apples-to-apples time series. I use the AAR fuel price index in my preferred regression specifications (i.e. as a component of $\mathbf{T}_{ojms}$) rather than the RCAF, because the RCAF provides only quarterly variation. Tables A.5.12 and A.5.30 demonstrate that my empirical results are not sensitive to this choice.

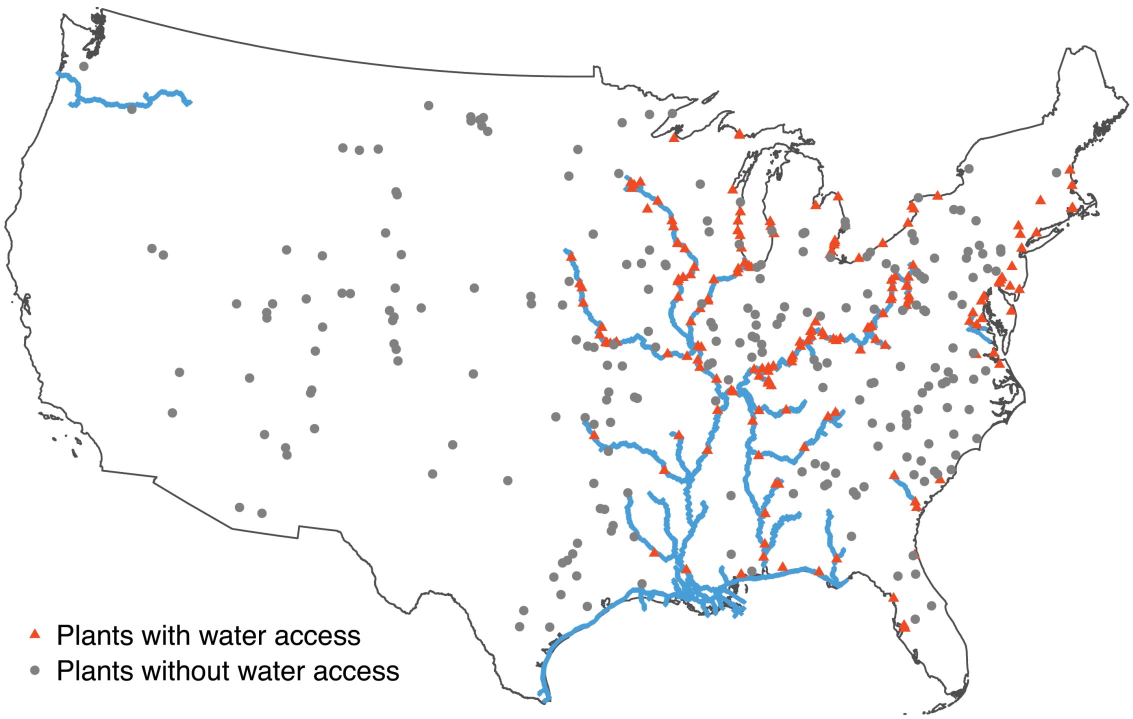
The left panel of Figure A.2.4 shows that the AAR survey-based price index closely tracks U.S. monthly average diesel prices, and the correlation between the two series is 0.99.[47] Regionally differentiated diesel prices offer the potential for capturing cross-sectional variation in rail shipping costs. However, PADD-specific price series are highly correlated, with all pairwise correlations greater than 0.99 (see the right panel of Figure A.2.4).[48] Hence, the AAR time series does not obsucre important cross-sectional variation in diesel prices.

[46] 2005–2015 RCAF indices is available at `https://www.aar.org/data-center/rail-cost-indexes`. For remaining years, I digitize PDFs from three locations: (i) `https://www.aar.org/Pages/RCAF-Quarterly-Filings.aspx`; (ii) `https://www.stb.gov/stb/industry/econ_rateindex.html`; and (iii) `https://www.stb.gov/Decisions/` (query "quarterly rail cost adjustment factor").

[47] Specifically, I use monthly No. 2 diesel retail price time series, which are available for download at `https://www.eia.gov/dnav/pet/PET_PRI_GND_A_EPD2D_PTE_DPGAL_M.htm`.

[48] During World War II, the U.S. created five Petroleum Administration for Defense Districts (PADDs) to manage the allocation of gasoline and diesel supply. These regions still exist for data-collection purposes. The five regions are defined as East Coast (PADD 1), Midwest (PADD 2), Gulf Coast (PADD 3), Rocky Mountains (PADD 4), and West Coast (PADD 5).

Figure A.2.5: Coal Plants with Access to Waterborne Shipments



Notes: This map plots all 430 coal power plants in my unrestricted sample, split into groups with/without access to coal-by-barge shipments. Red triangles indicate plants with a water option (i.e., $W_j = 1$), while grey circles denote plants without a water option (i.e., $W_j = 0$). Blue lines map all navigable U.S. rivers and the Gulf Intracoastal Waterway.

A.2.4 Distance to Navigable Waterway

I use GIS data to identify the subset of coal plants with access to waterborne coal shipments. These data come from three separate sources. First, I use the U.S. Army Corps of Engineers Waterway Mile Marker Database, which reports the locations of all navigable inland waterways in the U.S. (i.e., rivers and the Gulf Intracoastal Waterway).[49] Second, I use a shapefile of the Great Lakes from *Natural Earth*.[50] Third, I use a shapefile of U.S. coastlines, also from *Natural Earth*.[51] I calculate the distance of all coal plants to the nearest navigable river, Great Lake, and coastline, using the minimum of these three distances to determine plants' water access.

I construct a water access (or "water option") indicator (i.e. W_j from the main text) equal to 1 for all plants with a minimum distance to water of less than 1.5 miles. I choose this threshold because the vast majority of waterborne coal deliveries are to plants

[49] These data are available at `http://www.navigationdatacenter.us/data/datamile.htm`.

[50] These data are a subset of a global lakes shapefile, available at `http://www.naturalearthdata.com/http//www.naturalearthdata.com/download/10m/physical/ne_10m_lakes.zip`.

[51] These data are a subset of a global coastlines shapefile, available at `http://www.naturalearthdata.com/http//www.naturalearthdata.com/download/10m/physical/ne_10m_coastline.zip`.

within 1.5 miles of a navigable river, Great Lake, or coastline. At first pass, a threshold of 0 miles might seem appropriate, as barge deliveries typically go to plants physically adjacent to a river. However, the Army Corp of Engineers data record navigable rivers at 1-mile intervals, so this discreteness necessitates a non-zero threshold. I find that a 1.5-mile threshold correctly assigns several plants that receive exclusively waterborne deliveries despite being non-adjacent to navigable water—these plants have constructed long conveyor belts to carry coal overland from river barge offloading points to their coal stockpiles.

I cross-check this GIS-derived indicator with coal transportation modes listed in EIA's 423/923 delivery data. This reveals that 15 coal plants receiving a substantial share of coal deliveries by water are not located within 1.5 miles of a river, Great Lake, or coastline. After manually checking each of these plants in Google Earth, I find that they are all located a few miles from a Great Lake or coastline on small inlets (which do not appear in my GIS data). I correct each of these plants' indicators for water access, setting them equal to 1. Figure A.2.5 maps the full sample of 430 coal plants, split into water-access and non-water-access groups. I also map navigable inland waterways, which (importantly) represent only the largest U.S. rivers.[52]

A.2.5 Natural Gas Prices

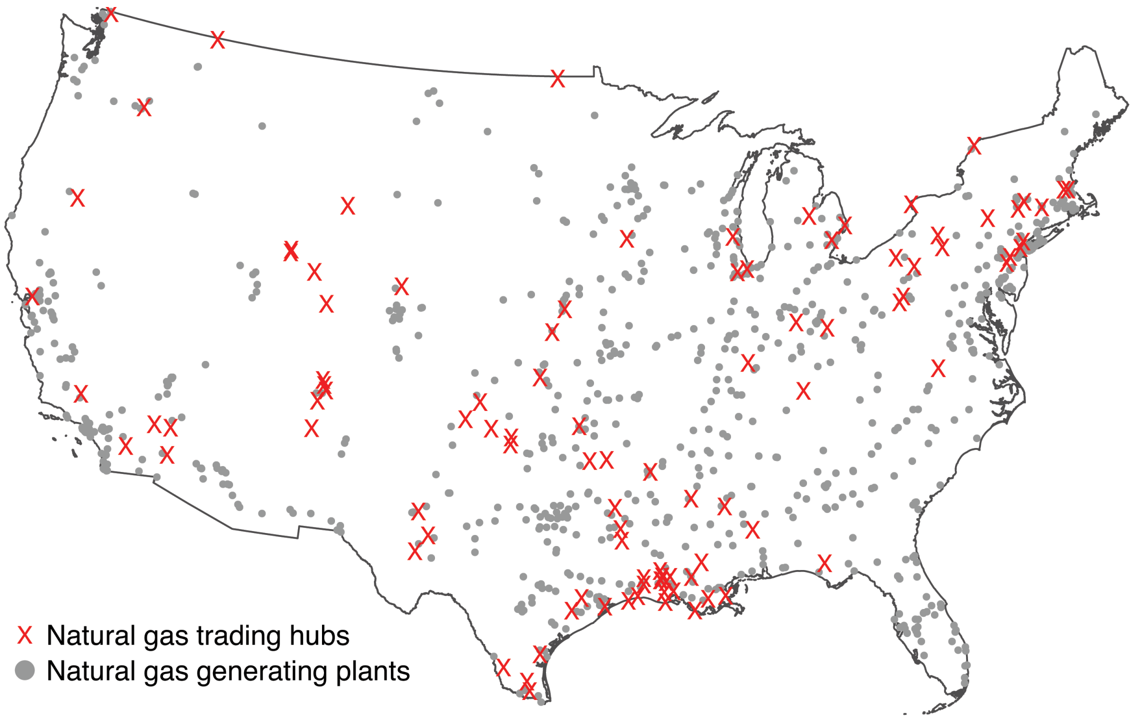
I use natural gas price data from SNL, which reports prices for 104 trading hubs at various locations throughout the U.S. gas pipeline network.[53] The primary trading hub is Henry Hub in Erath, Louisiana, which represents the standard commodity price for U.S. natural gas markets. I use the Henry Hub average monthly gas price as the time series component of my difference-and-differences regressions (Z_m^{HH} in Equation (1.9)).

Compared to coal, natural gas behave much more like a uniform-price commodity market—at least within the continental U.S. where there exists an extensive pipeline distribution network. However, regional variation in gas prices does exist due to variation in pipeline operating costs. For example, a plant far from gas producing regions will likely pay higher pipeline fees than a plant relatively close to gas production. More importantly, pipeline capacity constraints generate substantial dispersion in gas prices faced by power plants in different U.S. regions. This occurs frequently during winter months in New England, where sustained cold temperatures increase demand for natural gas heating beyond what the pipeline network can continuously supply.

[52]Water is a key input to thermal electricity generation—fossil fuel combustion creates heat, which boils water to create steam, and this pressurized steam rotates a turbine which drives an electrical generator. Hence, most coal power plants are located adjacent to some water source, such as a river or lake. Figure A.2.5 shows that many such water sources are not navigable, and could not feasibly convey coal barges.

[53]These data are available for download at `http://www.snl.com`, under the "Market Prices" menu (Advanced Search $\rightarrow$ Commodity), and require a subscription to access. The Henry Hub dail spot price series is widely available on line (e.g., at `http://www.eia.gov/dnav/ng/hist/rngwhhdd.htm`).

Figure A.2.6: Natural Gas Trading Hubs and Gas Plants



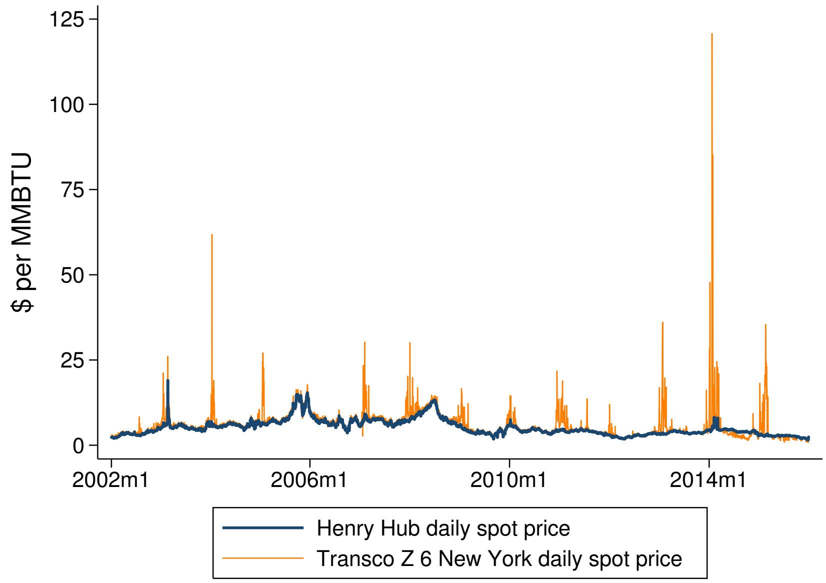
Notes: This map plots all natural gas trading hubs with available daily price time series from SNL (in red). It also plots all power plants with a CEMS gas-fired generating unit that is active during my sample period (in gray).

In my coal demand estimation, I incorporate both cross-sectional and time-series variation in gas prices. For each of SNL's 104 trading hubs, I manually assign geographic coordinates using SNL's mapping interface. Then, I match each gas-fired generating CEMS unit to its nearest trading hub (based on straight-line distance). SNL does not report complete daily time series for all 104 hubs, and I populate missing values in each series via linear projection.[54] Figure A.2.6 maps gas trading hubs and CEMS gas-fired generating units, which illustrates how I am able to assign very localized prices in some regions (i.e. Oklahoma) but not others (i.e. Florida). Figure A.2.7 illustrates how daily gas prices can wildly diverge across trading hubs.

While SNL reports hub-specific *wholesale* gas prices, power plants pay *retail* gas prices that reflect additional pipeline fees for the final portion of the gas distribution (i.e., the smaller pipelines that connect hubs to power plants). Electricity regulators and utilities also fund gas pipeline construction and investment cost recovery by increasing power plants' marginal fuel prices. I compare SNL monthly average hub prices to EIA 423/923

[54]Hubs that are geographically proximate tend to have highly correlated prices. In fact, using roughly 20 gas trading hubs (representing different regions of the pipeline network) would be sufficient to characterize nearly all of the cross-sectional dispersion in daily gas prices. My demand estimation would likely be quite similar if I matched gas plants only to trading hubs with complete daily time series in SNL.

Figure A.2.7: Natural Gas Daily Hub-Specific Prices (Example)



Notes: This figure plots daily SNL prices for two natural gas trading hubs: Henry Hub (in navy), and Transco Z 6 trading hub in New York (in orange). This illustrates how gas prices faced by power plants in the Mid-Atlantic region may diverge markedly (by over \$100/MMBTU) from gas prices faced by gas power plants in Louisiana, at the same point in time. The spikes in Transco Z 6 occur mostly in winter months, due to local gas supply shortages. For perspective, this Henry Hub price is the daily version of the monthly Henry Hub price series in Figure 1.2.2 from the main text (with a different vertical axis).

monthly average delivered gas prices, which allows me to estimate wholesale-to-retail price adjustment factors. I estimate price adjustment factors at the plant-month level (where possible) and then at the state-year level, using linear interpolation to populate missing plant-months. I then add each plant-month-specific adjustment factor to each plant's matched daily hub price to arrive at Z_{gd} in Equation (A.27), or the gas price paid by gas unit g on day d.

A.2.6 Allowance Prices for SO_2, NO_x, and CO_2

I construct time series of monthly allowance prices in order to monetize each fossil generating unit's marginal environmental costs. My primary data source for allowance prices is SNL, which reports average monthly prices for each of the emissions trading programs listed in Table A.2.1.[55] However, these data are incomplete and do not cover all months in my 2002–2015 sample. I supplement SNL allowance prices with several additional data sources, in order to arrive at full monthly time series:

[55] Allowance price data can be downloaded at `http://www.snl.com` from the "Market Prices" menu (SNL subscription required).

- ARP SO_2 allowance prices for 2002–2005 from BGC Environmental Consulting.[56]
- OTC seasonal NO_x allowance prices for 2002 from BGC Environmental Consulting.[57]
- NBP seasonal NO_x allowance prices for 2003–2005 from EPA annual progress reports.[58]
- CAIR annual NO_x allowance prices for 2009–2011 from EPA annual progress reports.[59]
- CSAPR prices for all three allowance types for 2015 from Evolution Markets.[60]
- RGGI quarterly CO_2 allowance auction results for 2008–2012 from RGGI.[61]

I use linear interpolation to populate months that are still missing. The resulting allowance price time series are obviously imperfect. Allowance markets were very thin during parts of my sample period, and low trading volumes likely explain many of the data gaps I encounter. In using prevailing allowance prices to monetize plants' marginal pollution costs, my goal is to approximate the contemporaneous allowance price signal to which plants optimized generation decisions. Even if SNL time series were complete across all sample months, these prices would likely still mismeasure plants' own (interpretation of their) shadow costs of SO_2, NO_x, and CO_2 emissions. Figure A.2.8 plots allowance price time series for SO_2, CO_2, seasonal NO_x, and annual NO_x. There has been tremendous variation in SO_2 and NO_x prices, and large spikes/drops reflect adjustments in expectations immediately following (announced) policy changes (Schmalensee and Stavins (2013)).

A.2.7 Temperature Data

Finally, I use PRISM weather data to assign the daily maximum temperature at each power plant location. The PRISM Climate Group at Oregon State University maintains daily (and monthly) spatial datasets of temperature, precipitation, dew point, and vapor pressure, for the conterminous United States. Each daily dataset incorporates readings

[56] http://www.bgcebs.com/registered/aphistory.htm

[57] http://www.bgcebs.com/registered/apnx0623.htm

[58] https://www.epa.gov/sites/production/files/2015-08/documents/noxreport03.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/ozonenbp-2004.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/2006-nbp-report.pdf

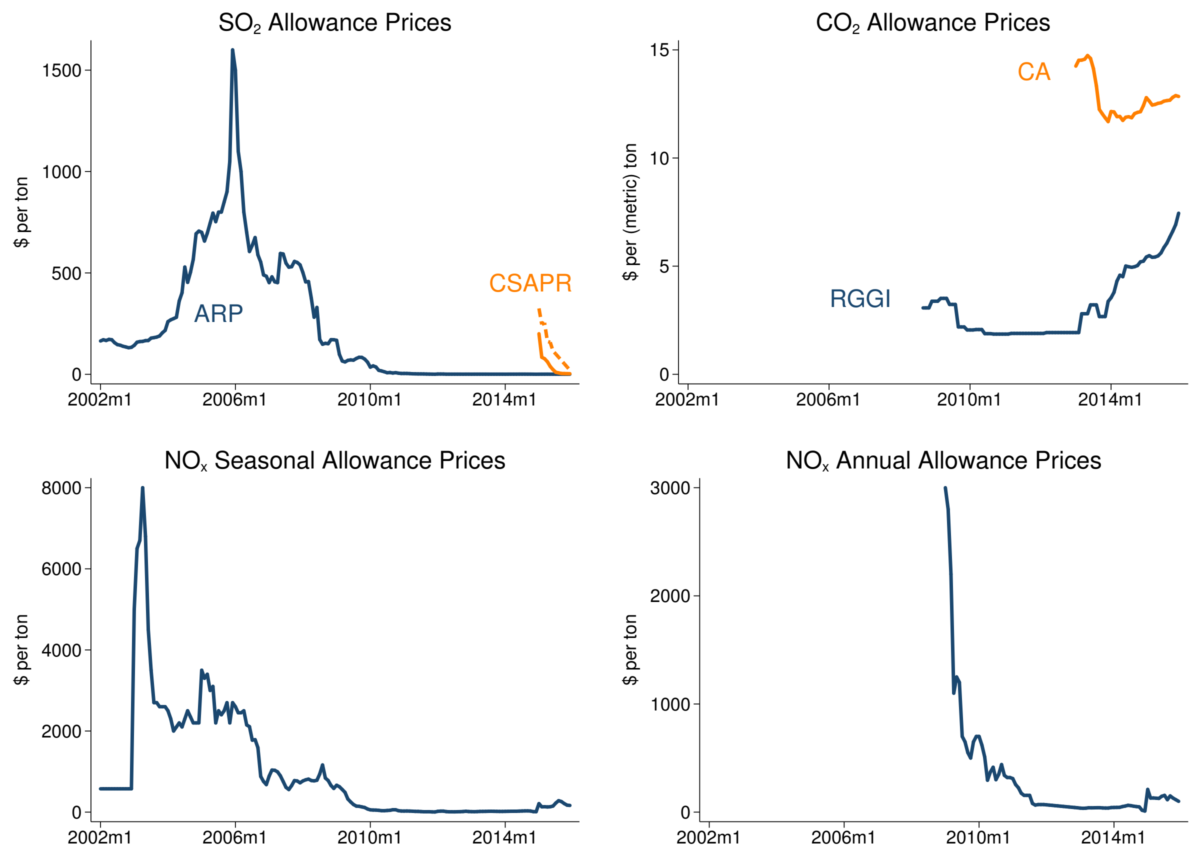
[59] https://www.epa.gov/sites/production/files/2015-08/documents/cair09_ecm_analyses.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/arpcair10_analyses.pdf; https://www.epa.gov/sites/production/files/2015-08/documents/arpcair11_analyses_0.pdf

[60] http://www.evomarkets.com/content/news/reports_10_report_file.pdf; http://www.evomarkets.com/content/news/reports_12_report_file.pdf

[61] http://www.rggi.org/market/co2_auctions/results

across 20 separate networks of weather stations, and applies a spatial interpolation algorithm to produce gridded rasters at 4km resolution.[62] I project power plant coordinates onto each day's maximum temperature raster to construct a panel of daily maximum temperature at each plant location.

Figure A.2.8: Allowance Price Time Series



Notes: This figure plots allowance price time series for the four tradable allowance types. Orange lines in the top-left panel plot SO_2 allowance price for two groups of plants under CSAPR. The orange line in the top-right panel plots California cap-and-trade allowance prices (in $ per metric ton), while the blue line plots RGGI allowance prices (in $ per short ton). For both type of NO_x allowance price series, I plot a single prevailing allowance price series that spans 2–4 distinct policy periods. All prices are nominal.

[62] I use PRISM's "AN81d" product, which exists at two resolutions: 4km (for free) and 800m (for a fee). PRISM data are available for download at `http://www.prism.oregonstate.edu`, along with extensive documentation.

A.3 Rail Graph Algorithm

In this section, I describe my algorithm estimating the shortest distance along the rail network between each coal-producing county and each coal power plant, and to construct an indicator for rail captiveness. I begin with a time-invariant GIS dataset of all active rail lines and rail terminal nodes in the contiguous U.S., which I describe in Appendix A.2.3 above. Next, I overlay a map with the coordinates of all coal power plants, and an additional map with the coordinates of the production-weighted centroids of each coal-producing county.[63] Then, I calculate the geographically closest (as the crow flies) rail node to each coal plant, and to each coal county's production-weighted centroid.[64]

I apply a graph algorithm to calculate the shortest distance along the rail network between each pairwise combination of origin nodes (i.e. rail nodes matched to county coordinates) and destination nodes (i.e. rail nodes matched to plant coordinates). I begin by converting the rail network from a GIS dataset into a graph object defined by three elements: (i) a list of nodes (i.e. rail nodes); (ii) a list of edges (i.e. rail lines), where each edge is defined by the two nodes that it connects; and (iii) a weight corresponding to each edge, equal to the distance corresponding to that edge (i.e. GIS-derived mileage of each rail line).[65] Using Dijkstra's algorithm, I calculate the shortest path along all possible edges that connect each pair of origin and destination nodes, weighting edges by their distance.[66] This shortest distance for each route oj is the variable that enters all of my markup regressions in $\mathbf{T}_{ojms}$, to control for the distance component of rail shipping costs.

Importantly, my strategy for calculating shortest rail shipping distance does not account for ownership of rail lines. This means that my algorithm may calculate a shortest route that is owned or operated by multiple Class I carriers. In practice, rail carriers have track-sharing agreements, and the Surface Transportation Board prevents carriers from extracting excessive rents for usage of short, pivotal segments of track. For example,

[63] I calculate the average geographic coordinates of all coal mines in a given county, weighted by total mine production during my sample period (using MSHA production and mines datasets). This yields one pair of coordinates for each coal county that is time-invariant, which more accurately approximates the location of coal production than a geographic county centroid. In principle, I could match rail nodes to individual coal mines, which I could then match to coal shipments based on their originating mines. Unfortunately, EIA data do not report mine identifiers prior to 2008, which would lead to measurement error that varied systematically across years of shipment data.

[64] I use the `SpatialPoints` function in `R`'s `sp` package; and the `distGeo` function in `R`'s `geosphere` package, which calculates distances between geographic coordinates using the WGS84 reference ellipsoid.

[65] I use the `readshpnw` function from `R`'s `shp2graph` package to convert the three-element GIS dataset of rail lines into a `SpatialLinesDataFrame` object, with indexed nodes and edges. Then, I apply the `nel2igraph` function (also from `R`'s `shp2graph` package) to convert this `SpatialLinesDataFrame` object into an `igraph` object.

[66] I use the `distance` function from `R`'s `igraph` package to calculate the length of each shortest paths, and the `shortest_paths` function (also from `R`'s `igraph` package) to extract an ordered list of edge identifiers along each shortest path. Hughes (2011) applies a similar algorithm to calculate the shortest rail distance for U.S. ethanol shipments. The author compares the estimated shortest-distance paths along BNSF's network to BNSF's own reported distances, and finds that the distribution of the ratio of estimated-to-actual distances has a mean of 0.96 and a standard deviation of 0.03. This suggests that GIS-derived shortest distances closely approximate (yet slightly understate) actual rail shipping distances.

consider a shipping route with 5 rail nodes, ordered A-B-C-D-E, where Carrier 1 owns segments AB, BC, and CD, but Carrier 2 owns segment DE. Regulators would prevent Carrier 2 from charging Carrier 1 excessive rents for the right to access segment DE on an ABCDE shipment, if DE is relatively short (U.S. Government Accountability Office (2006)). Even in the absence of regulatory oversight, Carrier 1 would likely have the ability to hold up Carrier 2 along a different shipping route, and track-sharing agreements would be the likely equilibrium outcome of this cooperative multiple-route repeated game. An alternative graph algorithm would force each shortest path to have uniform Class I ownership across all connecting segments; given that carriers share tracks in practice, this strategy would impose excessive (unrealistic) structure on the rail network.

To derive rail captiveness, I iterate the above algorithm 7 times, for 7 restricted rail networks. Each restricted network removes all rail nodes and lines that are owned or operated by 1 of the 7 Class I carriers. I define two concepts by comparing the unrestricted network to the 7 restricted networks: "node unconnectedness" and "route unconnectedness". A plant becomes "node-unconnected" if the removal of any single Class I carrier renders that plant unconnected from the rail network. A plant becomes "route-unconnected" if it becomes unconnected from all observed trading partners (i.e. origin counties) after removal of the dominant Class I carrier along each (unrestricted) shortest route. I define the set of captive as the union of the set of plants that become node-unconnected and the set of plants that become route-unconnected.

A.3.1 Node Unconnectedness

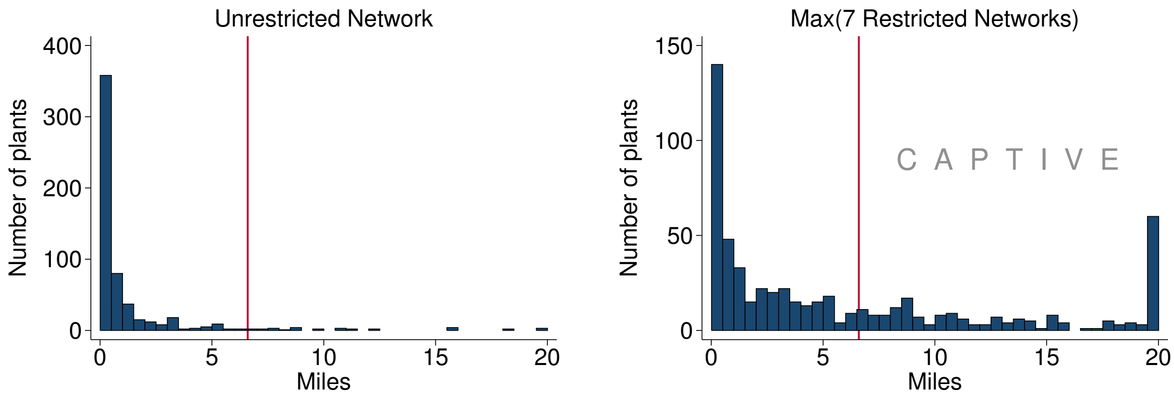
Conceptually, a plant becomes node-unconnected if a single Class I carrier controls all rail nodes from which it can potentially receive coal. In practice, I must establish a threshold beyond which a node is sufficiently far from a plant to be a feasible coal delivery point. I set a threshold of 6.6 miles for node unconnectedness, based on the 95th percentile of the distribution of plants' distance to the nearest node on the *unrestricted* network. In other words, using the full network of active rail lines, 95 percent of coal plants have a rail node that is within a 6.6 mile radius.[67]

This 6.6-mile threshold is exceedingly conservative in terms of geographic measurement error. It is possible that for a power plant with a large geographic footprint, the coordinates of the plant's flue gas stack (which corresponds to EPA's reported latitude and longitude) are far from the plant's physical stock pile of coal. However, it is extremely unlikely that a plant's coal pile is 6.6 miles away from its boilers. It is also possible, albeit highly unlikely, that the location of a plant's rail node is actually over 7 miles away from its physical coal offloading point.[68] In such cases, BTS shapefiles would be inadequate for

[67] In many cases, this nearest node is controlled by a smaller rail carrier, rather than one of 7 Class I carriers. Such a plant cannot become node-unconnected, since the 7 restricted networks only remove Class I nodes.

[68] A few plants have constructed small rail lines to ferry coal from the rail carrier's node directly to their coal offloading points, and these "driveways" are too small to appear in BTS shapefiles. However, rail nodes are virtually always closer to plants than 6.6 miles. ("Driveway" is not part of standard railroad

Figure A.3.9: Distance to Nearest Rail Node



Notes: The left panel reports a histogram of plants' distances to the nearest rail node on the unrestricted rail network. The red line denotes the 95th percentile of this distribution, which is 6.6 miles. The right panel reports the histogram of the plants' maximum distances to the nearest rail node across all 7 restricted rail networks (because the nearest unrestricted node can only be controlled by 1 of the 7 Class I rail carriers, I plot maximum across all 7 restricted networks). Both histograms are top-coded at 20 miles. For plants in the "captive" region of the right panel, *all* rail nodes within a 6.6-mile radius are controlled by the same Class I carrier.

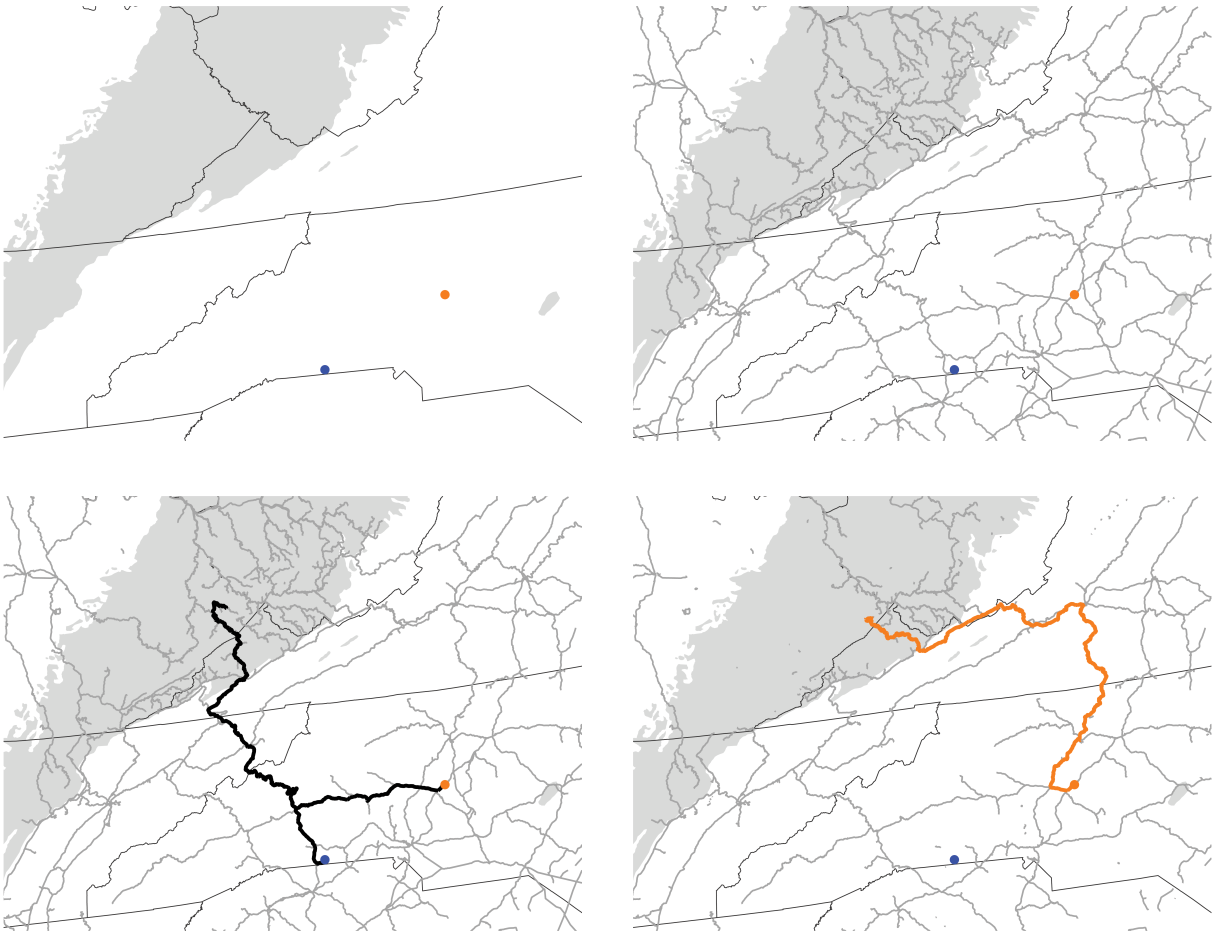
determining rail captiveness. Apart from geographic measurement error, 6.6 miles is an extremely conservative distance buffer. It would almost certainly be cost prohibitive to regularly transport multi-ton carloads of coal 6.6 miles over land by a mode *other* than rail.

Figure A.3.9 shows histograms of plants' distance to the nearest rail node, along with vertical lines marking my 6.6-mile threshold. The left panel shows how virtually all plants are within 3 miles of the nearest node, on the unrestricted rail network. The right panel shows the maximum distance to the nearest rail node, across all 7 restricted rail networks. This illustrates how for many plants, removing a single rail carrier can render the plant infeasibly far from the nearest rail node. I consider plants to the right of the 6.6-mile threshold in the right panel to be "captive" to a single rail carrier.[69] In Tables A.5.10 and A.5.28, I show that my econometric results are robust to a node unconnectedness threshold of 5 or 10 miles. Figure A.3.10 shows an example of how I determine node unconnectedness.

parlance, and I merely use this term as an analogy—coal plants may not be physically adjacent to the rail network just as a residence may be connected to its nearest roadway via a long driveway.)

[69]There are 13 plants in the left panel of Figure A.3.9, which are "node-unconnected" even in the unrestricted network. This likely reflects errors in my rail graph algorithm, and I classify these plants as "non-captive" throughout my analysis. In other words, I assume that plants are non-captive unless they are shown to be captive.

Figure A.3.10: Example of Node Unconnectedness



Notes: This figure illustrates how I calculate the shortest path along the rail network between originating coal counties and coal plants. It also applies my definition of route unconnectedness. The top-left panel shows two coal plants in North Carolina that are 85 miles apart (approximately the median distance between nearest-neighbor matched pairs). The top-right panel overlays the full (unrestricted) rail network. The bottom-left panel applies Dijkstra's algorithm to calculate each plant's shortest rail shipping path to the coal production-weighted centroid for Perry County, Kentucky (a county from which each plant purchases coal via rail). For both shortest routes (in black), the dominant Class I rail carrier is CSX. The bottom-right panel removes CSX rail lines and calculates the new shortest path to Perry County for the orange plant (in orange), which increases by only 24 miles compared to the unrestricted shortest path. However, after removing CSX rail lines, the blue plant is now 6.8 miles from its nearest rail node; this is past my threshold for node unconnectedness and renders the blue plant unconnected from the (restricted) rail network. Hence, I classify the blue plant as "captive".

A.3.2 Route Unconnectedness

Conceptually, a plant becomes route-unconnected if each of its coal shipping routes is controlled by a dominant Class I rail carrier. This need not the same rail carrier for each route, and the dominant carrier on a given route may not control 100 percent of the rail lines along the (shortest) route. In practice, I must evaluate the extent to which coal shipments between a given origin and destination are controlled by a single railroad.

I define route unconnectedness by comparing the lengths of the shortest oj-connecting paths calculated on unrestricted vs. restricted rail networks. For each oj pair, I compare the length of its unrestricted shortest path to the length on the restricted network that removes the route's dominant carrier (i.e. the Class I carrier that controls the most miles of track along the unrestricted shortest path). For example, suppose the shortest o-to-j path has 5 rail nodes, ordered A-B-C-D-E, where Carrier 1 owns segments AB, BC, and CD, but Carrier 2 owns segment DE. If the length of AB+BC+CD is greater than the length of DE, then I would compare length of the unrestricted path (i.e. AB+BC+CD+DE) to the length of a shortest path on the restricted network that removes Carrier 1. On the other hand, if the length of DE is greater than the length of AB+BC+CD, then I would compare the length of the unrestricted path (i.e. AB+BC+CD+DE) to the length of a shortest path on the restricted network that removes Carrier 2.

I base this comparison on each route's dominant (i.e. modal) Class I carrier, as this carrier is most likely be the firm that transacts coal deliveries. An alternative strategy would consider how shortest distance changes after removing *any* Class I carrier along an o-to-j shortest route. However, this would increase the dimensionality of route unconnectedness, while potentially overweighting very short rail segments. Using the above A-B-C-D-E example, suppose that removing Carrier 1 would increase the length of the shortest route by 50 miles, while removing Carrier 2 would increase the length of the shortest route by 500 miles. Under this alternative strategy, I would need to trade-off the 50- vs. 500-mile increases against the relative importance of Carrier 1 vs. Carrier 2 along the unrestricted shortest path. For this reason, I use the more straight-forward approach that limits route unconnectedness to a single unrestricted-vs.-restricted comparison per route.[70]

I apply a 300-mile threshold to determine route unconnectedness. That is, an origin-destination pair becomes route-unconnected if the length of its shortest connecting path increases by at least 300 miles after removing the dominant carrier along its unrestricted shortest path. This threshold is essentially arbitrary, and I choose 300 miles because a 300-mile increase in rail shipping distance would imply roughly a 20 percent increase

[70] As an artifact of my decision to not consider rail line ownership in calculating shortest paths, my algorithm often returns a shortest path with lines from multiple carriers, even though a similar slightly longer path also exists where a single carrier controls all lines. Using the same A-B-C-D-E example, there may exist a slightly longer route A-B-C-D-F-E in which all segments are owned by Carrier 1 (i.e. the combined lengths of DF+FE is slightly longer than segment DE). In this case, removing Carrier 2 from the network would only increase the shortest route slightly, and the route would not become unconnected. If my measure of route unconnectedness considered the removal of Carrier 2 *as well as* the removal of Carrier 1, it would overweight the importance of segment DE.

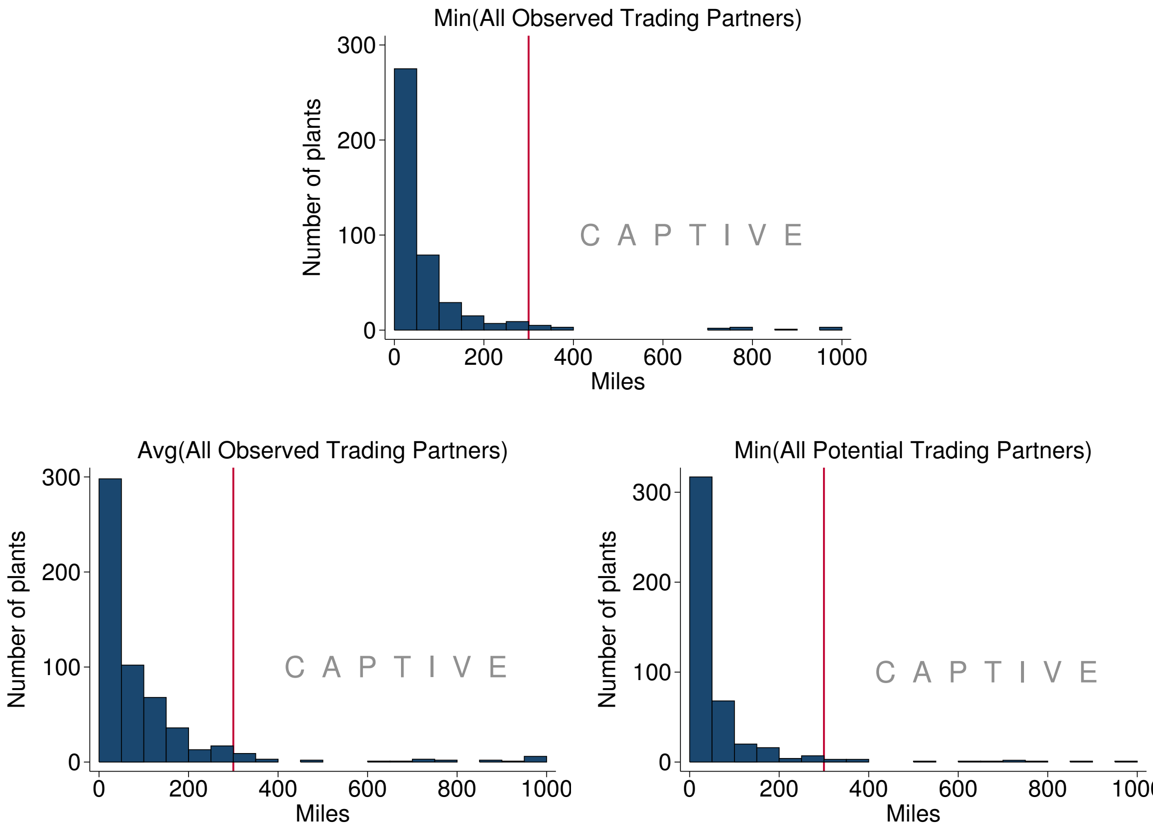
in the median delivered coal price in my sample.[71] However, in many cases, route-unconnectedness is not sensitive to a particular mileage threshold, because removal of the dominant carrier eliminates all paths connecting connecting origin o and destination j—rendering the restricted shortest path infinitely long.[72]

I consider plant j to be rail captive if it becomes route-unconnected from *all* observed trading partners, or *every* county from which it purchased coal between 2002 and 2015. The top panel of Figure A.3.11 reports a histogram of the minimum increase in mileage for each coal plant, or the route-specific component of my definition of rail captiveness. This reveals that route unconnectedness is relatively less important than node unconnectedness in my definition of rail captiveness. The bottom-left panel weakens the definition of captiveness, based on route-unconenctness of the *average* coal shipment to plant j; this prevents relatively underutilized coal routes from preventing an otherwise captive plant from being categorized as non-captive. On the other hand, the bottom-right panel strengthens the definition of captiveness, bassed on route-unconnectedness across all of plant j's *potential* trading partners; this allows for an un-utilized route that does not become unconnected to render a seemingly captive plant non-captive. Figure A.3.11 reveals that my choice of routes has little effect on which plants meet the definition of "captive". In Tables A.5.9 and A.5.27, I show that my econometric results are robust to a halving/doubling of the 300-mile threshold for route-unconnectedness, as well as to my choice of *which* routes to use to determine captiveness. Figure A.3.12 shows an example of how I determine route unconnectedness.

[71] During my sample period, the median rail shipping rate reported by EIA is \$0.025 per ton-mile. Given this rate, a 300-mile increase in rail shipping distance implies an increase in \$7.50 per ton, which is 20 percent of the the median delivered coal price of \$38/ton.

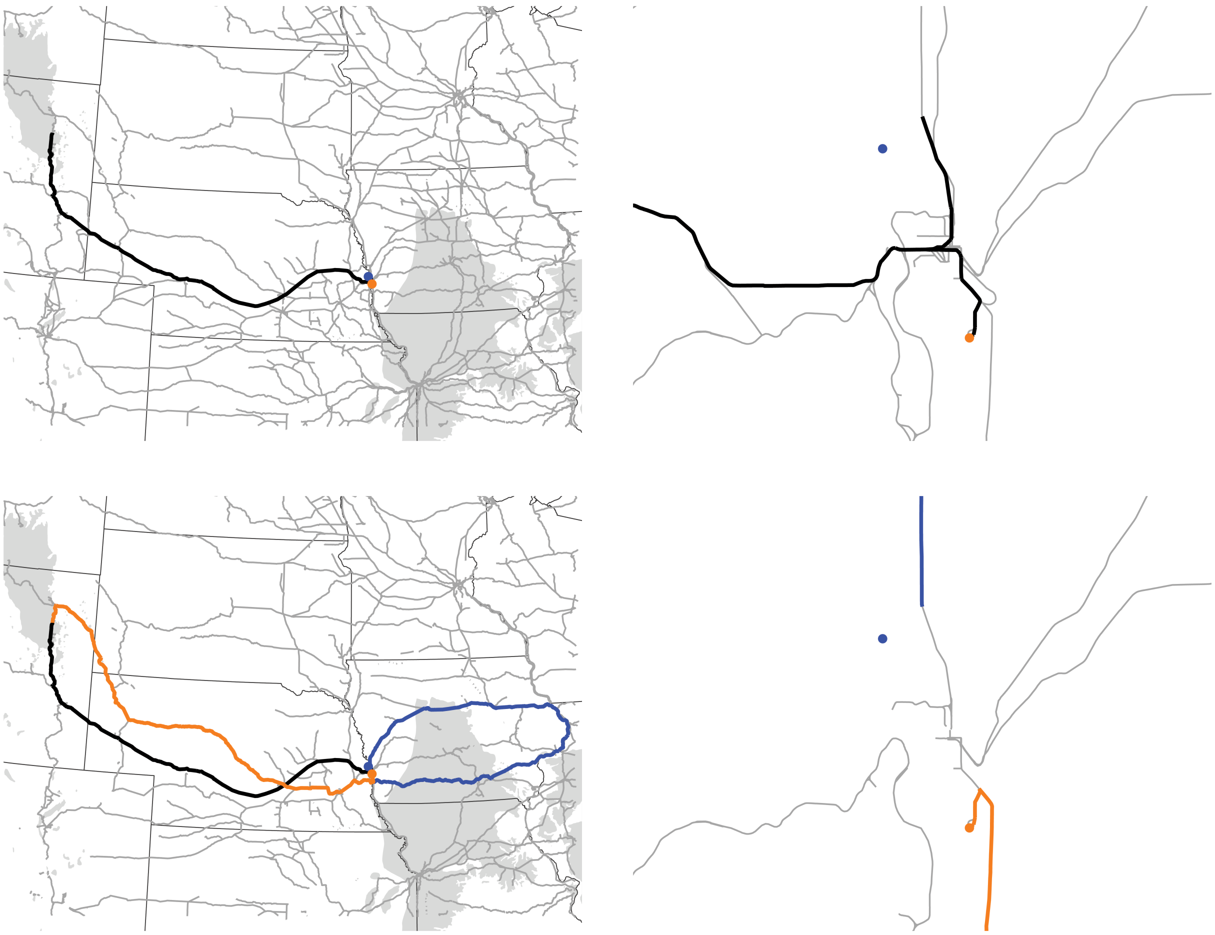
[72] The elimination of all connecting routes could result from a plant's nearest node becoming unconnected from a portion of the rail network, *or* from the removal of a plant's nearest node. I incorporate my threshold for node unconnectedness (discussed above) in determining whether plants j's closest node in the restricted network is indeed close enough to be a feasible coal delivery point. I also use a similar concept to define "origin-node-unconnectedness", whereby a coal-producing county becomes unconnected from the rail network. I apply an extremely conservative threshold of 42 miles, which is the 99th percentile of unrestricted nearest-node distance and greater than diameter of of virtually all coal counties.

Figure A.3.11: Distance to Nearest Rail Node



Notes: The top panel reports a histogram of plants' minimum increase in shortest shipping path, across all observed trading partners, where the red line denotes the 300-mile threshold. I consider the few plants to the right of this 300-mile threshold to be rail captive. The bottom panels show alternative captiveness definitions, either based on the average increase in the length of the shortest path (weight-averaged across all observed coal deliveries) or the minimum increase across all potential shortest paths (i.e. counties that produce coal with similar attributes to plant j's purchased coal, but which are not observed trading partners).

Figure A.3.12: Example of Route Unconnectedness



Notes: This figure shows an example of rail captiveness, for two plants that are 12 miles apart, near Omaha, Nebraska. Each plant purchases coal exclusively from the same county in Wyoming's Powder River Basin. The top-left panel maps the shortest path from this county to each plant, as calculated by applying my graph algorithm to the unrestricted rail network. The top-right panel zooms in, to illustrate how this shortest route bifurcates close to each plant. The blue (northern) plant is 3 miles from the nearest rail node, and I assign this closest node because it is within a 6.6-mile radius. (This plant receives 100 percent of its coal shipments by rail, and it would obviously be incorrect to consider this plant unconnected from the rail network. Google Earth images confirm that this plant is connected to the highlighted rail route via a small set of rail tracks, which are not included in BTS shapefiles.) Union Pacific is the dominant carrier along this shortest unrestricted route, and the bottom panels show the new shortest routes after restricting the network to exclude Union Pacific rail lines. The orange (southern) plant sees its new shortest route (in orange) increase by 40 miles. However, the blue (northern) plants sees its new shortest route (in blue) increase by 700 miles. This disparity arises because the blue plant must now be approached from the north, while the orange plant may be approached from the south. Because its increase in mileage is greater than 300 miles, I consider the blue plant to be captive.

A.4 Coal Demand Estimation

Here, I provide a more thorough treatment of my algorithm for estimating coal demand at the power plant level. Appendix A.4.1 describes each step in this estimation procedure in detail. Appendix A.4.2 provides additional results to supplement Figure 1.5.6 in the main text, while also reporting sensitivity analysis on Equation (1.6).

A.4.1 Demand Estimation Algorithm

In this section, I walk through my strategy for estimating plant-specific coal demand parameters as a function of the price of natural gas. First, I estimate a semi-parametric model of each coal generating unit's operations in each hour, conditional on the relative marginal costs of coal vs. natural gas generation. Next, I construct a distribution of counterfactual coal prices at which each unit would have been exactly marginal in electricity dispatch. Then, I transform and aggregate these distributions into quantity-price mappings, yielding plant-month-specific coal demand curves. Finally, I estimate how changes in natural gas prices affect both the level and the slope of each plant's inverse coal demand.

Step 1: I construct a coal-to-gas cost ratio by dividing each coal unit's marginal cost of generation by the generation-weighted average marginal cost of gas-fired units in its PCA. For both coal and gas units, I multiply unit-specific fuel prices (P for coal, Z for gas) by unit-specific heat rates (HR), and add the unit's marginal costs of environmental compliance (MC^{env}).[73] Gas plants typically have limited capacity to store fuel on site, meaning that short-run price changes can impact plant operating decisions. Hence, I assign each gas unit the daily spot price of its closest natural gas trading hub, in order to capture the effect of day-to-day price fluctuations.[74] By contrast, coal plants can cheaply store fuel on site, causing their opportunity cost of coal to respond more slowly to price changes. I use the average delivered coal price at the plant-month level, which

[73] MC^{env}_{um} captures unit u's opportunity cost of SO_2, NO_x, and CO_2 emissions in month m, scaled by average monthly allowance prices (A^z_m) and unit-specific emissions rates per MMBTU (E^z_{um}), for each pollutant z. It also includes the non-energy operating costs of scrubbers (i.e. flue gas desulfurization to reduce SO_2 emissions), net of the marginal revenues from selling the gypsum byproduct of the desulfurization process:

$$MC^{env}_{um} = \sum_{z \in \{SO_2, NO_x, CO_2\}} A^z_m E^z_{um} \cdot \mathbf{1}[\text{in } z \text{ trading program}]_{um} + MC^{scrubber}_{um}$$

I abstract from non-fuel, non-environmental variable operating costs, because these data are generally thought to be unreliable across all years and units.

[74] I use daily natural gas prices from SNL for 104 trading hubs, as described in Appendix A.2.5.

is the finest temporal resolution that EIA reports publicly.[75] I also assign heat rates and environmental costs at the monthly level, for each unit.

Indexing coal units by u, gas units by g, months by m, and days by d, the daily cost ratio is:

$$MC_{um}^{coal} \equiv HR_{um} \cdot (P_{jm} + MC_{um}^{env}) \tag{A.26}$$

$$MC_{ud}^{gas} \equiv \sum_{g \in \text{PCA}_u} \left(\frac{Q_{gm}^{elec} \cdot HR_{gm} \cdot (Z_{gd} + MC_{gm}^{env})}{\sum_{g \in \text{PCA}_u} Q_{gm}^{elec}} \right) \tag{A.27}$$

$$\Rightarrow \quad CR_{ud} = \frac{MC_{um}^{coal}}{MC_{ud}^{gas}} \tag{A.28}$$

Steps 1–3 of this estimation strategy treat the coal *unit* as the relevant unit of analysis, rather than the coal plant. This is because a single power plant may comprise both coal and gas generating units, and because individual coal units within the same plant have different heat rates and environmental costs, implying different marginal costs for a given fuel price.

Step 2: For each coal unit u, I estimate the following binned time series regression, for all hours h, from 2002 to 2015 (Equation (1.6) in the main text):

$$CF_{uh} = \sum_b \alpha_{ub} \mathbf{1}[G_{uh} \in b] + \sum_b \gamma_{ub} \mathbf{1}[G_{uh} \in b] \cdot CR_{ud} + \zeta_u CR_{ud} + \xi_u \mathbf{G}_{uh} + \omega_{uh} \tag{A.29}$$

$CF_{uh} \in [0, 1]$ is unit u's operating capacity factor in hour h, where $CF_{uh} = 0$ when the unit is off and $CF_{uh} = 1$ when the unit is generating at full capacity. G_{uh} is aggregate net generation in hour h, summed across all CEMS electric generating units in unit u's market region.[76] This is not equivalent to aggregate electricity "load", which includes non-CEMS generation such as nuclear, hydro, renewables. However, these other technologies typically

[75] Even if I had access to sub-monthly coal prices, coal's storability allows plants to arbitrage short-run fuel price fluctuations, and the relevant coal price for this application is plant j's opportunity cost of coal purchases. While other studies have characterized this opportunity cost based on spot market purchases only (e.g. Cicala (2017); Chu, Holladay, and LaRiviere (forthcoming)), I average P_{jm} across both contract and spot transactions. Many plants purchase coal exclusively on long-term contracts, and restricting P_{jm} to spot shipments would necessitate estimating unobserved spot prices for many plant-months. Given that my analysis hinges on *plant-specific* delivered coal prices, I choose to pool all observed coal prices and apply a standard definition for P_{jm} for all coal plants. I conduct sensitivity analysis using alternative coal price variables in Appendix A.4.2.

[76] Here, I aggregate generation across "market regions" (i.e. ISOs or NERC regions) rather than PCAs, because there is a substantial amount of trading between PCAs. This implies that in a given hour, the marginal CEMS unit, which predicts unit u's operating decision, is less likely to reside u's PCA than in u's broader market region. (There is comparatively much less trading across market regions, as I define them.) However, I follow Linn and Muehlenbachs (forthcoming) in defining average marginal costs at the PCA-level, because within-PCA comparisons more accurately exploit cross-sectional differences in natural gas prices that occur due to local pipeline constraints.

precede CEMS units in the dispatch order, meaning that the marginal operating unit will almost always be a CEMS unit when $G_{uh} > 0$.[77] Generation bins b allow me to flexibly estimate unit u's generation, both un-interacted ($\hat{\alpha}_{ub}$, following Davis and Hausman (2016)) and interacted with the cost ratio ($\hat{\gamma}_{ub}$). Because electricity demand is nearly perfectly inelastic, Equation (A.29) is unlikely to suffer from simultaneity bias between CF_{uh} and G_{uh}.

The matrix $\mathbf{G}_{uh}$ includes several time-varying factors that affect unit u's probability of operating conditional on G_{uh} and CR_{ud}. $\mathbf{G}_{uh}$ includes the daily sum, daily maximum, daily minimum, and daily standard deviation of G_{uh} to control for within-day dynamic operating constraints, because coal-fired boilers cannot instantaneously turn on or off (Cullen and Mansur (2017)). I also control for the daily maximum temperature at each power plant, because outdoor temperature directly impacts the thermal efficiency of coal boilers. $\mathbf{G}_{uh}$ includes hour-of-day fixed effects, to control for diurnal operating patterns; quarter-of-year fixed effects, to control for seasonality in electricity demand, relative fuel prices, and plant maintenance schedules; and year fixed effects, to capture long-run changes in unit u's operations. Finally, $\mathbf{G}_{uh}$ includes the interaction of year fixed effects with the daily sum of G_{uh}, to control for changes in electric generating capacity and unit u's position in the dispatch order. Because each time series regression includes multiple years of data at the hourly level, I am able to cluster standard errors by month, which accommodates arbitrary within-unit-month serial correlation.[78]

Step 3: After estimating a model of unit-specific operations as a function of both aggregate CEMS generation G_{uh} and the coal-to-gas marginal cost ratio CR_{ud}, I can predict the distribution of counterfactual coal prices at which unit u would have been exactly marginal. Let $\tilde{P}_{uh}$ denote the coal price at which unit u would have had a 50 percent probability of operating at full capacity in hour h, or $\widehat{CF}_{uh} = 0.5$. Then, rearranging Equation (A.29) post-estimation:

$$HR_{um} \cdot \left(\tilde{P}_{uh} + MC^{env}_{um}\right) = \frac{\left(0.5 - \hat{\alpha}_{ub,h} - \hat{\xi}_u \mathbf{G}_{uh}\right) \cdot MC^{gas}_{ud}}{\hat{\gamma}_{ub,h} + \hat{\delta}_u} \tag{A.30}$$

[77]Nuclear, wind, and solar generation are (virtually) always inframarginal in hours when a CEMS unit is also operating (i.e. $G_{uh} > 0$). Hydroelectric plants are subject to complex dynamic operating constraints, and in certain regions hydro has the potential to be marginal in electricity dispatch. However, these same regions have virtually no coal generation: from 2002–2015, six states (Washington, Oregon, California, New York, Montana, and Idaho) contributed 70 percent of U.S. hydro generation and only 2.5 percent of coal generation.

[78]Month is also the unit of variation in average coal price data, as I assign a single average coal price across all hours in each month. Clustering by plant-month (to accommodate correlated errors between units u within a single plant j) produces virtually identical standard errors, but is far more computationally intensive.

$$\text{(A.31)} \qquad \Rightarrow \quad \tilde{P}_{uh} = \left[\frac{\left(0.5 - \hat{\alpha}_{ub,h} - \hat{\xi}_u \mathbf{G}_{uh}\right) \cdot MC^{gas}_{ud}}{\hat{\gamma}_{ub,h} + \hat{\zeta}_u} \right] \Big/ HR_{um} - MC^{env}_{um}$$

Here $\hat{\alpha}_{ub,h}$ and $\hat{\gamma}_{ub,h}$ denote the binned coefficients associated with each hourly realization of G_{uh}. I assign a standard error to each predicted value of $\tilde{P}_{uh}$ by applying the delta method to the variance-covariance matrix from estimating Equation (A.29).

Step 4: Armed with a distribution of estimated $\tilde{P}_{uh}$'s for each coal unit, I construct a mapping between coal price and plant j's monthly coal demand. Summing across all hours of month m, and across each of plant j's constituent coal units, I define plant j's monthly coal demand function F_{jm} as:

$$\text{(A.32)} \qquad F_{jm}(P) = \sum_{u \in j} \sum_{h \in m} \mathbf{1}\left[P < \tilde{P}_{uh}\right] \cdot \bar{Q}^{coal}_{um}$$

where $\bar{Q}^{coal}_{um}$ is unit u's hourly coal consumption when operating at maximum capacity in month m. This simply assumes that for a given coal price P, plant j will demand the amount of coal required to operate each of its units at full capacity, in inframarginal hours only. Figure 1.5.5 in the main text plots two $F_{jm}(\,\cdot\,)$ curves, for a representative plant j, for two months m.

Step 5: I apply a kernel mean-smoothing algorithm to each $F_{jm}(\,\cdot\,)$ function, and define its smoothed inverse $F^{-1}_{jm}(\,\cdot\,)$. I then calculate local approximations of the first and second derivatives of $F^{-1}_{jm}(\,\cdot\,)$, which I denote as $\Delta F^{-1}_{jm}(\,\cdot\,)$ and $\Delta^2 F^{-1}_{jm}(\,\cdot\,)$, respectively. This lets me estimate empirical analogues of the three components of Equation (1.3): $\frac{\partial P_{oj}}{\partial Z_j}$, $\frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_{oj}$, and $E_{D_{oj}}$.

As each of these terms is a *partial* derivative, I estimate their empirical analogues based on *realized* coal prices for each plant-month. For each month $\ddot{m}$, the value $F_{j\ddot{m}}(P_{j\ddot{m}})$ is the quantity that corresponds with month $\ddot{m}$'s observed coal price $P_{j\ddot{m}}$, given that month's estimated demand function. Plugging this quantity into the functions $F^{-1}_{jm}(\,\cdot\,)$, $\Delta F^{-1}_{jm}(\,\cdot\,)$, and $\Delta^2 F^{-1}_{jm}(\,\cdot\,)$, for *all* months m, I construct the dependent variables for three plant-specific OLS regressions:

$$\text{(A.33)} \qquad \left\{F^{-1}_{jm}(F_{j\ddot{m}}(P_{j\ddot{m}}))\right\}_{j\ddot{m}m} = \lambda_{0j} Z_{jm} + \delta_{j\ddot{m}m} + \nu_{j\ddot{m}m}$$

$$\text{(A.34)} \qquad \left\{\Delta F^{-1}_{jm}(F_{j\ddot{m}}(P_{j\ddot{m}})) \times F_{j\ddot{m}}(P_{j\ddot{m}})\right\}_{j\ddot{m}m} = \lambda_{1j} Z_{jm} + \delta_{j\ddot{m}m} + \nu_{j\ddot{m}m}$$

(A.35)
$$\left\{\Delta^2 F^{-1}_{j\ddot{m}}(F_{j\ddot{m}}(P_{j\ddot{m}})) \div \Delta F^{-1}_{j\ddot{m}}(F_{j\ddot{m}}(P_{j\ddot{m}})) \times F_{j\ddot{m}}(P_{j\ddot{m}})\right\}_{j\ddot{m}} = \lambda_{2j} + \nu_{j\ddot{m}}$$

Z_{jm} is plant j's PCA-weighted natural gas price, averaged for month m, and $\delta_{j\ddot{m}m}$ are two sets of month fixed effects (for months $\ddot{m}$ and m).[79] The coefficient λ_{0j} recovers how changes in Z_{jm} affect the *level* of plant j's inverse demand; λ_{1j} recovers how changes in Z_{jm} affect the *slope* of plant j's inverse demand; and λ_{2j} summarizes plant j's average elasticity of the slope of inverse demand. These map directly to the three terms in Equation (1.3):

$$\text{(A.36)} \qquad \hat{\lambda}_{0j} \sim \frac{\partial P_{oj}}{\partial Z_j} \qquad \hat{\lambda}_{1j} \sim \frac{\partial^2 P_{oj}}{\partial Q_{oj} \partial Z_j} Q_j \qquad \hat{\lambda}_{2j} \frac{\theta_{oj}}{N_{oj}} \sim E_{D_{oj}}$$

The strength of this demand estimation strategy is that it avoids making any assumptions on coal plants' objective functions. The parameters $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, and $\hat{\lambda}_{2j}$ are the outcomes of a semi-parametric linear regression, and I make no functional form assumptions on coal demand to estimate $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, and $\hat{\lambda}_{2j}$.[80] A weakness is my assumption that counterfactual coal prices ($\tilde{P}_{uh}$) hold the rest of the electricity market constant, including prices faced by other coal plants. In reality, plant-specific markups make up only a small portion of delivered coal prices, and large changes to plant j's coal price (e.g. due to a global coal price shock, or a regional shortage in diesel) likely affect many plants simultaneously. This means that my demand estimates will likely only be informative for small changes in plant-specific coal prices. If rail carriers *jointly* reoptimize coal markups across multiple plants (i.e. plant j's markups move in the same direction as the markups of rival coal plants), then my estimated demand functions $F_{jm}(\cdot)$ may be too large (small) at low (high) coal prices. I also assume that coal plants operate their constituent units independently, that coal demand is either zero or at maximum capacity for each hour (yet summing over all hours in each month smooths this discreteness),[81] and that Equation (A.29) is not misspecified.

A.4.2 Results and Sensitivities

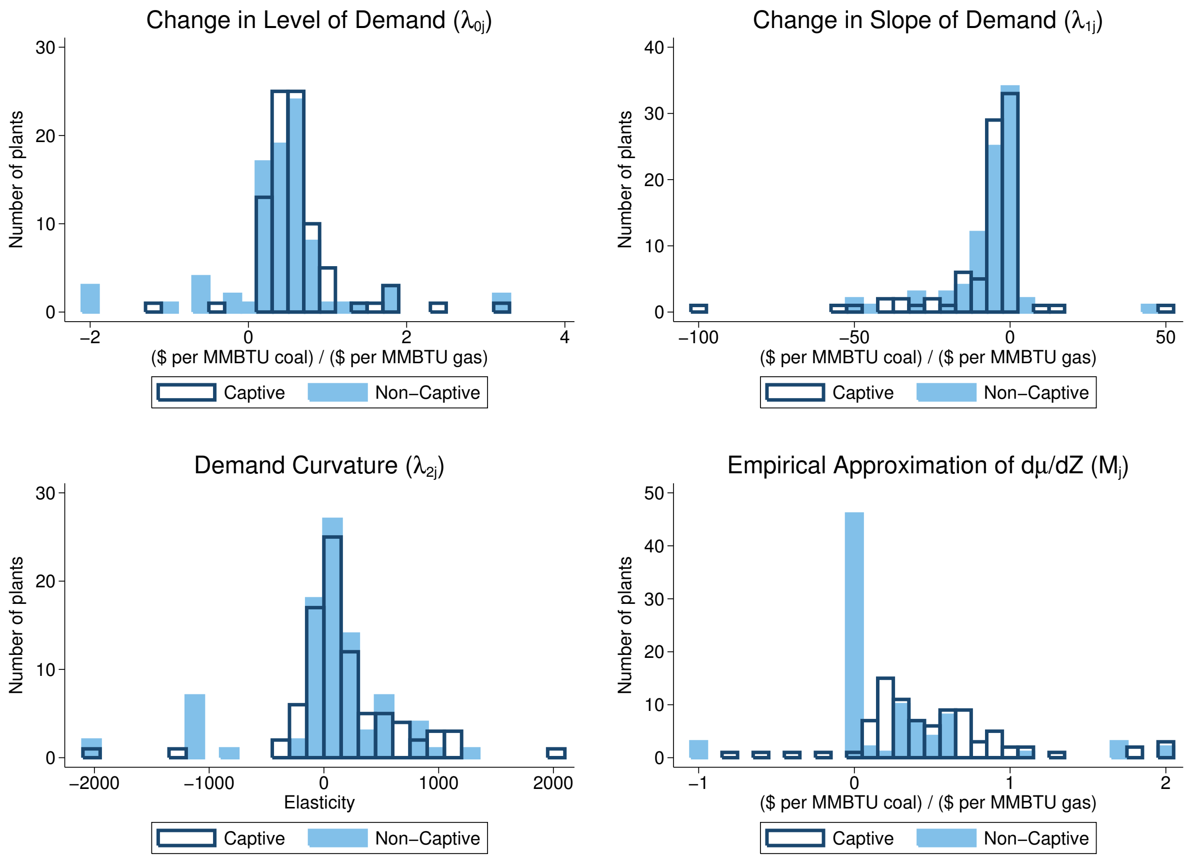
Figure 1.5.6 in the main text reports demand estimation results, after adjusting the sample of plants by $k = 3$ nearest-neighbor weights. I present the same four histograms weighted for $k = 1$ nearest neighbors (in Figure A.4.13), and unweighted (in Figure A.4.14). These results are broadly consistent, demonstrating that sample weights are not meaningfully influencing the the distributions of $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$, $\hat{\lambda}_{2j}$, or M_j. Figure A.4.15 maps coal plants

[79]Month m fixed effects control for month-on-month changes to plants j's coal demand, as well as month-specific estimation error in each $F_{jm}(\cdot)$ function. Month $\ddot{m}$ fixed effects control for errors in calibrating $F_{j\ddot{m}}(P_{j\ddot{m}})$ to plant j's actual coal consumption in month $\ddot{m}$.

[80]Binned linear predictions of power plant operations have become the state-of-the-art in estimating electricity market dispatch (Davis and Hausman (2016); Cicala (2017)). This is because (i) regulated electric utilities do not necessarily operate as profit-maximizing agents, (ii) electric generating units face complex operational constraints, and (iii) a combination of transmission constraints and real-time balancing requirements dictate when/where electricity must be generated to meet demand.

[81]Coal plants often operate at capacity factors between 0.5 and 1, and the median operating capacity factor is 0.91. While I effectively assume capacity factors of 0 or 1 at counterfactual coal prices, summing across all hours in each month smooths this discreteness in coal demand.

Figure A.4.13: Coal Demand Estimation Results ($k = 1$ Nearest Neighbors)

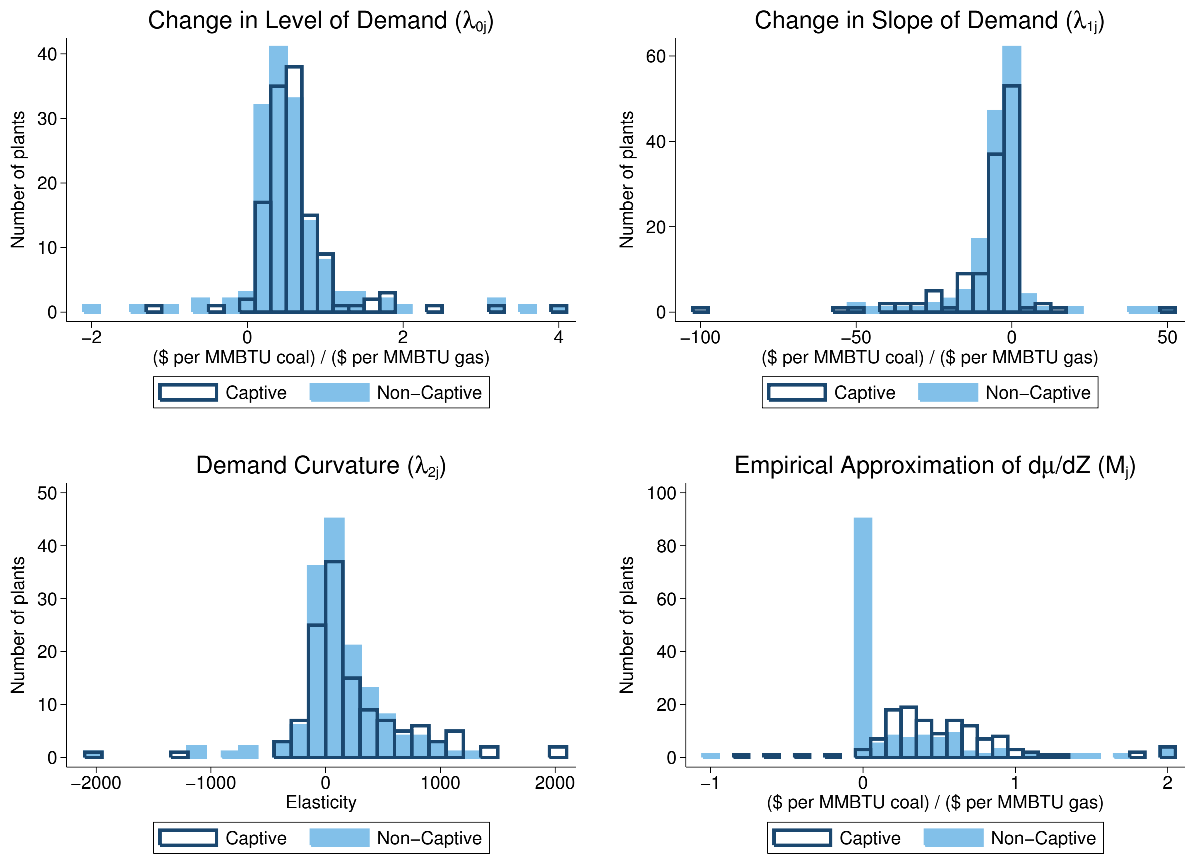


Notes: These histograms are identical to those in Figure 1.5.6 from the main text, except that I weight plants in my matched sample with $k = 1$ nearest neighbors.

color-coded by M_j, revealing that high/low M_j plants are not concentrated in a single region.

I conduct several sensitivity analyses on my demand estimation algorithm, which I report in Table A.4.2 and Figure A.4.16. For each sensitivity, I iterate Steps 1–3 above and store values of $\tilde{P}_{uh}$ for each coal unit u. Next, I calculate correlation coefficients between each unit's alternative $\tilde{P}_{uh}$'s and the same unit's $\tilde{P}_{uh}$'s from my preferred specification. Then, I construct a distribution of each unit's correlation coefficient, and report percentiles of this distribution in Table A.4.2 and histogram of correlation coefficients in Figure A.4.16. In calculating these correlation coefficients, I bottom-code at $\tilde{P}_{uh} = 0$, because Step 4 above uses a lower bound of $P = 0$ and effectively treats negative values of $\tilde{P}_{uh}$ as

Figure A.4.14: Coal Demand Estimation Results (Unweighted)



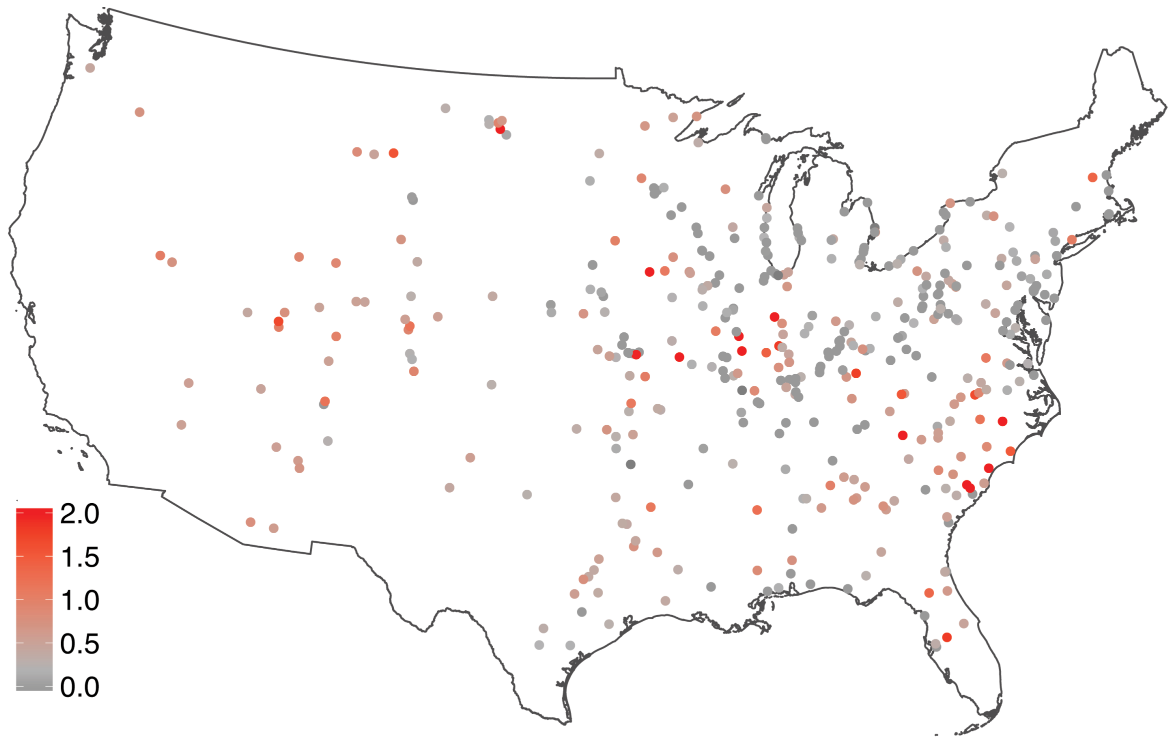
Notes: These histograms are identical to those in Figure 1.5.6 from the main text, except that I use the full sample of utility-owned coal plants, without applying weighting restrictions.

if they were 0. I also top-code at $\tilde{P}_{uh} = 35$, which is greater than the highest coal price ever observed in my data.[82]

First, I test for robustness to alternative definitions of unit u's marginal cost in Equation (A.26). My preferred specification uses each plant's average monthly coal price P_{jm}; however, this average price may not accurately characterize the plant's opportunity cost of coal. For example, suppose a plant purchases 75 percent of its coal on a (relatively expensive) long-term contract, and 25 percent of its coal on the (cheaper) spot market. For this plant, the relevant coal price influencing marginal operating decisions may be the cheaper spot price, which would be less than its average monthly price P_{jm}. I replace P_{jm} with each plant's *minimum* monthly delivered coal price, and report results in the first row of Table A.4.2. This reveals that my demand estimates are not sensitive to the

[82] $\tilde{P}_{uh}$ is often extremely small or extremely large. Extreme negative values of $\tilde{P}_{uh}$ imply that even at a coal price of $P = 0$, unit u would not generate in hour h. Extreme positive values of $\tilde{P}_{uh}$ imply that no feasible coal price would be sufficient to incentivize unit u to generate in hour h.

Figure A.4.15: Coal Plants by M_j



Notes: This map plots 430 coal plants, color-coded by their value of M_j (calculated from Equation 1.8 based on my demand parameter estimates). I top-code (bottom-code) the color scale at 2.0 (0.0) for ease of presentation.

average vs. minimum price distinction: for the median coal unit, $\tilde{P}_{uh}$'s estimated using minimum prices have a correlation of 0.99 with $\tilde{P}_{uh}$'s from my preferred specification. For a coal unit in the 5th percentile, using the minimum price does yields slightly different values of $\tilde{P}_{uh}$.

Storage represents another way in which P_{jm} might mischaracterize plants' opportunity costs of coal. If plants have abundant coal stockpiles, then this true opportunity cost may in fact be close to zero. However, plants understand that coal price changes may persist, and they also value a buffer stock of coal to hedge against potential supply disruptions (Jha (2017)). As I lack detailed data on coal inventories to test this hypothesis, I test for dynamic coal price effects using one-month-lagged P_{jm}, reported in second row of Table A.4.2. My demand estimates are slightly more sensitive to lagged coal prices than to minimum coal prices, yet over half of coal units have $\tilde{P}_{uh}$ correlations of at least 0.90.[83]

My main specification includes marginal environmental costs (for which I have data) and excludes non-fuel variable costs (for which I do not). The third row of Table A.4.2 reports results after I remove MC^{env} from Equations (A.26)–(A.27); the fourth row reports

[83]This may simply reflect autocorrelation in coal prices across months.

Table A.4.2: Demand Estimation – Specification Sensitivities

	Correlation of $\tilde{P}_{uh}$ with Preferred Specification				
	5^{th} pctile	25^{th} pctile	50^{th} pctile	75^{th} pctile	95^{th} pctile
Minimum monthly coal prices	0.24	0.79	0.99	1.00	1.00
1-month lagged coal prices	0.11	0.62	0.90	0.99	1.00
No environmental costs	0.03	0.59	0.90	0.99	1.00
Adding non-fuel variable costs	0.47	0.90	1.00	1.00	1.00
Twice as many bins b	0.46	0.73	0.87	0.97	1.00
Half as many bins b	0.46	0.76	0.90	0.98	1.00
Month-of-year fixed effects	0.16	0.64	0.87	0.97	0.99
Removing year $\times$ load controls	0.06	0.43	0.70	0.90	0.98

Notes: This table reports sensitivity analysis on my demand estimation algorithm. Each row conducts a separate sensitivity, for which I iterate Steps 1–3 of the algorithm from Section A.4.1. I calculate unit-wise correlations between these new values of $\tilde{P}_{uh}$ and the values from my preferred specification, and report percentiles of this distribution of correlations. For example the fourth column reports the 75th percentile of unit-specific correlation coefficients between preferred $\tilde{P}_{uh}$'s and alternative $\tilde{P}_{uh}$'s. If this value is 0.97 for a given sensitivity, then 25 percent of coal units have $\tilde{P}_{uh}$ correlations greater than 0.97. Figure A.4.16 reports histograms of these unit-specific correlations, with each panel corresponding to a row of this table. I describe each sensitivity in detail in the surrounding text.

results after adding technology-specific defaults for non-fuel variable costs to Equations (A.26)–(A.27), as described in Appendix A.2.2.3. These two sensitivities reveal, not surprisingly, that non-fuel costs are second-order relative to fuel costs and heat rates. Removing environmental costs does affect $\tilde{P}_{uh}$ for a small share of coal units, and the 0.03 $\tilde{P}_{uh}$ correlation for the 5th percentile suggests that MC^{env} is closer to first-order for 1 in 20 units.[84] Adding (assumed) non-fuel variable costs (e.g,, labor, maintenance) has virtually no effect on $\tilde{P}_{uh}$'s for the vast majority of coal units.

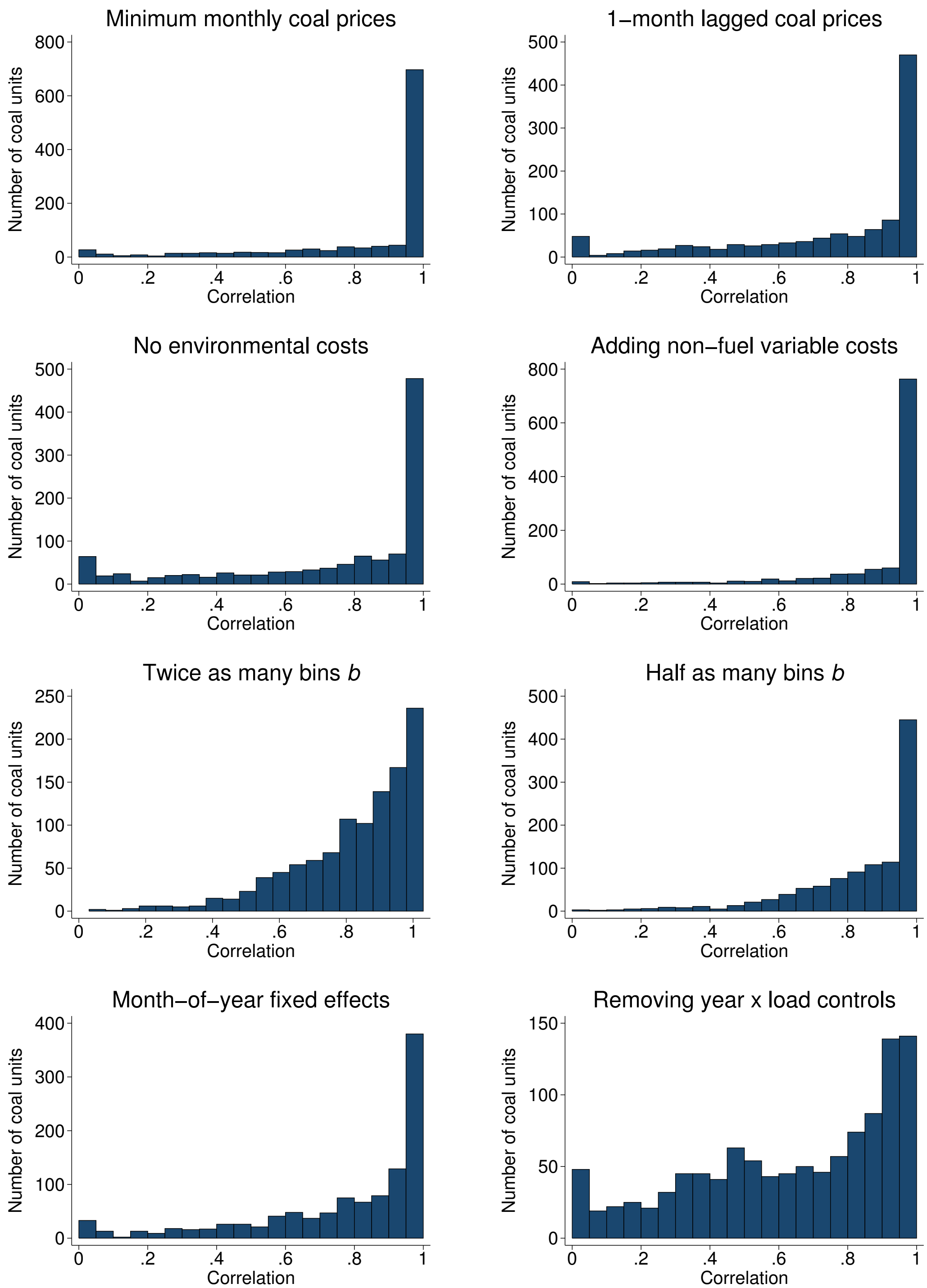
Next, I test for sensitivity to the number of generation bins b in Equations (1.6) and (A.29). I use a preferred bin size of 1000 MWh for the six largest electricity market regions, in decreasing order of size (average CEMS hourly generation): SERC (54,072 MW), PJM (53,869 MW), MISO (51,930 MW), SPP (23,832 MW), ERCOT (23,236 MW), FRCC (17,148 MW). For the five smallest market regions, I reduce the bin sizes such that at least 90 percent of hours in each year have at least 10 bins: NWPP (14,787 MW), SRSG (11,785 MW), CAISO (9,520 MW), NEISO (6,086 MW), NYISO (5,881 MW). I conduct

[84]This implies that measurement error in MC^{env} is unlikely to be seriously influencing my demand estimation results.

two sensitivities on bin size, where I double and halve the number of bins, respectively. The resulting $\tilde{P}_{uh}$ distributions are very highly correlated across the majority of coal units, for both sets of alternative bin sizes.

Finally, I conduct two sensitivities on the fixed effects that enter Equations (1.6) and (A.29) through $\mathbf{G}_{uh}$. My preferred specification controls for seasonal variation using quarter-of-year fixed effects. I prefer quarter fixed effects to month fixed effects, as the latter remove much of the variation in coal prices. However, $\tilde{P}_{uh}$'s estimated with month fixed effects are still quite highly correlated with my preferred $\tilde{P}_{uh}$ estimates, as reported in the seventh row of Table A.4.2. $\mathbf{G}_{uh}$ also includes both year fixed effects (to control for changes in each unit's average operations over time) and the interaction of year fixed effects with total daily generation (to control for changes over time in the relationship between each unit's operations and aggregate market generation). I include year-by-load interactions because electricity markets are not static throughout my sample period, and the 10 GW generation bin may imply dramatically different electricity prices in 2005 vs. 2015—and hence, dramatically difference probabilities that unit u decides to generate. The eighth row of Table A.4.2 shows that removing these year-by-load interactions substantially changes $\tilde{P}_{uh}$'s for a large number of coal units.

Figure A.4.16: Demand Sensitivities – Histograms of Correlations



Notes: Each panel displays a histogram of unit-specific correlations for each sensitivity listed in Table A.4.2. Each histogram contains 1,097 coal units and reports the correlation between a unit's preferred $\tilde{P}_{uh}$'s and its $\tilde{P}_{uh}$'s for each sensitivity. Table A.4.2 reports percentiles of these distributions of correlations. For the few cases were these correlations are negative, I bottom-code at 0 to present consistent horizontal axes.

A.5 Sensitivity Analysis

A.5.1 Nearest-Neighbor Matching

I match captive plants to the k non-captive plants with the closest geographic proximity, and I enforce a maximum distance of 200 miles between matched plants. I force exact matches on the preferred (modal) coal rank that each plant consumed between 2002–2006, which ensures that matched plants do not purchase coal from predominantly opposite sides of the country (i.e. bituminous coal from the East vs. sub-bituminous coal from the West). I also exclude the few plants with covariates that do not overlap with the opposite group.

Formally, let $D_j = 1$ if plant j is captive and $D_j = 0$ if plant j is non-captive. For each number of matches k, I assign nearest-neighbor weights $w_j(k)$ for non-captive plants by summing the inverse of the number of matches $n_\ell(k)$ across all captive plants ℓ:

$$w_j(k) = \begin{cases} 1 & \text{if } \ D_j = 1\,,\ n_j(k) \in \{1, \ldots, k\} \\ 0 & \text{if } \ D_j = 1\,,\ n_j(k) = 0 \\ \displaystyle\sum_{\substack{\ell \,|\, D_\ell = 1, \\ j \in n_\ell(k)}} \frac{1}{n_\ell(k)} & \text{if } \ D_j = 0\,,\ j \in n_\ell(k) \text{ for some } \ell \,|\, D_\ell = 1 \\ 0 & \text{if } \ D_j = 0\,,\ j \notin n_\ell(k) \text{ for all } \ell \,|\, D_\ell = 1 \end{cases} \tag{A.37}$$

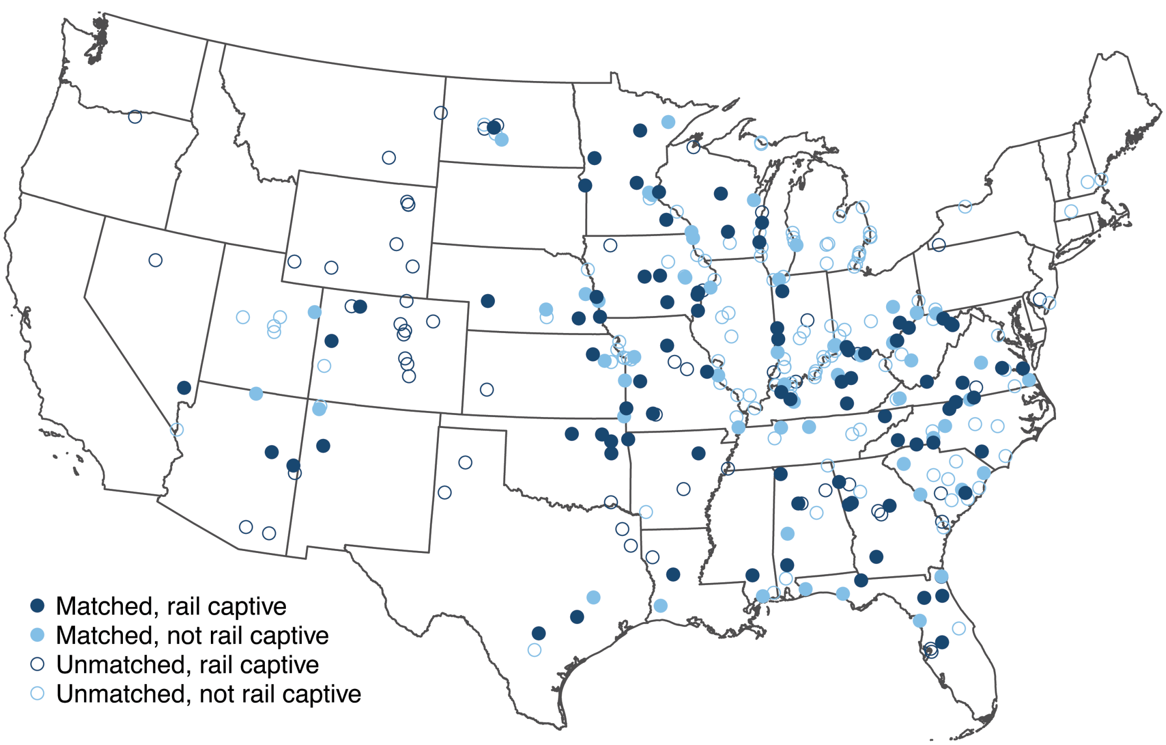
That is, all matched captive plants receive weights of $w_j(k) = 1$, and all unmatched plants receive weights of $w_j(k) = 0$. Matched non-captive plants receive weights that adjust the share of non-captive plants to equal the number of matched captive plants. For example, suppose a non-captive plant is one of 3 matches for captive plant A and one of 2 matches for captive plant B, for $k = 3$. This plant would receive a weight of $w_j(3) = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}$.

In the main text, I map nearest-neighbor matches and report summary statistics for $k = 3$ matches (Figure 1.5.4 and Table 1.5.1). Figure A.5.17 is an analogous map for $k = 1$, and I present the summary statistics for the $k = 1$ matched sample in Table A.5.3. This reveals that geographic overlap and covariate balance are not sensitive to the my choice of the number of nearest neighbors. Table A.5.4 reports sensitivity on the distance cutoff of 200, restricting all match to be within 100 miles. This yields nearly identical results, while reducing the sample size from 86 to 71 matched plants. Figure A.5.18 demonstrates that even with a 200-mile distance cutoff, the majority of matches are within 100 miles in geographic proximity.

Table A.5.5 demonstrates that covariate balance does not depend on my definition of "route unconnectedness", as defined in Appendix A.3. My preferred threshold for route unconnectedness is 300 miles, or the increase in rail shipping distance after removing a dominant rail carrier that renders a coal plant "captive". Table A.5.5 redefines captiveness based on either a 150- or 600-mile threshold for route unconnectedness, and this has little effect on the composition of my matched sample. Table A.5.6 likewise shows that covariate balance does not depend on my definition of "node unconnectedness". My preferred threshold is 6.6 miles, which is the 95th percentile of the distribution of plants' distance to their closest rail node. Table A.5.6 redefines captiveness based on either a 5- or 10-mile threshold for node unconnectedness, with little effect on the average covariates across captive vs. non-captive groups.

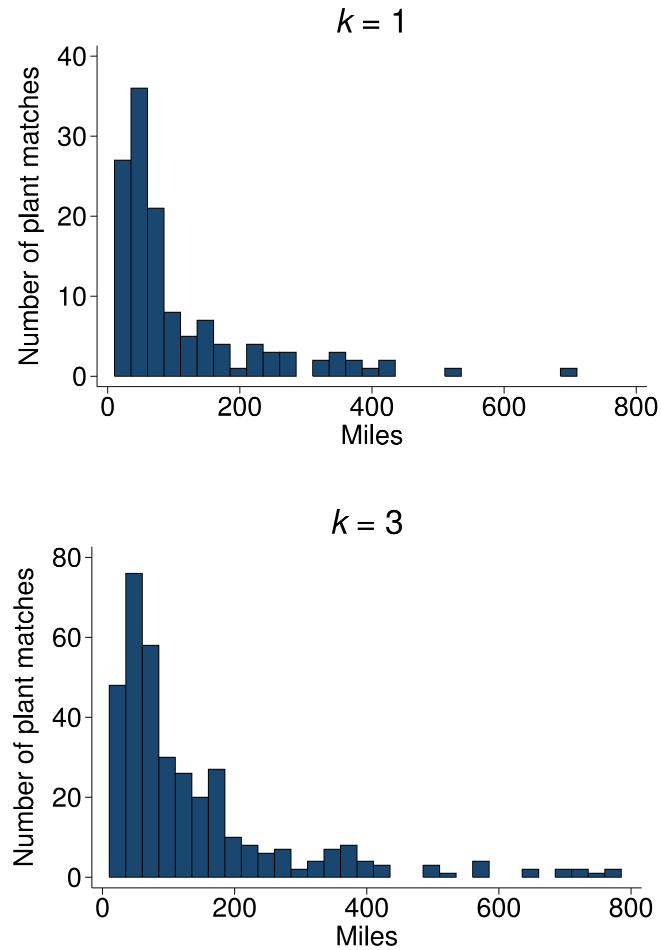
Finally, Tables A.5.7–A.5.8 reported summary statistics for a "balanced" panel of coal plants, or the subset of coal plants receiving at least one coal delivery in each year between 2002 and 2015. These matched samples correspond to the estimation samples in Tables 1.6.4–1.7.7 in the main text, and there are no detectable differences between captive vs. non-captive plants after nearest-neighbor matching this balanced sample. Table 1.6.6 also illustrates why I slightly prefer estimates with $k = 3$ nearest neighbor matches: with $k = 1$ match per captive plant, 60 captive plants have just 36 unique non-captive matches.

Figure A.5.17: Nearest Neighbor Matching ($k = 1$)



Notes: This map is identical to Figure 1.5.4 from the main text, except that captive plants have only $k = 1$ nearest neighbors. Matches have a maximum distance of 200 miles; exact matches on coal rank; and omit non-utility and non-rail plants.

Figure A.5.18: Distance to Nearest k Neighbors



Notes: Each panel displays a histogram of the distances to each captive plant's k nearest neighbors, with exact matches on coal rank, and removing non-utility and non-rail plants. My analysis restricts matched distances to be within 200 miles.

Table A.5.3: Summary Statistics, $k = 1$ Nearest Neighbors

	Full sample			Matched sample ($k = 1$)		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	908.19	865.73	42.46	900.89	964.85	−63.96
	(780.72)	(742.77)	[0.57]	(69.32)	(147.14)	[0.69]
Coal-fired capacity (MW)	806.13	760.84	45.29	815.64	858.49	−42.85
	(738.72)	(703.91)	[0.52]	(66.58)	(140.80)	[0.78]
Number of coal units	2.36	2.62	−0.26	2.59	2.65	−0.06
	(1.32)	(1.64)	[0.08]*	(0.15)	(0.23)	[0.82]
Coal unit vintage (year)	1968.85	1962.88	5.97	1966.25	1961.22	5.03
	(13.90)	(13.34)	[0.00]***	(1.39)	(1.79)	[0.03]**
Annual capacity factor	0.63	0.60	0.03	0.63	0.63	−0.00
	(0.17)	(0.17)	[0.04]**	(0.01)	(0.02)	[0.97]
Heat rate (MMBTU/MWh)	11.09	11.06	0.03	10.97	10.90	0.06
	(1.40)	(1.52)	[0.86]	(0.14)	(0.22)	[0.81]
Scrubber installed	0.36	0.29	0.07	0.28	0.28	0.00
	(0.48)	(0.45)	[0.12]	(0.05)	(0.07)	[1.00]
Market participant	0.49	0.71	−0.22	0.46	0.51	−0.05
	(0.50)	(0.46)	[0.00]***	(0.05)	(0.08)	[0.64]

	Full sample			Matched sample ($k = 1$)		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	48.82	44.00	4.82	47.44	47.14	0.30
	(47.90)	(43.70)	[0.29]	(4.10)	(8.00)	[0.97]
Sulfur content (lbs/MMBTU)	0.87	1.02	−0.15	0.79	0.86	−0.07
	(0.61)	(0.79)	[0.03]**	(0.06)	(0.10)	[0.55]
Ash content (lbs/MMBTU)	8.46	8.96	−0.50	8.08	8.19	−0.11
	(4.21)	(8.24)	[0.46]	(0.37)	(0.68)	[0.89]
Share spot market	0.19	0.19	−0.00	0.18	0.14	0.03
	(0.29)	(0.25)	[0.87]	(0.03)	(0.03)	[0.40]
Share contracts expiring ≤ 2 years	0.22	0.24	−0.01	0.19	0.19	−0.00
	(0.25)	(0.26)	[0.61]	(0.02)	(0.03)	[0.99]
Share sub-bituminous	0.41	0.31	0.10	0.43	0.41	0.01
	(0.47)	(0.42)	[0.03]**	(0.05)	(0.08)	[0.88]
Average rail distance (miles)	554.91	620.34	−65.43	565.67	595.15	−29.48
	(385.90)	(417.90)	[0.12]	(40.37)	(60.55)	[0.69]

	Full sample			Matched sample ($k = 1$)		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	94	149	243	49	34	83
Preferred coal rank: sub-bituminous	77	76	153	36	19	55
Non-rail plants	17	14	31	0	0	0
Utility plants	148	176	324	87	54	141
Total plants	190	240	430	87	54	141

Notes: This table is identical to Table 1.5.1 from the main text, except it uses $k = 1$ nearest-neighbor matches. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.4: Summary Statistics, 100-Mile Distance Cutoff

	Matched sample ($k = 1$)			Matched sample ($k = 3$)		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	882.27	932.56	−50.28	882.27	939.63	−57.36
	(77.57)	(157.71)	[0.77]	(77.57)	(110.79)	[0.67]
Coal-fired capacity (MW)	790.25	835.23	−44.98	790.25	827.71	−37.46
	(75.12)	(149.19)	[0.79]	(75.12)	(104.20)	[0.77]
Number of coal units	2.55	2.67	−0.12	2.55	2.58	−0.03
	(0.16)	(0.25)	[0.69]	(0.16)	(0.16)	[0.90]
Coal unit vintage (year)	1966.21	1960.06	6.15	1966.21	1961.92	4.29
	(1.56)	(1.80)	[0.01]**	(1.56)	(1.55)	[0.05]*
Annual capacity factor	0.62	0.62	0.00	0.62	0.62	0.00
	(0.02)	(0.02)	[0.88]	(0.02)	(0.01)	[0.93]
Heat rate (MMBTU/MWh)	11.02	10.98	0.04	11.02	10.78	0.25
	(0.16)	(0.25)	[0.89]	(0.16)	(0.18)	[0.30]
Scrubber installed	0.27	0.23	0.04	0.27	0.23	0.04
	(0.05)	(0.06)	[0.60]	(0.05)	(0.05)	[0.56]
Market participant	0.51	0.56	−0.06	0.51	0.57	−0.06
	(0.06)	(0.09)	[0.59]	(0.06)	(0.07)	[0.49]
	Matched sample ($k = 1$)			Matched sample ($k = 3$)		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	45.66	44.97	0.69	45.66	43.76	1.91
	(4.59)	(8.02)	[0.94]	(4.59)	(5.46)	[0.79]
Sulfur content (lbs/MMBTU)	0.85	0.90	−0.05	0.85	0.91	−0.06
	(0.07)	(0.11)	[0.70]	(0.07)	(0.08)	[0.55]
Ash content (lbs/MMBTU)	8.12	7.90	0.23	8.12	7.82	0.30
	(0.42)	(0.42)	[0.70]	(0.42)	(0.30)	[0.56]
Share spot market	0.17	0.15	0.02	0.17	0.16	0.02
	(0.03)	(0.03)	[0.64]	(0.03)	(0.02)	[0.66]
Share contracts expiring $\leq$ 2 years	0.20	0.20	−0.00	0.20	0.20	−0.01
	(0.02)	(0.03)	[0.90]	(0.02)	(0.02)	[0.85]
Share sub-bituminous	0.40	0.38	0.02	0.40	0.33	0.07
	(0.06)	(0.08)	[0.85]	(0.06)	(0.06)	[0.41]
Average rail distance (miles)	561.54	595.26	−33.71	561.54	576.60	−15.06
	(45.23)	(58.95)	[0.65]	(45.23)	(49.66)	[0.82]
	Matched sample ($k = 1$)			Matched sample ($k = 3$)		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	42	31	73	42	53	95
Preferred coal rank: sub-bituminous	27	17	44	27	31	58
Non-rail plants	0	0	0	0	0	0
Utility plants	71	49	120	71	85	156
Total plants	71	49	120	71	85	156

Notes: This table is identical to Table 1.5.1 from the main text, except it restricts the maximum match distance to 100 miles. The left three columns use $k = 1$ nearest neighbors, and the right three colums use $k = 3$ nearest neighbors. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.5: Summary Statistics, Alternative Captiveness Thresholds (Routes)

	$\geq$ 150-mile increase			$\geq$ 600-mile increase		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	888.67	927.57	−38.89	900.89	943.67	−42.77
	(68.55)	(91.55)	[0.73]	(69.32)	(93.60)	[0.71]
Coal-fired capacity (MW)	797.96	821.40	−23.44	815.64	824.40	−8.77
	(66.22)	(87.99)	[0.83]	(66.58)	(90.44)	[0.94]
Number of coal units	2.53	2.66	−0.13	2.59	2.60	−0.01
	(0.15)	(0.15)	[0.54]	(0.15)	(0.15)	[0.94]
Coal unit vintage (year)	1965.91	1962.28	3.63	1966.25	1962.27	3.98
	(1.37)	(1.38)	[0.06]*	(1.39)	(1.44)	[0.05]*
Annual capacity factor	0.63	0.63	0.01	0.63	0.63	−0.00
	(0.01)	(0.01)	[0.80]	(0.01)	(0.01)	[0.94]
Heat rate (MMBTU/MWh)	10.89	10.76	0.14	10.97	10.76	0.21
	(0.12)	(0.12)	[0.42]	(0.14)	(0.13)	[0.28]
Scrubber installed	0.27	0.27	−0.00	0.28	0.26	0.01
	(0.05)	(0.06)	[0.95]	(0.05)	(0.06)	[0.84]
Market participant	0.46	0.50	−0.04	0.46	0.50	−0.04
	(0.05)	(0.06)	[0.59]	(0.05)	(0.06)	[0.65]

	$\geq$ 150-mile increase			$\geq$ 600-mile increase		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	46.37	44.62	1.75	47.44	45.74	1.70
	(4.07)	(5.15)	[0.79]	(4.10)	(5.22)	[0.80]
Sulfur content (lbs/MMBTU)	0.81	0.84	−0.04	0.79	0.84	−0.05
	(0.06)	(0.06)	[0.66]	(0.06)	(0.07)	[0.58]
Ash content (lbs/MMBTU)	8.04	7.91	0.13	8.08	7.99	0.09
	(0.36)	(0.41)	[0.81]	(0.37)	(0.42)	[0.87]
Share spot market	0.18	0.18	0.00	0.18	0.16	0.02
	(0.03)	(0.03)	[0.97]	(0.03)	(0.02)	[0.63]
Share contracts expiring $\leq$ 2 years	0.19	0.20	−0.01	0.19	0.19	−0.00
	(0.02)	(0.02)	[0.74]	(0.02)	(0.02)	[0.91]
Share sub-bituminous	0.42	0.39	0.02	0.43	0.40	0.02
	(0.05)	(0.06)	[0.77]	(0.05)	(0.06)	[0.77]
Average rail distance (miles)	563.19	579.39	−16.20	565.67	582.89	−17.22
	(39.79)	(41.21)	[0.78]	(40.37)	(42.69)	[0.77]

	$\geq$ 150-mile increase			$\geq$ 600-mile increase		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	51	64	115	49	59	108
Preferred coal rank: sub-bituminous	36	36	72	36	35	71
Non-rail plants	0	0	0	0	0	0
Utility plants	89	101	190	87	95	182
Total plants	89	101	190	87	95	182

Notes: This table is identical to Table 1.5.1 from the main text, except both sets of columns are nearest-neighbor matched using alternative distnace thresholds to define (as defined in Appendix A.3). The left (right) three columns consider a plant to be captive if removing the dominant rail carrier along along each origin-destination route increases the shortest path by at least 150 (600) miles. My preferred threshold is an increase of 300 miles, which implies roughly a 20 percent increase in median rail shipping costs. Matching criteria: up to k nearest neighbors ($k = 3$), with a maximum distance of 200 miles; exact matches on preferred coal rank; and removing non-utility and non-rail plants. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.6: Summary Statistics, Alternative Captiveness Thresholds (Nodes)

	5-mile threshold			10-mile threshold		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	919.62	932.77	−13.15	844.63	900.34	−55.71
	(66.97)	(93.40)	[0.91]	(81.50)	(90.83)	[0.65]
Coal-fired capacity (MW)	820.42	827.53	−7.11	757.64	825.29	−67.64
	(64.12)	(91.81)	[0.95]	(77.82)	(89.54)	[0.57]
Number of coal units	2.59	2.58	0.01	2.39	2.48	−0.09
	(0.15)	(0.15)	[0.96]	(0.14)	(0.16)	[0.68]
Coal unit vintage (year)	1966.20	1963.26	2.94	1967.12	1963.55	3.57
	(1.36)	(1.42)	[0.14]	(1.62)	(1.50)	[0.11]
Annual capacity factor	0.63	0.64	−0.01	0.63	0.62	0.01
	(0.01)	(0.02)	[0.78]	(0.02)	(0.01)	[0.76]
Heat rate (MMBTU/MWh)	10.88	10.68	0.21	11.10	10.81	0.30
	(0.13)	(0.12)	[0.24]	(0.17)	(0.13)	[0.16]
Scrubber installed	0.27	0.28	−0.01	0.33	0.27	0.06
	(0.05)	(0.06)	[0.89]	(0.06)	(0.06)	[0.47]
Market participant	0.44	0.47	−0.03	0.48	0.49	−0.01
	(0.05)	(0.06)	[0.66]	(0.06)	(0.06)	[0.86]

	5-mile threshold			10-mile threshold		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	46.87	45.44	1.43	45.10	45.72	−0.62
	(4.05)	(5.09)	[0.83]	(4.87)	(5.30)	[0.93]
Sulfur content (lbs/MMBTU)	0.83	0.88	−0.05	0.85	0.85	−0.00
	(0.07)	(0.08)	[0.63]	(0.08)	(0.07)	[0.99]
Ash content (lbs/MMBTU)	8.07	7.88	0.20	8.21	8.12	0.09
	(0.36)	(0.40)	[0.72]	(0.46)	(0.50)	[0.90]
Share spot market	0.19	0.16	0.03	0.14	0.16	−0.02
	(0.03)	(0.02)	[0.43]	(0.03)	(0.03)	[0.53]
Share contracts expiring $\leq$ 2 years	0.19	0.18	0.01	0.18	0.20	−0.03
	(0.02)	(0.02)	[0.60]	(0.02)	(0.02)	[0.36]
Share sub-bituminous	0.41	0.40	0.01	0.44	0.42	0.02
	(0.05)	(0.06)	[0.87]	(0.06)	(0.06)	[0.81]
Average rail distance (miles)	572.89	570.12	2.76	542.55	560.74	−18.19
	(38.85)	(45.58)	[0.96]	(46.34)	(42.66)	[0.77]

	5-mile threshold			10-mile threshold		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	54	57	111	36	59	95
Preferred coal rank: sub-bituminous	37	35	72	29	37	66
Non-rail plants	0	0	0	0	0	0
Utility plants	91	92	183	67	97	164
Total plants	91	92	183	67	97	164

Notes: This table is identical to Table 1.5.1 from the main text, except both sets of columns are nearest-neighbor matched using alternative distnace thresholds to define (as defined in Appendix A.3). The left (right) three columns consider a plant to be captive if all rail terminal nodes within a 5-mile (10-mile) radius are controlled by a single Class I carrier. My preferred threshold is 6.6 miles, which is the 95th percentile of the the distance to each plant's (unrestricted) nearest node. Matching criteria: up to k nearest neighbors ($k = 3$), with a maximum distance of 200 miles; exact matches on preferred coal rank; and removing non-utility and non-rail plants. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.7: Summary Statistics, $k = 3$ Nearest Neighbors, Balanced Panel

	Full sample			Matched sample ($k = 3$)		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	1137.98	1097.62	40.36	1116.89	1183.58	−66.68
	(734.98)	(741.99)	[0.67]	(77.07)	(115.68)	[0.63]
Coal-fired capacity (MW)	1058.79	1008.70	50.09	1028.73	1057.03	−28.29
	(726.28)	(734.03)	[0.59]	(76.02)	(116.50)	[0.84]
Number of coal units	2.45	2.79	−0.33	2.61	2.64	−0.04
	(1.29)	(1.68)	[0.09]*	(0.18)	(0.19)	[0.89]
Coal unit vintage (year)	1971.84	1966.32	5.52	1968.95	1965.62	3.33
	(12.22)	(12.26)	[0.00]***	(1.52)	(1.65)	[0.14]
Annual capacity factor	0.69	0.65	0.04	0.67	0.67	0.00
	(0.12)	(0.12)	[0.01]**	(0.01)	(0.02)	[0.89]
Heat rate (MMBTU/MWh)	10.54	10.57	−0.03	10.50	10.38	0.12
	(0.79)	(0.91)	[0.81]	(0.08)	(0.12)	[0.43]
Scrubber installed	0.41	0.36	0.05	0.31	0.38	−0.07
	(0.49)	(0.48)	[0.45]	(0.06)	(0.08)	[0.48]
Market participant	0.42	0.68	−0.25	0.44	0.50	−0.06
	(0.50)	(0.47)	[0.00]***	(0.06)	(0.08)	[0.56]
	Full sample			Matched sample ($k = 3$)		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	65.65	59.61	6.04	60.97	60.29	0.68
	(46.06)	(45.22)	[0.30]	(4.73)	(6.60)	[0.93]
Sulfur content (lbs/MMBTU)	0.80	1.04	−0.24	0.73	0.95	−0.22
	(0.59)	(0.82)	[0.01]**	(0.07)	(0.09)	[0.06]*
Ash content (lbs/MMBTU)	8.50	8.63	−0.13	7.82	8.14	−0.32
	(4.20)	(4.90)	[0.82]	(0.35)	(0.58)	[0.64]
Share spot market	0.15	0.14	0.01	0.17	0.13	0.04
	(0.25)	(0.16)	[0.72]	(0.03)	(0.02)	[0.32]
Share contracts expiring ≤ 2 years	0.17	0.20	−0.03	0.16	0.15	0.01
	(0.20)	(0.24)	[0.24]	(0.02)	(0.02)	[0.71]
Share sub-bituminous	0.51	0.35	0.16	0.54	0.43	0.11
	(0.48)	(0.42)	[0.01]**	(0.06)	(0.07)	[0.27]
Average rail distance (miles)	579.52	633.65	−54.13	611.75	567.36	44.40
	(413.91)	(421.50)	[0.33]	(49.38)	(55.36)	[0.55]
	Full sample			Matched sample ($k = 3$)		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	45	78	123	28	40	68
Preferred coal rank: sub-bituminous	57	51	108	32	26	58
Non-rail plants	7	8	15	0	0	0
Utility plants	97	107	204	61	66	127
Total plants	113	136	249	61	66	127

Notes: This table is identical to Table 1.5.1 from the main text, except it includes only plants that receive at least 1 coal delivery in each year, from 2002 to 2015. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.8: Summary Statistics, $k = 1$ Nearest Neighbors, Balanced Panel

	Full sample			Matched sample ($k = 1$)		
A. Plant characteristics	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Total plant capacity (MW)	1137.98	1097.62	40.36	1116.89	1342.17	−225.27
	(734.98)	(741.99)	[0.67]	(77.07)	(171.47)	[0.23]
Coal-fired capacity (MW)	1058.79	1008.70	50.09	1028.73	1224.72	−195.98
	(726.28)	(734.03)	[0.59]	(76.02)	(168.41)	[0.29]
Number of coal units	2.45	2.79	−0.33	2.61	2.92	−0.31
	(1.29)	(1.68)	[0.09]*	(0.18)	(0.30)	[0.38]
Coal unit vintage (year)	1971.84	1966.32	5.52	1968.95	1965.43	3.52
	(12.22)	(12.26)	[0.00]***	(1.52)	(2.02)	[0.17]
Annual capacity factor	0.69	0.65	0.04	0.67	0.67	0.00
	(0.12)	(0.12)	[0.01]**	(0.01)	(0.02)	[0.89]
Heat rate (MMBTU/MWh)	10.54	10.57	−0.03	10.50	10.36	0.14
	(0.79)	(0.91)	[0.81]	(0.08)	(0.15)	[0.41]
Scrubber installed	0.41	0.36	0.05	0.31	0.37	−0.06
	(0.49)	(0.48)	[0.45]	(0.06)	(0.09)	[0.58]
Market participant	0.42	0.68	−0.25	0.44	0.52	−0.08
	(0.50)	(0.47)	[0.00]***	(0.06)	(0.09)	[0.50]

	Full sample			Matched sample ($k = 1$)		
B. Coal deliveries	Rail Captive	Not Rail Captive	Difference in means	Rail Captive	Not Rail Captive	Difference in means
Deliveries (million MMBTU/year)	65.65	59.61	6.04	60.97	69.29	−8.32
	(46.06)	(45.22)	[0.30]	(4.73)	(9.48)	[0.43]
Sulfur content (lbs/MMBTU)	0.80	1.04	−0.24	0.73	1.03	−0.30
	(0.59)	(0.82)	[0.01]**	(0.07)	(0.14)	[0.06]*
Ash content (lbs/MMBTU)	8.50	8.63	−0.13	7.82	8.80	−0.98
	(4.20)	(4.90)	[0.82]	(0.35)	(0.98)	[0.35]
Share spot market	0.15	0.14	0.01	0.17	0.11	0.05
	(0.25)	(0.16)	[0.72]	(0.03)	(0.03)	[0.18]
Share contracts expiring ≤ 2 years	0.17	0.20	−0.03	0.16	0.15	0.01
	(0.20)	(0.24)	[0.24]	(0.02)	(0.04)	[0.81]
Share sub-bituminous	0.51	0.35	0.16	0.54	0.37	0.16
	(0.48)	(0.42)	[0.01]**	(0.06)	(0.09)	[0.14]
Average rail distance (miles)	579.52	633.65	−54.13	611.75	558.05	53.71
	(413.91)	(421.50)	[0.33]	(49.38)	(80.51)	[0.57]

	Full sample			Matched sample ($k = 1$)		
C. Number of plants	Rail Captive	Not Rail Captive	Total	Rail Captive	Not Rail Captive	Total
Preferred coal rank: bituminous	45	78	123	28	23	51
Preferred coal rank: sub-bituminous	57	51	108	32	13	45
Non-rail plants	7	8	15	0	0	0
Utility plants	97	107	204	61	36	97
Total plants	113	136	249	61	36	97

Notes: This table is identical to Table 1.5.1 from the main text, except it uses $k = 1$ nearest-neighbor matches, and includes only plants that receive at least 1 coal delivery in each year, from 2002 to 2015. Standard deviations are in parentheses, and p-values [in brackets] are clustered at the plant level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.5.2 Markup Levels

In this section, I conduct sensitivity analysis for my estimates of markup levels, which I report in Section 1.6.1 of the main text. Each sensitivity analysis alters a single element of Equation (1.5). I report results only for a balanced panel of coal plants, using $k = 3$ nearest neighbors; these results are broadly consistent for $k = 1$ nearest neighbors or for an unbalanced sample.[85]

I begin by estimating Equation (1.5) using alternative definitions of D_j, the indicator for rail captiveness. Appendix A.3 describes my method for constructing this variable, which necessitates imposing an arbitrary cutoff of 300 miles for "route unconnectedness", or the distance a plant's shortest route must increase for the route to become "unconnected", after removing the dominant carrier along the route. I classify a plant as captive if it becomes route-unconnected (after removing *each* route's dominant carrier one-by-one) from *all* observed trading partners. Table A.5.9 shows that my markup estimates are robust to halving or doubling this 300-mile route unconnectedness threshold, comparing Columns (2)–(3) to my preferred specification in Column (1). Column (4) weakens my definition of captiveness, such that only the *average* shipment of coal need become route-unconnected. This prevents largely un-utilized routes (i.e. for singleton coal shipments) from influencing a plant's captiveness designation, yet has little effect my results. By contrast, Column (5) strengthens the definition of captiveness to include routes to all *potential* originating coal counties with similar coal attributes to a plant's observed purchases. The prevents a plant from being designated as captive if it *could have* purchased coal from a county from which it does *not* become route-unconnected.[86] This likewise has little effect on my results. Finally, Column (6) defines captiveness based only on "node unconnectedness", which is a more straightforward (though less nuanced) distinction that ignores coal routes. Here, a plant is captive if removing any single Class I rail carrier renders that plant unconnected to any rail node. This captiveness definition yields similar results.

Table A.5.10 conducts an analogous sensitivity analysis for my threshold for "node unconnectedness". My preferred cutoff is 6.6 miles, which is the 95th percentile of the distribution of plants' distance to their nearest terminal rail node; I define a plant as captive is all rail nodes within a 6.6-mile radius are controlled by a single Class I rail carrier. Column (1) of Table A.5.10 strengthens my definition my captiveness by reducing this threshold to 5 miles, while Column (2) weakens my definition of captiveness by increasing this threshold to 10 miles. In both cases, my results are largely unchanged and retain statistical significance. Columns (3)–(4) perform the same sensitivities on the

[85] Recall that my "balanced" sample is not fully balanced, because not every coal plant records a coal purchase in each month that it operates. I "balance" this panel by keep only those plants that report at least 1 coal delivery in each calendar year from 2002 to 2015, to remove coal plants that retired during my sample period.

[86] In other words, one might worry that my preferred definition of captiveness would result in many false positives. If a coal plant only purchases coal from county A (from which it becomes route-unconnected), but it *could have* purchased identical coal from county B (from which it does *not* become route-unconnected), then I would want to classify this plant as non-captive.

Table A.5.9: Markup Levels – Sensitvity to Definition of Captiveness

	Unconnected across all observed routes			Average route	All potential routes	Only node connected-ness
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\text{Captive}]_j$	2.301***	2.163***	2.265***	1.907***	2.374***	2.004***
	(0.655)	(0.641)	(0.641)	(0.616)	(0.666)	(0.619)
k nearest neighbors	3	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Route unc. cutoff (miles)	300	150	600	300	300	
Mean of dep var	36.85	37.37	36.45	36.05	37.05	39.05
Plants	127	129	127	132	127	119
Captive plants	61	61	61	84	60	54
Observations	77,115	78,997	77,420	74,390	77,094	72,398

Notes: This table is identical to Table 1.6.2 from the main text, except that it uses alternative definitions for rail captiveness. Column (1) uses my preferred definition described in Appendix A.3, reported in Column (5) of Table 1.6.2. Columns (2)–(3) halve and double my preferred 300-mile threshold for route unconnectedness. Column (4) weakens the definition of captiveness such that only the plant's *average* route need be unconnected (rather than *all* observed routes). Column (5) strengthens the definition of captiveness to include all *potential* routes with observationally similar coal, even if I do not observe any deliveries along such routes. Column (6) defines captiveness based on node (un)connectedness only. Each column re-constructs nearest-neighbor weights consistent with its respective definition of captiveness. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

threshold for node-unconnectedness, making this the the only determinant of captiveness (i.e., ignoring coal routes as in Column (6) of Table A.5.9). This has little effect on my results.

Table A.5.11 conducts sensitivity analysis on coal commodity controls (i.e. $\mathbf{C}_{ojms}$ in Equation (1.5)). Column (2) forces the coefficient on the county-year average mine-mouth price to be equal to 1, while Column (3) interacts this price with a spot market indicator to allow the mine-mouth sales price to vary for contract vs. spot sales. Neither has a meaningful effect on my results. Because sulfur content is a major driver of dispersion in coal prices, Column (4) adds an interaction between the county-year average mine-mouth price and each shipment's average sulfur content. Column (5) allows a different coefficient on sulfur content for each calendar year, to accommodate for changes to the shadow price of SO_2 emissions due to SO_2 allowance markets. Neither has any effect on my results. Finally, while Equation (1.5) includes coal county fixed effects, it is possible that time-varying factors relating to county-specific coal production have caused me to misspecify $\mathbf{C}_{ojms}$. Column (6) tests for this possibility by adding several time-varying controls for coal production in each county, but my results change very little.[87]

[87] These county-by-year controls include mine age, seam thickness and depth (which influence extraction costs), the share of coal produced from (more expensive) underground mines, the share of mine

Table A.5.10: Markup Levels – Sensitvity to Definition of Captiveness

	Unconnected across all observed routes		Only node connectedness	
	(1)	(2)	(3)	(4)
$\mathbf{1}[\text{Captive}]_j$	1.585**	2.496***	1.560**	1.574**
	(0.718)	(0.700)	(0.653)	(0.705)
k nearest neighbors	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes
Route unc. cutoff (miles)	300	300		
Node unc. cutoff (miles)	5	10	5	10
Mean of dep var	39.78	36.60	40.55	39.52
Plants	129	115	130	87
Captive plants	64	43	60	33
Observations	78,927	68,105	77,140	52,576

Notes: This table is identical to Table 1.6.2 from the main text, except that it uses alternative definitions for rail captiveness (see description in Appendix A.3). My preferred definition for node unconnectedness is 6.6 miles, the 95th percentile of plants' distance to nearest nodes. Columns (1) and (3) apply a 5-mile threshold for node unconnectedness, while Columns (2) and (4) apply a 10-mile threshold for node unconnectedness. Columns (3)–(4) define captiveness based on node (un)connectedness only, as in Column (6) of Table A.5.9. Each column re-constructs nearest-neighbor weights consistent with its respective definition of captiveness. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.12 conducts sensitivity on shipping cost controls (i.e. $S(\mathbf{T}_{ojms})$ in Equation (1.5)). My preferred specification (Column (1)) controls flexibly controls for shipping costs using the 4-way interaction of: (i) shortest rail shipping distance (an outcome of my rail graph algorithm in Appendix A.3), (ii) monthly diesel fuel cost index (to control for time series variation in shipping costs), (iii) the log of the quantity of coal shipped (to all for increasing returns to scale in rail freight), and (iv) the share of route-miles along rail lines reporting high traffic density (to account for higher costs due to rail network congestion). Assuming increasing returns to scale ignores the potential for rail capacity constraints, whereby adding a marginal rail car may increase average costs. To test for robustness to this assumption, Column (2) replaces the log of shipment size with shipment size in levels. Column (3) replaces the AAR diesel cost index with the STB's Rail Cost Adjustment Factor (RCAF), which incorporates changes in non-fuel variable cost of rail

employees working underground (which increases labor costs), and hours worked per ton of coal produced. I weight-average each variable for each quarter, based on quarterly production across all mines in each county. As the composition of coal production shifts across mines, this will cause even time invariant controls (e.g. seam depth) to become time-varying.

Table A.5.11: Markup Levels – Sensitvity to Commoditiy Controls

	Preferred commod. controls	Force avg orig price $\beta = 1$	Interact orig price $\times$ $\mathbf{1}$[spot]	Interact orig price $\times$ sulfur	Interact sulfur $\times$ year FEs	Adding mine controls
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\text{Captive}]_j$	2.301***	2.124***	2.304***	2.305***	2.270***	2.300***
	(0.655)	(0.649)	(0.662)	(0.658)	(0.653)	(0.654)
k nearest neighbors	3	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	36.85	36.85	36.85	36.85	36.85	36.99
Plants	127	127	127	127	127	127
Captive plants	61	61	61	61	61	61
Observations	77,115	77,115	77,115	77,115	77,115	76,634

Notes: This table is identical to Table 1.6.2 from the main text, except that it conducts sensitivity on commodity controls ($\mathbf{C}_{ojms}$ in Equation (1.5)). Column (1) reproduces my preferred specification (Column (5) of Table 1.6.2), which uses the following uninteracted linear controls: BTU content, sulfur content, ash content, a dummy for spot shipments, a dummy for contracts expiring within 2 years, a dummy for bituminous coal, and the average annual mine-mouth price in each originating county. Column (2) forces the coefficient on average mine-mouth price to be 1. Column (3) allows separate coefficients on average mine-mouth price for spot vs. contract shipments. Column (4) interacts the average mine-mouth price with each shipment's average sulfur content. Column (5) allows year-specific coefficients on sulfur content, which accommodates changes in the shadow price of SO_2 emissions. Column (6) controls for time-varying characteristics in each coal mining county, weight-averaged by monthly coal production: mine age, seam thickness, seam depth, share of coal mined underground, share of mine employees working underground, and hours worked per ton of coal produced. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

shipping.[88] Column (4) removes rail traffic density from the interaction, as this variable may serve as a poor proxy of rail network congestion costs. Finally, Column (5) allows all shipping cost coefficeints to vary by coal rank, which effectively eliminates any potentially confounding differences between western vs. eastern rail shipping costs. My results are quite robust to each of these alternative versions of $S(\mathbf{T}_{ojms})$.

Taken together, Tables A.5.11–A.5.12 demonstrate that the cost controls in Equation (1.5) are unlikely to be misspecified in a way that biases my estimates of average markups. Because my estimates are not sensitive to changes in either $\mathbf{C}_{ojms}$ or $S(\mathbf{T}_{ojms})$, this supports my interpretation of the estimated coefficient $\hat{\tau}$ as the average difference in *markups*, rather than simply the average difference in conditional coal price.

Table A.5.13 includes additional fixed effects, beyond the county and month-of-sample fixed effects in Equation (1.5). This reveals that my results are robust to county-specific

[88]Busse and Keohane (2007) use RCAF to control for time series variation in shipping costs, which uses the AAR fuel price index as an input. However, the RCAF is only published quarterly, and does not provide monthly variation.

Table A.5.12: Markup Levels – Sensitvity to Shipping Cost Controls

	4-way interaction (preferred)	Replace ln(quantity) w/ quantity	Replace diesel index w/ RCAF	Remove traffic density	5-way interaction w/ **1**[sub-bitum.]
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}[\text{Captive}]_j$	2.301***	2.063***	2.286***	2.711***	1.823***
	(0.655)	(0.621)	(0.657)	(0.745)	(0.615)
k nearest neighbors	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes
Mean of dep var	36.85	36.85	36.85	36.85	36.85
Plants	127	127	127	127	127
Captive plants	61	61	61	61	61
Observations	77,115	77,115	77,115	77,115	77,115

Notes: This table is identical to Table 1.6.2 from the main text, except that it conducts sensitivity on shipping cost controls ($S(\mathbf{T}_{ojms})$ in Equation (1.5)). Column (1) reproduces my preferred specification (Column (5) of Table 1.6.2), which uses a four-way linear interaction of: shorest-route shipping distance, AAR fuel price index, log of shipment size, and the proportion of each shortest route on rail lines with high traffic density. Column (2) replaces ln(shipment size) with shipment size in levels. Column (3) replaces the AAR diesel fuel price index with the quarterly Rail Cost Adjustment Factor. Column (4) removes controls for rail traffic density along each route, and uses only a 3-way interaction. Column (5) allows shipping costs to vary across a fifth interacted variable: a dummmy for sub-bituminous coal shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

time tends, county-by-plant-region fixed effects, month-by-shipment type fixed effects, and month-by-coal rank fixed effects. Table A.5.14 estimates Equation (1.5) for each coal rank, and removing coal plants from each region. Average differential markup estimates are larger for bituminous shipments than for sub-bituminous shipments, scaling with their difference in average delivered coal price ($53.92 for bituminous vs. $23.45 for sub-bituminous). My results are largely consistent and retain statistical significance across all five split samples.

Table A.5.15 estimates four additional sensitivities which relate to my identifying assumptions for Equation (1.5). Column (2) restricts the nearest-neighbor matching threshold from 200 miles to 100 miles. This removes the 10 matched captive plants with the greatest distance to their non-captive counterparts, and the resulting point estimate is slightly attenuated but still statistically significant. Column (3) restricts to sample to plants built before 1980, the year that the Staggers Act loosened railroad price regulation, which effectively legalized price discrimination. My point estimate increases slightly for this subset of plants that could not have influenced their rail captiveness; had this coefficient attenuated and lost significance, that would have indicated a likely violation of my identifying assumption for $\hat{\tau}$. Finally Columns (4)–(5) perform sensitivity analysis on my (interacted) regression weights. Column (4) uses only nearest-neighbor weights and treats large and small coal shipments equally; Column (5) uses only shipment-size weights and includes *all* 190 coal plants from Figure 1.5.4. In both cases, my point estimate change only slightly.

Table A.5.13: Markup Levels – Sensitvity to Alternative Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\text{Captive}]_j$	2.301***	2.135***	1.652***	2.220***	2.224***	1.558**
	(0.655)	(0.666)	(0.628)	(0.663)	(0.656)	(0.630)
k nearest neighbors	3	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Coal county FEs	Yes	Yes		Yes	Yes	
County time trends		Yes				
County × plant region FEs			Yes			Yes
Month-of-sample FEs	Yes	Yes	Yes			
Month × shipment type FEs				Yes		
Month × coal rank FEs					Yes	Yes
Mean of dep var	36.85	36.85	36.85	36.85	36.85	36.85
Plants	127	127	127	127	127	127
Captive plants	61	61	61	61	61	61
Observations	77,115	77,115	77,105	77,115	77,115	77,105

Notes: This table is identical to Table 1.6.2 from the main text, except that it uses alternative fixed effects. Column (1) reproduces my preferred specification (Column (5) of Table 1.6.2), which uses originating county fixed effects and month-of-sample fixed effects. County-by-plant region fixed effects subsume county fixed effects. Likewise, month-by-shipment type (i.e. contract vs. spot) and month-by-coal rank fixed effects each subsume month-of-sample fixed effects. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.14: Markup Levels – Sensitvity to Split Samples

	Split by Coal Grade		Removing Plants in Region		
	Bituminous	Sub-bituminous	West	Midwest	South/East
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}[\text{Captive}]_j$	3.866***	1.450**	1.690**	1.564**	2.125***
	(0.784)	(0.690)	(0.780)	(0.682)	(0.682)
k nearest neighbors	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes
Mean of dep var	54.12	23.43	49.55	37.57	28.48
Plants	87	73	89	68	97
Captive plants	32	37	39	39	44
Observations	43,703	32,807	58,376	46,949	48,896

Notes: This table is identical to Table 1.6.2 from the main text, except that it splits the sample by coal rank and plant region. Column (1) includes only shipments of bituminous coal, while Column (2) includes only sub-bituminous coal. Columns (3)-(5) each remove plants from a given coal region, using the same regional definitions as in Table 1.6.5. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.15: Markup Levels – Additional Sensitivities

	Preferred specification	Matching $\leq$ 100 miles	Pre-1980 plants	No shipment size weights	No nearest neighbor wts
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}[\text{Captive}]_j$	2.301***	1.548**	2.483***	2.218***	2.398***
	(0.655)	(0.671)	(0.710)	(0.682)	(0.707)
k nearest neighbors	3	3	3	3	
Balanced panel	Yes	Yes	Yes	Yes	Yes
Mean of dep var	36.85	43.24	36.67	45.46	39.17
Plants	127	109	105	127	190
Captive plants	61	50	49	61	91
Observations	77,115	67,610	66,913	78,487	116,477

Notes: This table is identical to Table 1.6.2 from the main text, except that it conducts four sensitivities. Column (1) reproduces my preferred specification (Column (5) of Table 1.6.2). Column (2) restricts the sample of nearest-neighbor matches to be less than 100 miles apart. Column (3) restricts the sample to plants build prior to the 1980 Staggers Act, which loosened railroad price regulation. While my preferred specification weights by the product of nearest-neighbor weights and shipment size, Columns (4)-(5) each remove one of those weights. Column (4) uses nearest-neighbor weights but counts each observation equally, regardless of shipment size. Column (5) weights by shipment size but ignores nearest-neighbor weights (i.e., including all filled and hollow plants represented in Figure 1.5.4). Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.5.3 Markup Changes, Captiveness Only

I estimate markup changes by interacting the Henry Hub natural gas price (Z_m^{HH}) with M_j, a theoretically informed predictor of how plant j's markups should change with the gas price. However, given that markup *levels* are higher for captive plants, it is natural to ask how markups *change* differentially for captive vs. non-captive plants. I estimate the following lag difference-in-differences specification:

$$\begin{aligned} P_{ojms} &= \tau D_j \cdot Z_{m-L}^{HH} + \sum_{\ell=0}^{L-1} \tau_\ell D_j \cdot \Delta Z_{m-\ell}^{HH} \quad \dots \\ &\quad + \beta_C \mathbf{C}_{ojms} + S(\mathbf{T}_{ojms}\,; \beta_T) + \beta_X \mathbf{X}_{jm} + \eta_{oj} + \delta_m + \varepsilon_{ojms} \end{aligned} \tag{A.38}$$

This specification is identical to Equation (1.9) from the main text, after replacing the continuous predictor M_j with the binary captiveness indicator D_j.

Table A.5.16 reports results for the cumulative effect of gas price changes on coal markups, from estimating Equation (A.38) with $L = 36$ lags. This reveals no statistically detectable changes in markups for captive plants relative to non-captive plants. Figure A.5.19 plots lagged coefficients $\hat{\tau}_\ell$ for each regression in Table A.5.16, where $\hat{\tau}_\ell$ represents the cumulative effect through ℓ months. The results are mostly imprecise across all 36 lags. While results with $k = 1$ nearest neighbors are sometimes negative and statistically significant, they lose significance with additional lags and are not robust to $k = 3$ nearest neighbors.[89]

The bottom-right histogram in Figure 1.5.6 illustrates the main reason why I fail to detect differential markup changes between captive vs. non-captive plants. Markups have changed for *both* groups of plants — there are many non-captive plants who likely also experience markup changes. This means that the $D_j = 0$ group is not "uncontaminated" by market power, and Equation (A.38) compares plants facing effective rail monopolies ($D_j = 1$) to plants facing effective rail duopolies ($D_j = 0$).[90] Based solely on the distribution of M_j, many captive plants should have seen only small markup changes, while many non-captive plants should have experienced relatively large markup changes. Hence, my preferred difference-in-difference specification uses M_j as the cross-sectional component, which incorporates for variation in plants' demand sensitivity, aligns more closely with my oligopoly model.

[89] Negative point estimates would suggest that rail carriers responded to a negative demand shock by *raising* markups to increase profits in coal shipping, which theory predicts should occur only if coal demand is very convex or if the fracking boom *decreased* the elasticity of coal demand. My demand estimates in Figure 1.5.6 reveal that neither of these conditions is likely to hold.

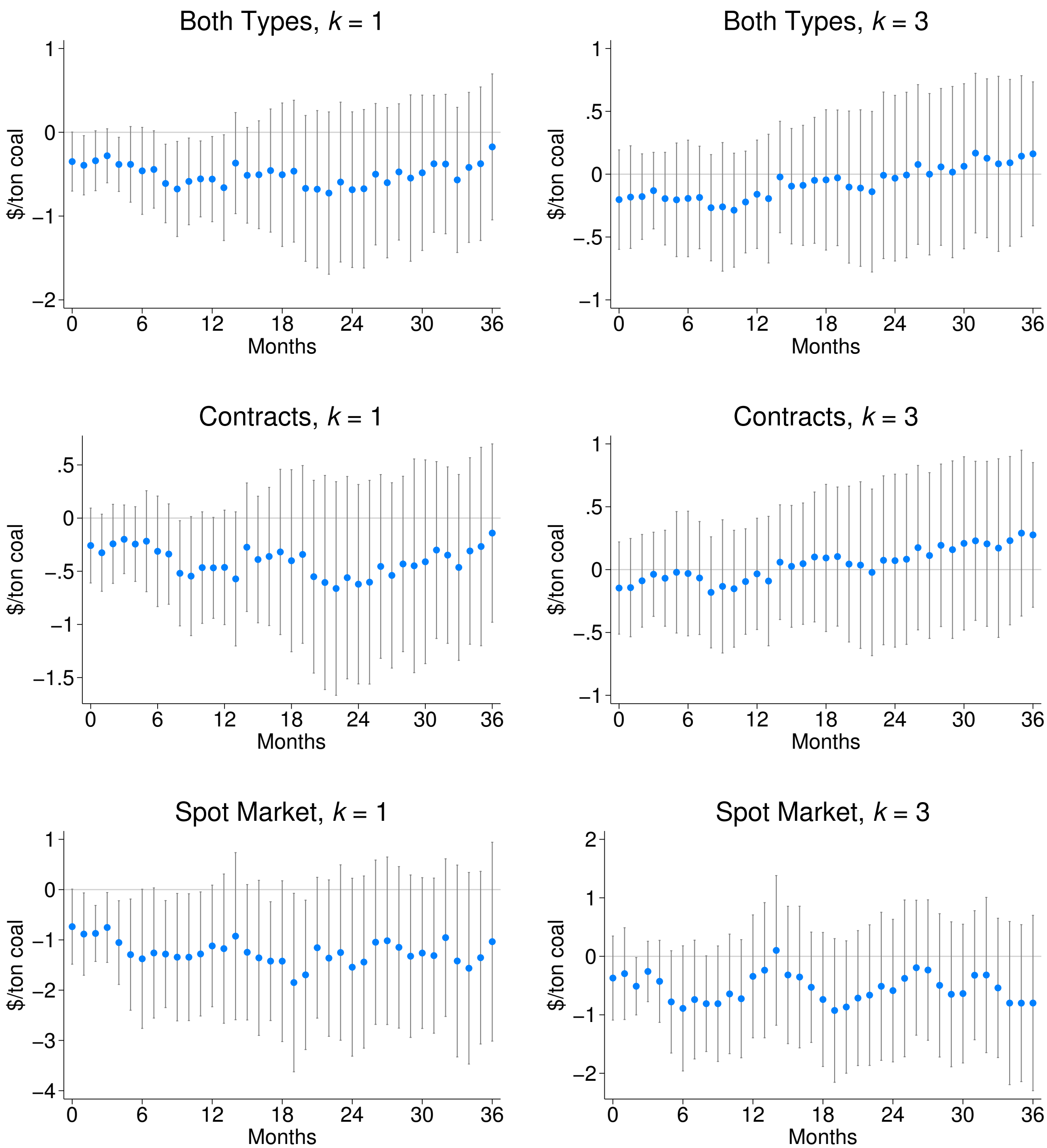
[90] In principle, I could use non-captive plants with a water delivery option as a comparison group, as these plants enjoy the most competitive shipping regimes. Unfortunately, this regression would be underpowered: too few captive plants have matchable non-captive plants that are also located on navigable waterways.

Table A.5.16: Markup Difference in Differences – Captiveness Interacted with Gas Price

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[\text{Captive}]_j \times (\text{Gas Price})_m$	−0.174	0.162	−0.141	0.276	−1.035	−0.799
	(0.439)	(0.289)	(0.422)	(0.291)	(0.995)	(0.757)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	39.86	36.83	39.34	36.30	44.25	40.64
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: Each regression estimates Equation (A.38) at the coal shipment level, with delivered coal price ($ per short ton) as the dependent variable. These regressions are identical to Table 1.6.4, except that they replace M_j with the binary captiveness indicator D_j. This table reports estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.19 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. Matching criteria: up to k nearest neighbors within a 200-mile radius; exact matches on coal rank; and removing non-utility and non-rail plants. Regressions apply nearest-neighbor weights, and also weight each observation by the quantity of coal transacted. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for non-captive plants only. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.5.19: Markup Changes – Cumulative Effects, Captiveness Only



Notes: Each panel corresponds to a column in Table A.5.16, which reports $\hat{\tau}$ only (i.e. the rightmost point in each graph). Each coefficient estimates the interaction of the rail captiveness indicator (D_j) with the ℓ-month lagged difference in natural gas prices ($\Delta Z^{HH}_{m-\ell}$), such that each dot represents the cumulative effect through ℓ months. Whiskers denote 95 percent confidence intervals for each point estimate, with standard errors clustered by plant.

A.5.4 Markup Changes

A.5.4.1 Additional Results

Here, I present results omitted from the main text for the sake of brevity. Figure A.5.20 plots the lagged coefficients $\hat{\tau}_\ell$ for each split sample regression in Table 1.6.5. These plots confirm that bituminous coal and plants in the South/East contribute most of my estimated markups changes. While the sub-bituminous coal estimates in the top-right panel only gain significance after 36 months, the steady upward trend in cumulative effects suggests that the cumulative effects may continue to grow with additional lags.[91] Two factors likely explain why sub-bituminous markups change less than bituminous markups due to gas price changes. First, sub-bituminous coal prices are much lower on average (\$22/ton in my sample, compared to \$48/ton for bituminous coal), and I estimate small markup levels for sub-bituminous coal (see Table A.5.14). Hence, lower delivered prices and lower markups should translate to small changes in markup levels. Second, the western rail network is relatively sparse, and sub-bituminous shipping routes from Wyoming's Powder River Basin may be more likely to face binding rate regulation, which would limit M_j's predictive power.

In Section 1.6.2 of the main text, I report results for balanced panels only.[92] Table A.5.17, Figure A.5.21, and Table A.5.18 report the same results as Table 1.6.4, Figure 1.6.7, and Table 1.6.6, respectively. This reveals that my estimates for markup changes are not sensitive to the removal of retiring coal plants. However, my preferred estimates (from the main text) exclude these plants, as they are no longer relevant for future policy projections.

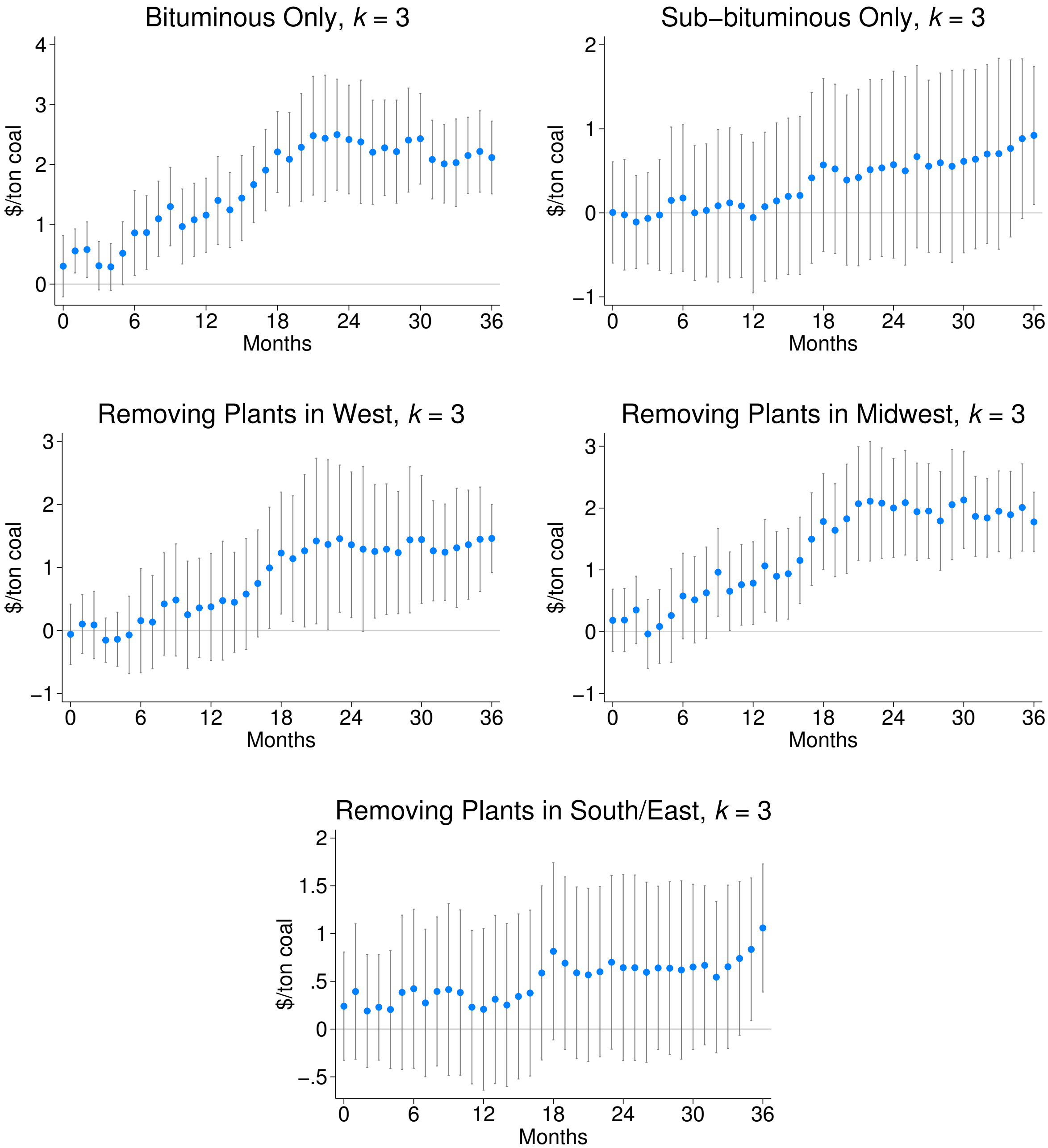
Finally, Tables A.5.19–A.5.20 report results that are analogous to Table 1.6.4, except that the dependent variable is in units of \$/MMBTU rather than \$/ton. As expected, these point estimates are roughly 1/20 the magnitude of those in Table 1.6.4, reflecting the average MMBTU-to-ton conversion rate. Tables A.5.19–A.5.20 use consistent units across P_{ojms} and M_j, and these point estimates serve as inputs into my pass-through estimates in Table 1.7.7 (and Tables A.5.37–A.5.39 below). However, these estimates may misspecify transportation costs, which depend on coal's weight rather than its energy content. Hence, my preferred specifications use an outcome variable in units of \$/ton.

[91] Indeed, cumulative effects for sub-bituminous coal increase in magnitude and remain statistically significant, for $L = 48$ lags.

[92] Recall that what I refer to as a "balanced" panel is not fully balanced, because not every plant purchases coal every month.

‘

Figure A.5.20: Contract Shipments & Split Samples, Cumulative Effects



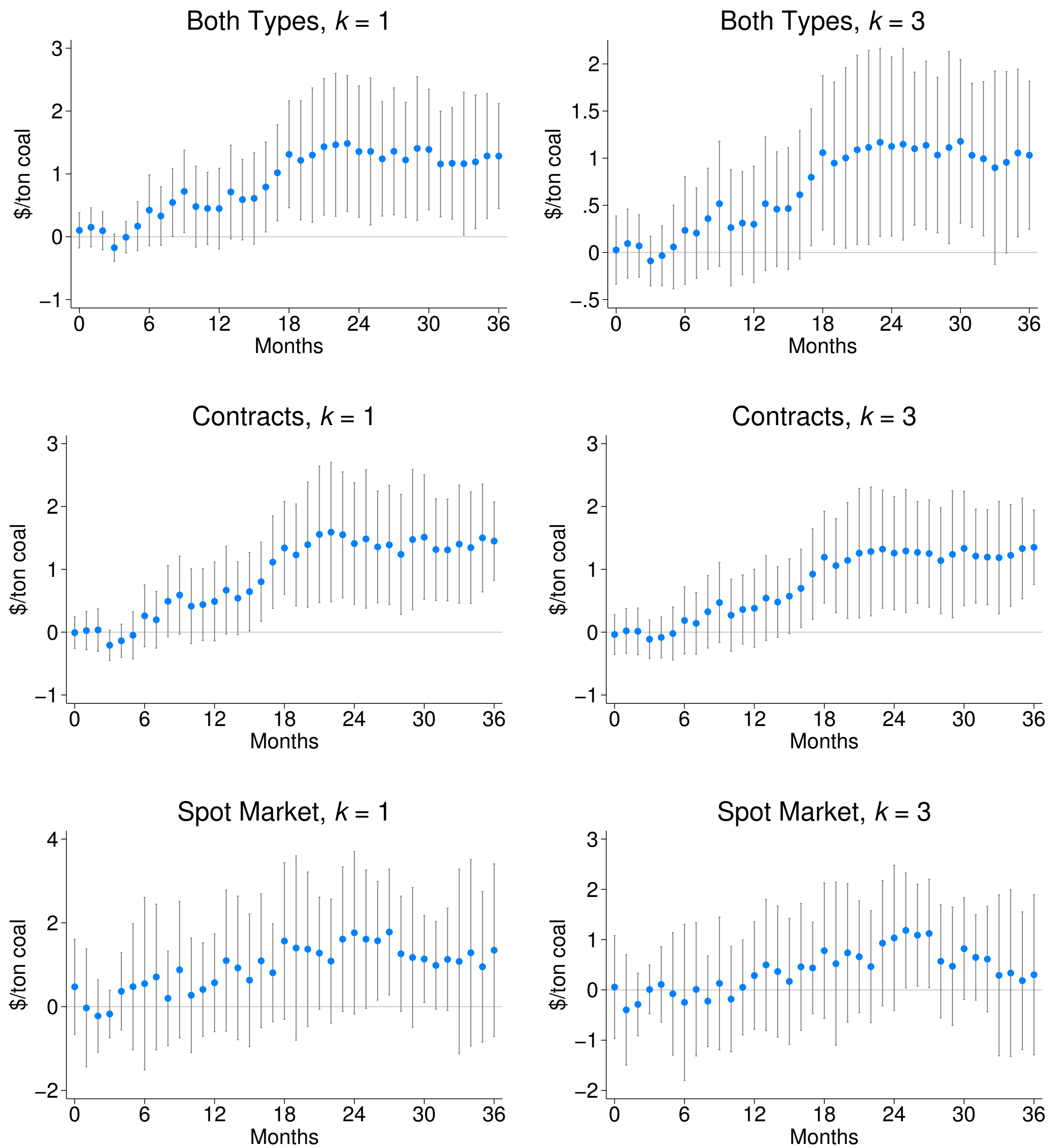
Notes: This figure plots 36 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{35}$) and $\hat{\tau}$, from 5 separate regressions of Equation (1.9) with $L = 36$ lags. Each panel corresponds to a column in Table 1.6.5 in the main text, which reports $\hat{\tau}$ only (i.e. the rightmost point in each graph). Each coefficient estimates the interaction of M_j with the ℓ-month lagged difference in natural gas prices ($\Delta Z^{HH}_{m-\ell}$), such that each dot represents the cumulative effect through ℓ months. Whiskers denote 95 percent confidence intervals for each point estimate, with standard errors clustered by plant. See the notes below Table 1.6.5 for further details on the estimation.

Table A.5.17: Markup Changes – Sensitivity to Unbalanced Panel

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.284***	1.031**	1.450***	1.351***	1.348	0.301
	(0.422)	(0.398)	(0.315)	(0.301)	(1.040)	(0.807)
k nearest neighbors	1	3	1	3	1	3
Balanced panel						
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.70	34.51	32.45	33.36	42.11	41.88
Plants	135	177	131	173	120	160
Plant-county-months	31,159	42,774	25,997	35,533	8,153	11,367
Observations	64,022	85,788	50,392	67,368	13,513	18,262

Notes: This table is identical to Table 1.6.4, except that each regression uses the full (unbalanced) panel of coal plants. I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.21 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.5.21: Markup Changes – Cumulative Effects, Sensitivity to Unbalanced Panel



Notes: This figure plots 36 lag-differenced DD coefficient estimates $(\hat{\tau}_0, \ldots, \hat{\tau}_{35})$ and $\hat{\tau}$, for each regression in Table A.5.17. It is analogous to Figure 1.6.7 from the main text.

Table A.5.18: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, Unbalanced Panel

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{GasPrice})_m$	0.010	−0.082	0.201	0.198	−0.844*	−1.363**
	(0.206)	(0.181)	(0.207)	(0.178)	(0.488)	(0.623)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{GasPrice})_m$	0.236	0.081	0.458*	0.393*	−0.328	−1.046*
	(0.231)	(0.200)	(0.236)	(0.209)	(0.443)	(0.564)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{GasPrice})_m$	0.667**	0.474**	0.692***	0.589***	0.864	0.121
	(0.270)	(0.238)	(0.249)	(0.212)	(0.791)	(0.905)
$\mathbf{1}\Big[M_j \in (0.70, 2.00]\Big] \times (\text{GasPrice})_m$	1.203**	0.914**	1.362***	1.190***	1.074	−0.128
	(0.492)	(0.366)	(0.462)	(0.358)	(1.054)	(0.920)
k nearest neighbors	1	3	1	3	1	3
Balanced panel						
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	31.84	35.18	31.07	34.88	34.46	36.26
Plants	135	177	131	173	120	160
Plant-county-months	31,159	42,774	25,997	35,533	8,153	11,367
Observations	64,022	85,788	50,392	67,368	13,513	18,262

Notes: This table is identical to Table 1.6.6, except that each regression uses the full (unbalanced) panel of coal plants. I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.19: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, \$/MMBTU, Balanced Panel

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{GasPrice})_m$	0.002	−0.001	0.013	0.014	−0.041	−0.060*
	(0.011)	(0.010)	(0.011)	(0.009)	(0.027)	(0.031)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{GasPrice})_m$	0.013	0.007	0.026**	0.023**	−0.017	−0.044*
	(0.011)	(0.010)	(0.011)	(0.010)	(0.022)	(0.026)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{GasPrice})_m$	0.033**	0.026**	0.036***	0.033***	0.042	0.015
	(0.013)	(0.011)	(0.011)	(0.010)	(0.036)	(0.040)
$\mathbf{1}\Big[M_j \in (0.70, 2.00]\Big] \times (\text{GasPrice})_m$	0.062***	0.049***	0.071***	0.063***	0.061	0.010
	(0.021)	(0.017)	(0.021)	(0.016)	(0.049)	(0.043)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	1.58	1.69	1.53	1.67	1.74	1.77
Plants	94	124	92	122	85	115
Plant-county-months	26,060	35,651	22,000	29,806	6,796	9,630
Observations	56,219	75,089	44,651	59,178	11,487	15,797

Notes: This table is identical to Table 1.6.6, except that it uses coal prices in dollars per MMBTU (i.e. coal's value to power plants as a fuel), rather than dollars per ton (i.e. coal's value to freight shippers). I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months, and I use these point estimates to construct pass-through rates in Tables 1.7.7 and A.5.37. I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.20: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, \$/MMBTU, Unbalanced

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{GasPrice})_m$	-0.000	-0.004	0.010	0.010	-0.044^{*}	-0.065^{**}
	(0.011)	(0.010)	(0.011)	(0.010)	(0.027)	(0.031)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{GasPrice})_m$	0.011	0.004	0.023^{**}	0.019^{*}	-0.021	-0.049^{*}
	(0.011)	(0.010)	(0.011)	(0.010)	(0.022)	(0.025)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{GasPrice})_m$	0.029^{**}	0.021^{*}	0.032^{***}	0.027^{***}	0.034	0.006
	(0.013)	(0.011)	(0.012)	(0.010)	(0.035)	(0.039)
$\mathbf{1}\Big[M_j \in (0.70, 2.00]\Big] \times (\text{GasPrice})_m$	0.055^{**}	0.042^{***}	0.064^{***}	0.056^{***}	0.046	-0.002
	(0.022)	(0.016)	(0.020)	(0.016)	(0.047)	(0.041)
k nearest neighbors	1	3	1	3	1	3
Balanced panel						
Plant $\times$ county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	1.59	1.70	1.54	1.67	1.76	1.79
Plants	135	177	131	173	120	160
Plant-county-months	31,159	42,774	25,997	35,533	8,153	11,367
Observations	64,022	85,788	50,392	67,368	13,513	18,262

Notes: This table is identical to Table A.5.18, except that it uses coal prices in dollars per MMBTU (i.e. coal's value to power plants as a fuel), rather than dollars per ton (i.e. coal's value to freight shippers). I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months, and I use these point estimates to construct pass-through rates in Tables A.5.38 and A.5.39. I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.5.4.2 Specification Robustness

In this section, I test my difference-in-differences results for sensitivity to the specification of Equation (1.9). I report sensitivities for the pooled estimates in Table 1.6.4, and sensitivity results are similar for estimates that discretize M_j into quintiles (as in Table 1.6.6).

My main specification estimates markup changes using $L = 36$ monthly lags. However, this number is arbitrary, as I have no reason to believe that railroads should adjust markups in three-year intervals. Here, I show that my results are consistent if I use $L = 24$ lags (Table A.5.21 and Figure A.5.22) or $L = 48$ lags (Table A.5.22 and Figure A.5.23). Allowing for effects to accumulate into a fourth year increases my effect sizes, while demonstrating that decreases to markups are quite persistent.

Figure A.5.24 converts my main specification into the style of an event study. Rather than control for lagged changes to the gas price, these estimates interact M_j with quarter-of-sample dummies in the following specification (indexing quarters by q):

$$\begin{aligned} P_{ojms} = &\sum_q \tau_q M_j \cdot \mathbf{1}[m \in q] \quad \ldots \\ &+ \beta_C \mathbf{C}_{ojms} + S(\mathbf{T}_{ojms}\,; \beta_T) + \beta_X \mathbf{X}_{jm} + \eta_{oj} + \delta_m + \varepsilon_{ojms} \end{aligned} \tag{A.39}$$

These plots reveal a level shift in point estimates after 2010 which parallels the decrease in gas prices (in light blue)—as gas prices fell, markups differentially fell for plant with higher M_j. High point estimates in the 6 months following the 2008 gas price spike are consistent with the 6-month lag in markup changes from Figure 1.6.7. The vertical height of the dark blue point estimates is not meaningful in these event study figures, as they are calibrated to an arbitrary omitted bin; the purpose of these plots is illustrate how *changes* in these point estimates align with variation in the gas price time series.

Equation (1.9) controls for origin-by-destination fixed effects, which subsume coal country fixed effects and plant fixed effects. However, by controlling for the average price within each route, my estimates cannot capture changes to the *composition* of plant routes. It is possible that the average markup on a given oj route has changed, but the prices that plant j receives have not changed (if plant j now purchases coal from a different county o). Table A.5.23 and Figure A.5.25 report results from estimating Equation (1.9) with separate (uninteracted) county and plant fixed effects. These point estimates are slightly larger in magnitude, which suggests that in addition to decreasing coal markups along individual oj routes, low gas prices may have incentivized coal plants to seek out less expensive coal shipments from other counties.

Table A.5.21: Markup Changes – Sensitivity to $L = 24$ Lags

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.414***	1.203***	1.541***	1.471***	1.669*	0.690
	(0.425)	(0.418)	(0.353)	(0.346)	(0.986)	(0.799)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.27	34.20	32.07	33.04	41.76	41.39
Plants	94	124	92	122	85	115
Plant-county-months	26,060	35,651	22,000	29,806	6,796	9,630
Observations	56,219	75,089	44,651	59,178	11,487	15,797

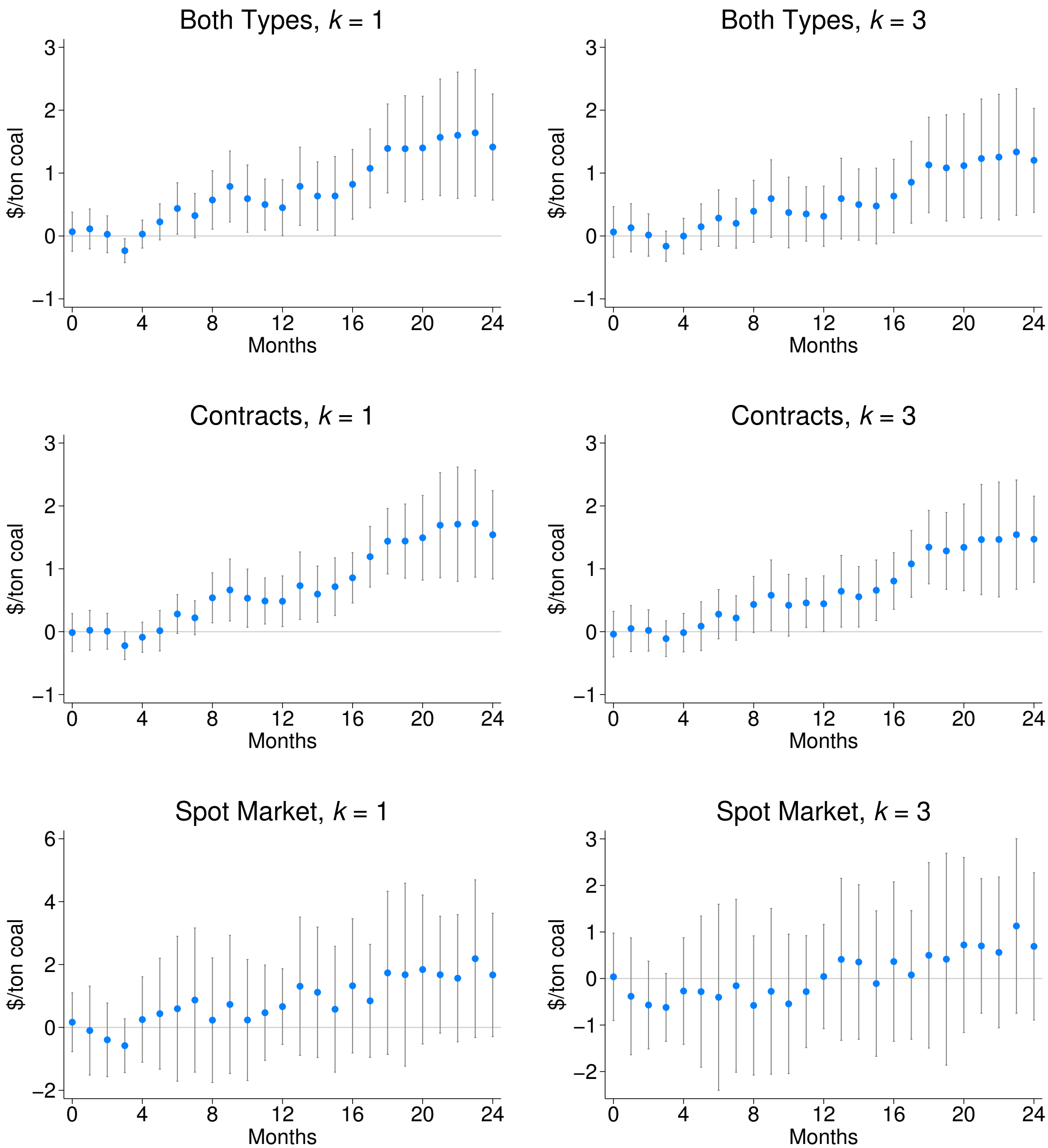
Notes: This table is identical to Table 1.6.4, except that each regression uses $L = 24$ monthly lags rather than $L = 36$ monthly lags. Figure A.5.22 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.22: Markup Changes – Sensitivity to $L = 48$ Lags

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.599***	1.263***	1.877***	1.730***	1.429	0.362
	(0.481)	(0.470)	(0.331)	(0.309)	(1.010)	(0.908)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.27	34.20	32.07	33.04	41.76	41.39
Plants	94	124	92	122	85	115
Plant-county-months	26,060	35,651	22,000	29,806	6,796	9,630
Observations	56,219	75,089	44,651	59,178	11,487	15,797

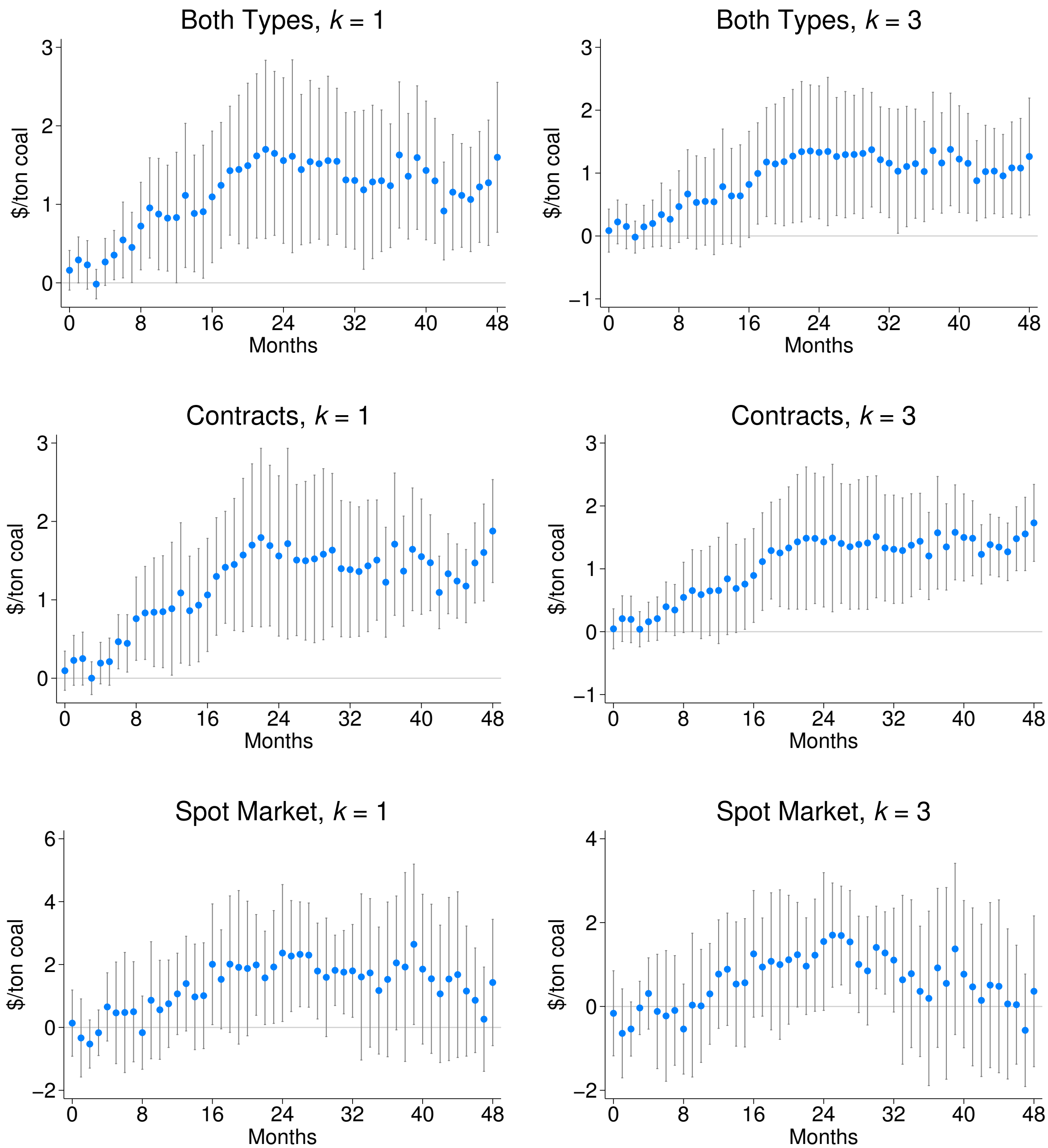
Notes: This table is identical to Table 1.6.4, except that each regression uses $L = 48$ monthly lags rather than $L = 36$ monthly lags. Figure A.5.23 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.5.22: Markup Changes – Cumulative Effects, Sensitivity to $L = 24$ Lags



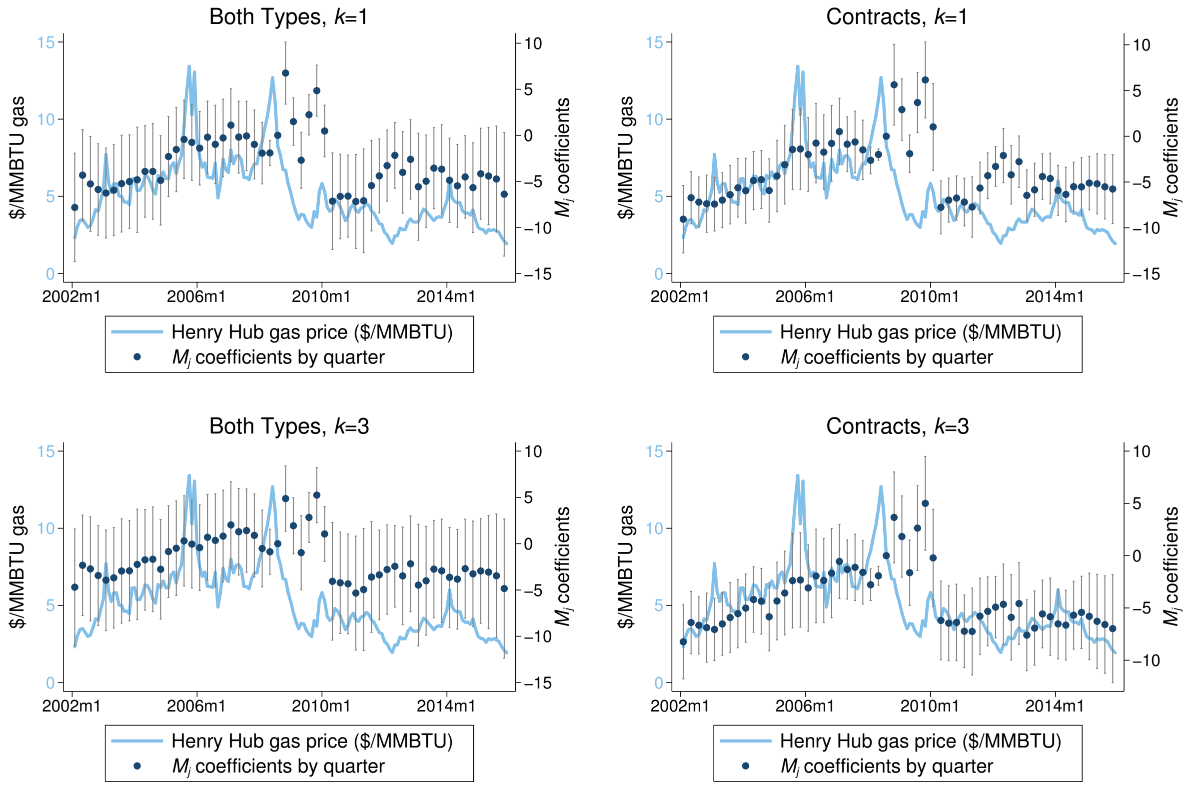
Notes: This figure plots 24 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{23}$) and $\hat{\tau}$, for each regression in Table A.5.21. It is analogous to Figure 1.6.7 from the main text.

Figure A.5.23: Markup Changes – Cumulative Effects, Sensitivity to $L = 48$ Lags



Notes: This figure plots 48 lag-differenced DD coefficient estimates $(\hat{\tau}_0, \ldots, \hat{\tau}_{47})$ and $\hat{\tau}$, for each regression in Table A.5.22. It is analogous to Figure 1.6.7 from the main text.

Figure A.5.24: Markup Changes – Event-Study Style Plot



Notes: This figure converts Equation (1.9) from a lagged difference-in-differences specification into an event-study style specification (Equation (A.39)). Instead of including monthly lags of the gas price, I interact M_j with quarter-of-sample dummies, and plot these coefficients against the gas price time series (in light blue). The vertical height of the dark blue point estimates is not meaningful, as they are calibrated to an arbitrary omitted bin (quarter 3 of 2008, the last quarter of high gas prices). Panels correspond to Columns (1)–(4) of Table 1.6.4, using a balanced panel of plants. Standard errors are clustered by plant, and whiskers denote 95 percent confidence intervals.

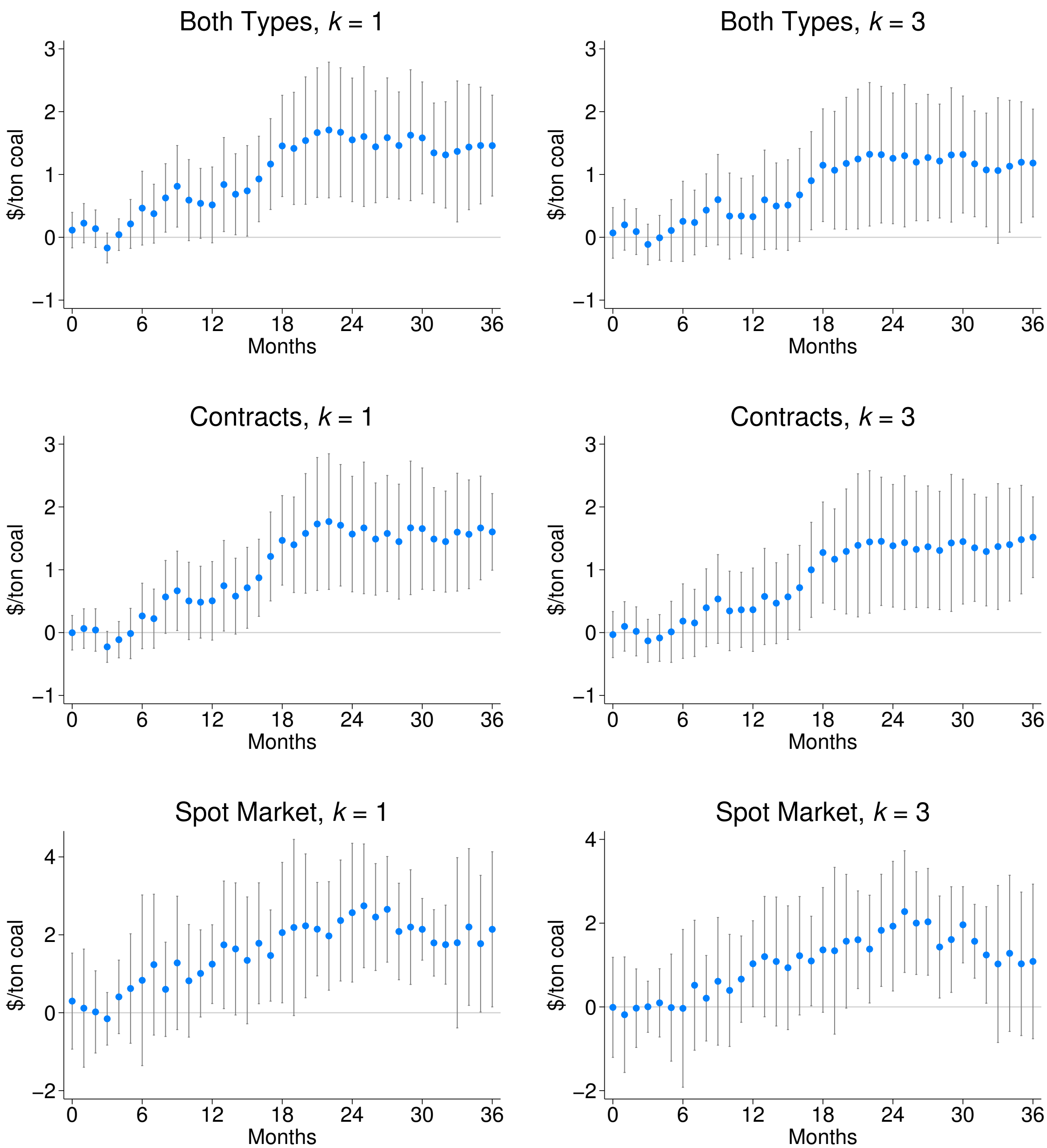
As with any panel fixed effects regression, time-varying unobservables have the potential to bias my estimates of $\hat{\tau}$ in Equation (1.9). Table A.5.24 and Figure A.5.26 interact month-of-sample fixed effects with coal basin fixed effects, which has little effect on my results.[93] This suggests that time-varying unobservables related to differences in regional coal markets are unlikely to bias my point estimates. Table A.5.25 and Figure A.5.27 interact month-of-sample fixed effects with fixed effects for each plant's electricity market region, which also has little effect on my results. This shows that time-varying confounders related to differences in regional electricity markets are similarly unlikely.

[93] I define 8 coal basins, following EIA's coal mine datasets: Northern Appalachia, Central Appalachia, Southern Appalachia, Illinois Basin, Interior, Powder River Basin, Uinta Region, and Western.

Figures A.5.28 –A.5.29 plots pre-fracking trends in delivered coal prices, revealing that my assumption of parallel trends broadly holds across different quintiles of M_j.

'

Figure A.5.25: Markup Changes – Cumulative Effects, Sensitivity to County + Plant FEs



Notes: This figure plots 36 lag-differenced DD coefficient estimates $(\hat{\tau}_0, \ldots, \hat{\tau}_{35})$ and $\hat{\tau}$, for each regression in Table A.5.23. It is analogous to Figure 1.6.7 from the main text.

Table A.5.23: Markup Changes – Sensitivity to Separate County + Plant Fixed Effects

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.460***	1.182***	1.603***	1.518***	2.141**	1.086
	(0.404)	(0.434)	(0.307)	(0.325)	(1.001)	(0.932)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant FEs	Yes	Yes	Yes	Yes	Yes	Yes
Coal county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.31	34.25	32.11	33.08	41.89	41.47
Plants	94	124	92	122	86	116
Plant-county-months	26,185	35,834	22,083	29,933	6,903	9,785
Observations	56,344	75,272	44,734	59,305	11,594	15,952

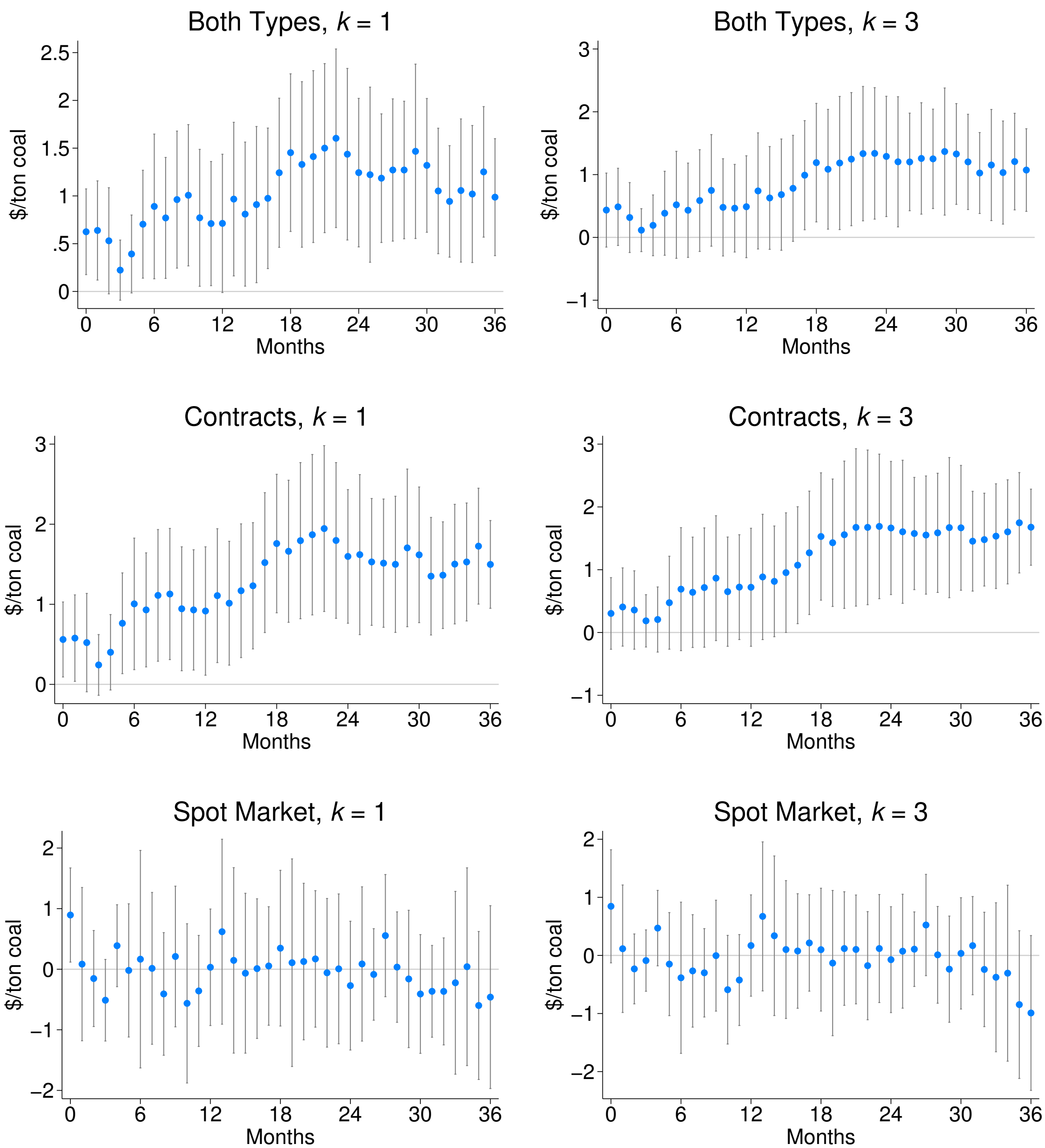
Notes: This table is identical to Table 1.6.4, except that it uses separate uninteracted county and plant fixed effects, rather than interacted county-by-plant fixed effects. I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.25 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.24: Markup Changes – Sensitivity to Basin-by-Month Fixed Effects

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	0.987***	1.072***	1.498***	1.677***	−0.461	−0.991
	(0.308)	(0.332)	(0.276)	(0.306)	(0.759)	(0.673)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month × basin FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.26	34.19	32.07	33.03	41.13	41.21
Plants	94	124	92	122	85	115
Plant-county-months	25,996	35,587	21,925	29,732	6,644	9,497
Observations	56,155	75,025	44,576	59,104	11,335	15,664

Notes: This table is identical to Table 1.6.4, except that it interacts month-of-sample fixed effects with coal basin fixed effects. I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.26 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.5.26: Markup Changes – Cumulative Effects, Sensitivity to Basin-by-Month FEs



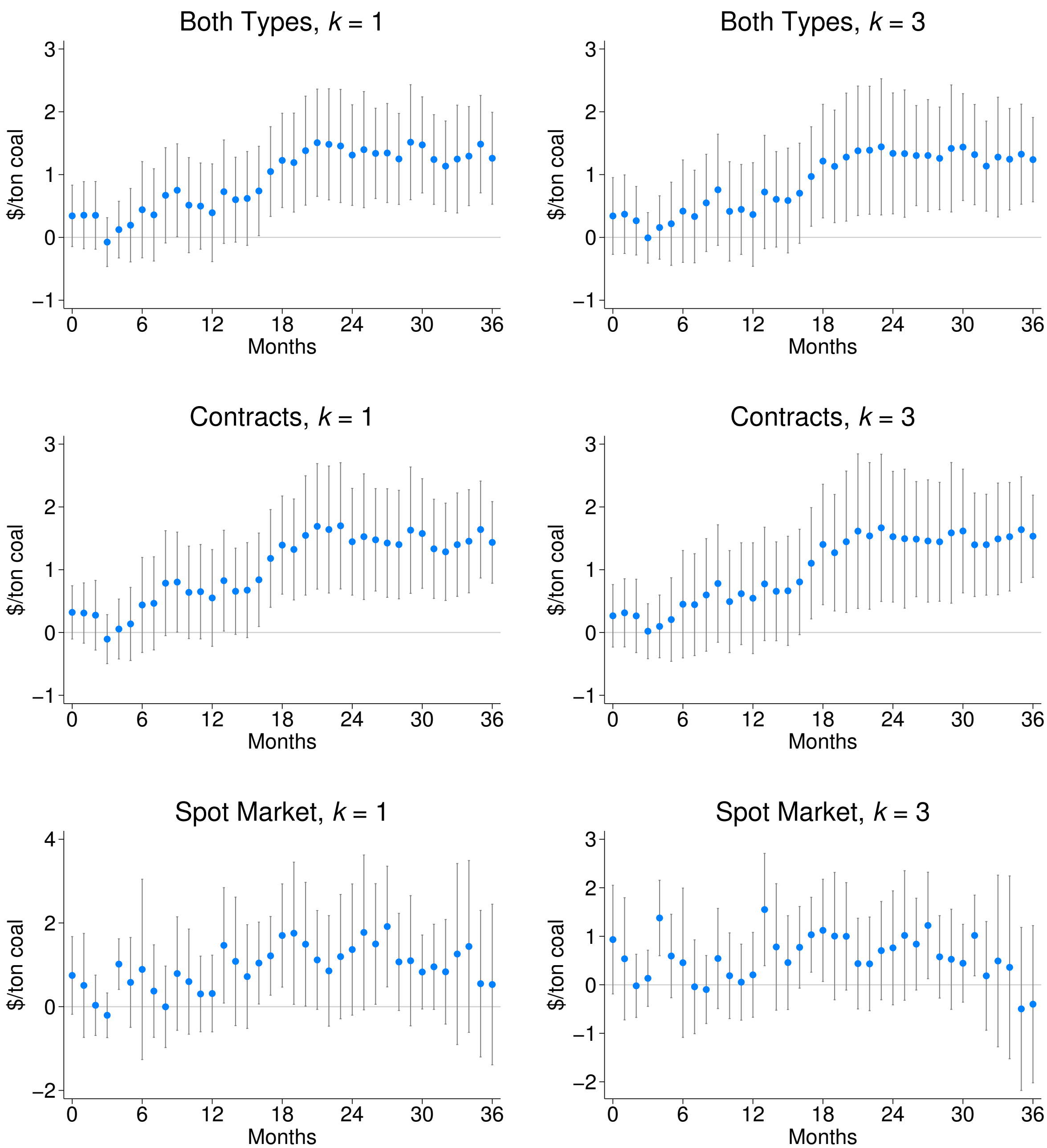
Notes: This figure plots 36 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{35}$) and $\hat{\tau}$, for each regression in Table A.5.24. It is analogous to Figure 1.6.7 from the main text.

Table A.5.25: Markup Changes – Sensitivity to Market Region-by-Month Fixed Effects

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta \widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.260***	1.238***	1.435***	1.533***	0.530	−0.400
	(0.369)	(0.339)	(0.327)	(0.331)	(0.965)	(0.818)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month × plant region FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.27	34.20	32.07	33.04	41.20	40.91
Plants	94	124	92	122	85	115
Plant-county-months	26,018	35,609	21,958	29,764	6,693	9,530
Observations	56,177	75,047	44,609	59,136	11,384	15,697

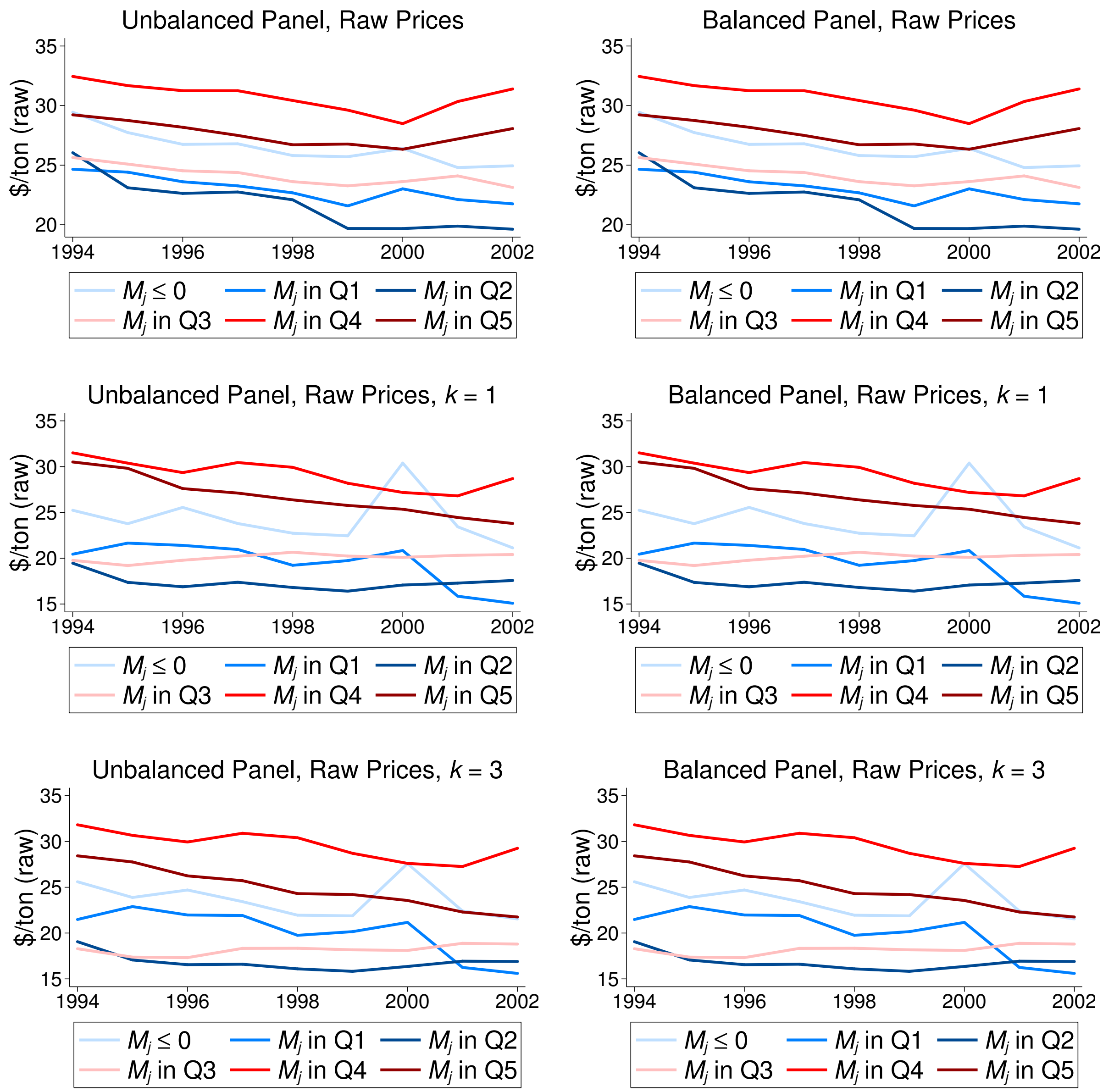
Notes: This table is identical to Table 1.6.4, except that it interacts month-of-sample fixed effects with fixed effects for each electricity market region (i.e. ISO or NERC region). I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.27 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.5.27: Markup Changes – Cumulative Effects, Region-by-Month FEs



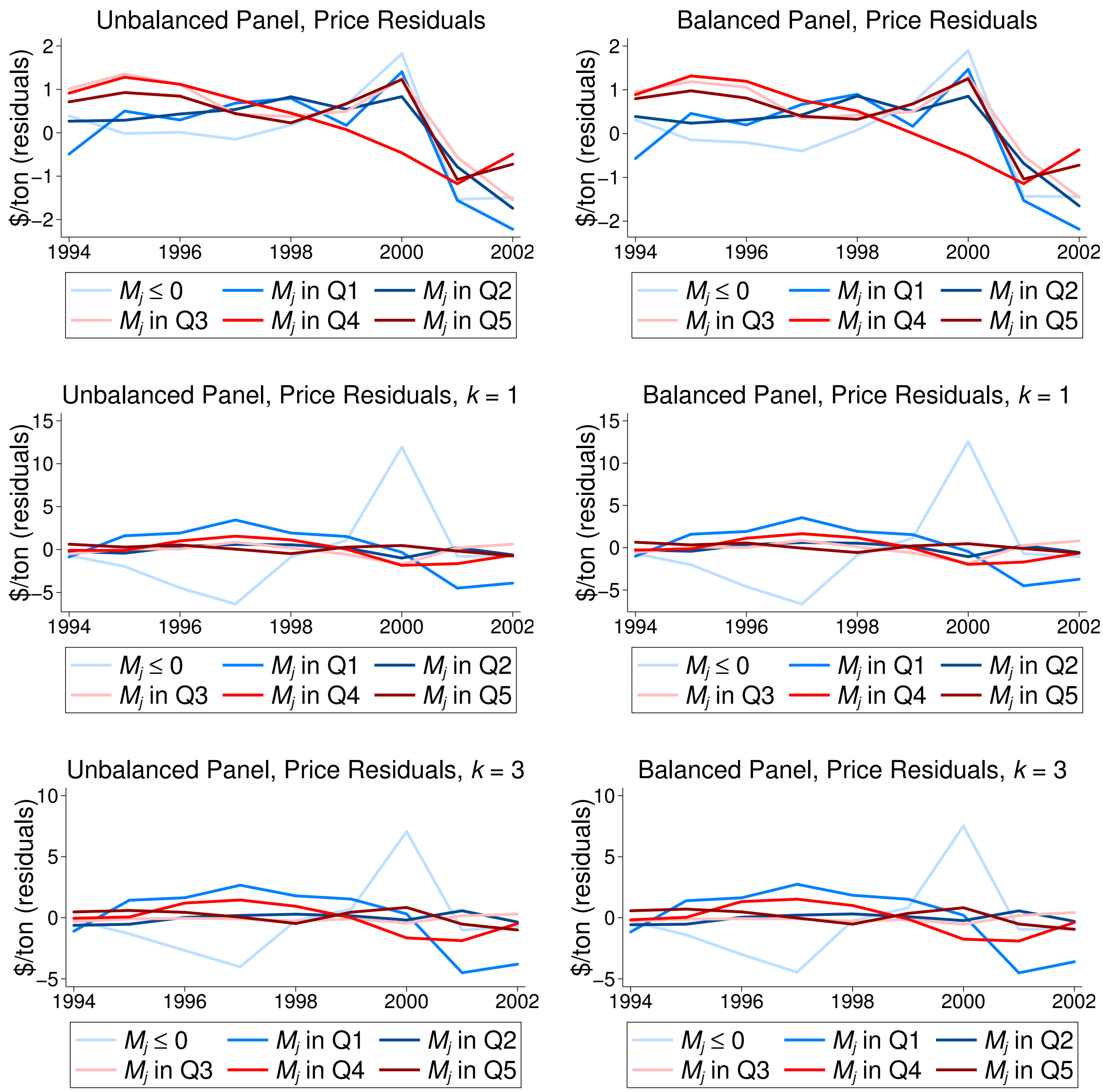
Notes: This figure plots 36 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{35}$) and $\hat{\tau}$, for each regression in Table A.5.25. It is analogous to Figure 1.6.7 from the main text.

Figure A.5.28: Pre-Trends in Delivered Coal Prices (Raw Coal Prices)



Notes: This figures reports annual average delivered coal prices for 1994–2002, for six groups of coal plants: the five quintiles of the positive support of M_j (four of which are reported in Table 1.6.6); and a sixth group of plants with $M_j \leq 0$ (the omitted comparison group in Table 1.6.6, along with the quintile 1). The outcome variable is raw average annual coal prices for plants in each group. The left column uses an unbalanced panel of plants, while the right column includes only the 2002–2015 balanced panel. I impose no additional sample restrictions to account for missing plant-years prior to 2002. Some group-years contain very few observations (most notably for the $M_j \leq 0$ group in 2000), which generate yearly deviations from trend that are not persistent. The top row does not weight plants; the middle and bottom rows apply nearest-neighbor weights.

Figure A.5.29: Pre-Trends in Delivered Coal Prices (Residualized Prices)



Notes: This figures reports annual average delivered coal prices for 1994–2002, for six groups of coal plants: the five quintiles of the positive support of M_j (four of which are reported in Table 1.6.6); and a sixth group of plants with $M_j \leq 0$ (the omitted comparison group in Table 1.6.6, along with the quintile 1). The outcome variable is average annual prices after partialing out all right-hand-side variables in Equation (1.9) other than the M_j interactions. The left column uses an unbalanced panel of plants, while the right column includes only the 2002–2015 balanced panel. I impose no additional sample restrictions to account for missing plant-years prior to 2002. Some group-years contain very few observations (most notably for the $M_j \leq 0$ group in 2000), which generate yearly deviations from trend that are not persistent. The top row does not weight plants; the middle and bottom rows apply nearest-neighbor weights.

My preferred difference-in-differences specification (Equation (1.9)) interacts the cross-sectional predictor M_j with the time series of Henry Hub natural gas spot price $Z^{HH}_{m-\ell}$. However, natural gas prices vary cross-sectionally due to regional heterogeneity in both proximity to gas production and pipeline capacity (see Appendix A.2.5). I prefer a pure time series of gas prices for econometric identification—localized changes in the gas price could reflect differential competitiveness of coal across regions, which could introduce a time-varying confounder in a manner similar to bad controls.[94] Table A.5.26 and Figure A.5.30 replace $Z^{HH}_{m-\ell}$ with state-specific average monthly gas prices for the electric power sector (from EIA's *Electric Power Monthly*). This has little effect on my results.

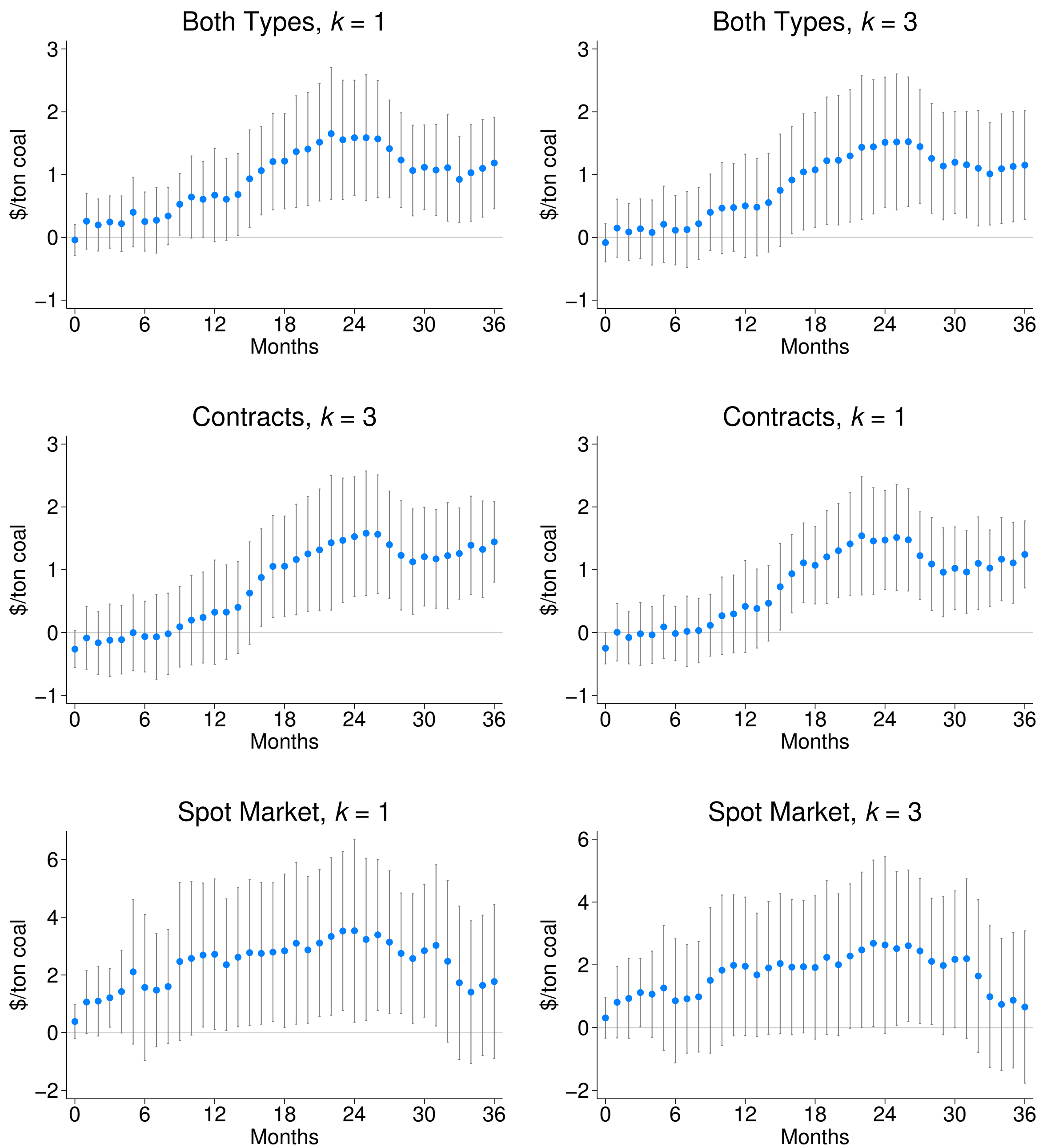
Table A.5.26: Markup Changes – Sensitivity to State-Specific Gas Prices

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.185***	1.150***	1.243***	1.443***	1.771	0.656
	(0.367)	(0.437)	(0.268)	(0.324)	(1.341)	(1.224)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	35.50	37.05	34.12	35.47	46.72	48.90
Plants	91	121	90	120	79	106
Plant-county-months	19,926	26,924	16,856	22,743	4,922	6,773
Observations	42,012	55,337	33,874	44,552	8,066	10,683

Notes: This table is identical to Table 1.6.4, except that it replaces the pure time series of Henry Hub natural gas spot prices ($Z^{HH}_{m-\ell}$ from Equation (1.9)) with state-specific average monthly gas prices. These are state-by-month average prices of natural gas delivered to the Electric Power Sector, from EIA's *Electric Power Monthly* (see Appendix A.2.1.2). I report estimates for $\hat{\tau}$, or the cumulative effects over $L = 36$ months. Figure A.5.30 plots each lagged coefficient $\hat{\tau}_\ell$, which reports the cumulative effect through ℓ months. I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[94]For example, consider two electricity market regions with both coal and gas generation. In Region A, coal generators are extremely efficient and would face little competitive pressure from low gas prices; in Region B, coal generators are less efficient and are likely to become marginal as gas prices fall. A negative shock to the Henry Hub gas price could increase gas demand in Region B and weaken the effect on of the negative shock on localized gas prices in that region. In this scenario, using regionally differentiated gas prices could introduce positive correlation between $\hat{\tau}$ and the error term.

Figure A.5.30: Markup Changes – Cumulative Effects, State-Specific Gas Prices



Notes: This figure plots 36 lag-differenced DD coefficient estimates ($\hat{\tau}_0, \ldots, \hat{\tau}_{35}$) and $\hat{\tau}$, for each regression in Table A.5.26. It is analogous to Figure 1.6.7 from the main text.

Tables A.5.27–A.5.31 conduct sensitivities analogous to those reported above in Appendix A.5.2.[95] Table A.5.27 shows that my estimates for markup changes are not sensitive to the definition of "route unconnectedness" or the subset of shipping routes that I use to define rail captiveness. Likewise, Columns (1)–(2) of Table A.5.28 show that my estimates are not sensitive to my 6.6-mile threshold for "node unconnectedness". However, a captiveness definition based solely on proximity to rail nodes cannot capture the full effect of heterogeneous railroad market power (see Column (6) of Table A.5.27, and Columns (3)–(4) of Table A.5.28). Table A.5.29 shows that my estimates are not sensitive to controls for coal commodity costs in $\mathbf{C}_{ojms}$, while Table A.5.30 shows that they are not sensitive to shipping cost controls in $S(\mathbf{T}_{ojms})$. Finally, Table A.5.31 shows that my results are robust to: a stricter nearest-neighbor distance threshold (Column (2)); restricting the sample to plants built prior to the 1980 Staggers Act (Column (3)); removing shipment size weights and treating each observation (rather than each coal ton) equally (Column (4)); and removing nearest-neighbor weights to include all plants represented in Figure 1.5.4 (Column (5), albeit with slightly attenuated magnitudes).

Table A.5.27: Markup Changes – Sensitivity to Definition of Captiveness

	Unconnected across all observed routes			Average route	All potential routes	Only node connectedness
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\widehat{\text{Markup}})_j \times (\text{GasPrice})_m$	1.187***	1.170***	1.191***	0.941**	1.190***	0.836*
	(0.408)	(0.403)	(0.404)	(0.420)	(0.407)	(0.481)
k nearest neighbors	3	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Route unc. cutoff (miles)	300	150	600	300	300	
Mean of dep var	34.20	34.25	33.73	32.95	34.26	36.12
Plants	124	126	124	129	123	117
Plant-county-months	35,651	36,890	35,714	35,343	35,613	34,131
Observations	75,089	76,966	75,396	72,339	75,062	70,444

Notes: This table is identical to Table 1.6.4, except that it uses alternative definitions for rail captiveness to construct M_j. Column (1) uses my preferred definition described in Appendix A.3, reported in Column (2) of Table 1.6.4. Columns (2)–(3) halve and double my preferred 300-mile threshold for route unconnectedness. Column (4) weakens the definition of captiveness such that only a plant's *average* route need be unconnected (rather than *all* observed routes). Column (5) strengthens the definition of captiveness to include all *potential* routes with observationally similar coal, even if I do not observe any deliveries along such routes. Column (6) defines captiveness based on node (un)connectedness only. I report means of the dependent variable for plants with $M_j = 0$, and regressions pool contract and spot shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[95] I report results only for the balanced panel with $k = 3$ nearest-neighbor matches; results are similar for $k = 1$ nearest neighbors and for unbalanced panels. Please see Appendix A.5.2 for detailed descriptions of how I modify my main specification for each of these sensitivity regressions.

Table A.5.28: Markup Changes – Sensitivity to Definition of Captiveness

	Unconnected across all observed routes		Only node connectedness	
	(1)	(2)	(3)	(4)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.230***	0.889**	0.894*	0.594
	(0.390)	(0.341)	(0.455)	(0.392)
k nearest neighbors	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes
Route unc. cutoff (miles)	300	300		
Node unc. cutoff (miles)	5	10	5	10
Mean of dep var	35.20	34.64	35.54	36.59
Plants	125	112	128	85
Plant-county-months	36,704	31,364	36,999	23,902
Observations	76,897	66,101	75,175	50,675

Notes: This table is identical to Table 1.6.4, except that it uses alternative definitions for rail captiveness to construct M_j (see description in Appendix A.3). My preferred definition for node unconnecteness is 6.6 miles, which is the 95th percentile of plants' distance to nearest nodes. Columns (1) and (3) apply a 5-mile threshold for node unconnectedness, while Columns (2) and (4) apply a 10-mile threshold for node unconnectedness. Columns (3)–(4) define captiveness based on node (un)connectedness only, as in Column (6) of Table A.5.27. I report means of the dependent variable for plants with $M_j = 0$, and each regression pools both contract and spot shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.29: Markup Changes – Sensitivity to Commodity Controls

	Preferred commodity controls	Force avg orig price $\beta = 1$	Interact orig price $\times$ **1**[spot]	Interact orig price $\times$ sulfur	Interact sulfur $\times$ year FEs	Adding mine controls
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.187***	1.150***	1.173***	1.190***	1.212***	1.288***
	(0.408)	(0.410)	(0.425)	(0.409)	(0.410)	(0.405)
k nearest neighbors	3	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	34.20	34.20	34.20	34.20	34.20	34.42
Plants	124	124	124	124	124	124
Observations	75,089	75,089	75,089	75,089	75,089	74,620

Notes: This table is identical to Table 1.6.4, except that it conducts sensitivity on commodity controls ($\mathbf{C}_{ojms}$ in Equation (1.9)). Column (1) reproduces my preferred specification (Column (2) of Table 1.6.4), which uses the following uninteracted linear controls: BTU content, sulfur content, ash content, a dummy for spot shipments, a dummy for contracts expiring within 2 years, a dummy for bituminous coal, and the average annual mine-mouth price in each originating county. Column (2) forces the coefficient on average mine-mouth price to be 1. Column (3) allows separate coefficients on average mine-mouth price for spot vs. contract shipments. Column (4) interacts the average mine-mouth price with each shipment's average sulfur content. Column (5) allows year-specific coefficients on sulfur content, which accommodates changes in the shadow price of SO_2 emissions. Column (6) controls for time-varying characteristics in each coal mining county, weight-averaged by monthly coal production: mine age, seam thickness, seam depth, share of coal mined underground, share of mine employees working underground, and hours worked per ton of coal produced. I report means of the dependent variable for plants with $M_j = 0$, and each regression pools both contract and spot shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.30: Markup Changes – Sensitivity to Shipping Cost Controls

	4-way interaction (preferred)	Replace ln(quantity) w/ quantity	Replace diesel index w/ RCAF	Remove traffic density	5-way interaction w/ $\mathbf{1}$[sub-bituminous]
	(1)	(2)	(3)	(4)	(5)
$(\Delta \widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.187***	1.181***	1.196***	1.215***	1.131***
	(0.408)	(0.424)	(0.423)	(0.401)	(0.403)
k nearest neighbors	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes
Mean of dep var	34.20	34.20	34.20	34.20	34.20
Plants	124	124	124	124	124
Observations	75,089	75,089	75,089	75,089	75,089

Notes: This table is identical to Table 1.6.4, except that it conducts sensitivity on shipping cost controls ($S(\mathbf{T}_{ojms})$ in Equation (1.9)). Column (1) reproduces my preferred specification (Column (2) of Table 1.6.4), which uses a four-way linear interaction of: shorest-route shipping distance, AAR fuel price index, log of shipment size, and the proportion of each shortest route on rail lines with high traffic density. Column (2) replaces ln(shipment size) with shipment size in levels. Column (3) replaces the AAR diesel fuel price index with the quarterly Rail Cost Adjustment Factor. Column (4) removes controls for rail traffic density along each route, and uses only a 3-way interaction. Column (5) allows shipping costs to vary across a fifth interacted variable: a dummmy for sub-bituminous coal shipments. I report means of the dependent variable for plants with $M_j = 0$, and each regression pools both contract and spot shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.31: Markup Changes – Additional Sensitivities

	Preferred specification	Matching $\leq$ 100 miles	Pre-1980 plants	No shipment size weights	No nearest neighbor wts
	(1)	(2)	(3)	(4)	(5)
$(\Delta\,\widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	1.187***	1.023**	1.054**	1.589***	0.758**
	(0.408)	(0.426)	(0.434)	(0.462)	(0.318)
k nearest neighbors	3	3	3	3	3
Balanced panel	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes
Month × plant region FEs	Yes	Yes	Yes	Yes	Yes
Mean of dep var	34.20	37.72	34.00	43.73	35.50
Plants	124	106	103	124	187
Plant-county-months	35,651	30,582	31,180	35,995	57,366
Observations	75,089	65,609	66,143	76,391	114,342

Notes: This table is identical to Table 1.6.4, except that it conducts four sensitivities. Column (1) reproduces my preferred specification (Column (2) of Table 1.6.4). Column (2) restricts the sample of nearest-neighbor matches to be less than 100 miles apart. Column (3) restricts the sample to plants build prior to the 1980 Staggers Act, which loosened railroad price regulation. While my preferred specification weights by the product of nearest-neighbor weights and shipment size, Columns (4)-(5) each remove one of those weights. Column (4) uses nearest-neighbor weights but counts each observation equally, regardless of shipment size. Column (5) weights by shipment size but ignores nearest-neighbor weights (i.e., including filled and hollow plants represented in Figure 1.5.4). I report means of the dependent variable for plants with $M_j = 0$, and each regression pools both contract and spot shipments. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.5.4.3 Estimation Error in M_j

My estimates for markup changes rely on M_j, which I construct using $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$ and $\hat{\lambda}_{2j}$ estimates from my demand estimation algorithm. These means that measurement error due to estimation error in M_j may bias my empirical results. Classical measurement error of this right-hand-side variable should bias my point estimates of $\hat{\tau}$ towards zero, making it more difficult to detect differential changes in coal markups. I have no reason to suspect that measurement error in M_j is non-classical.

My preferred estimates of markup changes exclude three coal plants with extremely low or high values of M_j (specifically, two plants with $M_j > 2$ and one plant with $M_j < -2$). These outliers almost certainly reflect estimation error in $\hat{\lambda}_{0j}$, $\hat{\lambda}_{1j}$ and $\hat{\lambda}_{2j}$, as it is not plausible that a \$1/MMBTU change in gas prices will cause a change in coal markups greater than \$2/MMBTU—or \$40 per ton of coal.[96] If I reestimate Equation (1.9) including these three plants, my point estimates for $\hat{\tau}$ attenuate and lose statistical significance. This is not surprising, given that these M_j outliers have a strong influence over the (parametric) linear relationship between M_j and P_{ojms}. I report these results in Table A.5.32, which is otherwise identical to Table 1.6.4 from the main text.

Table A.5.32: Markup Changes – Sensitivity to Extremely High/Low M_j

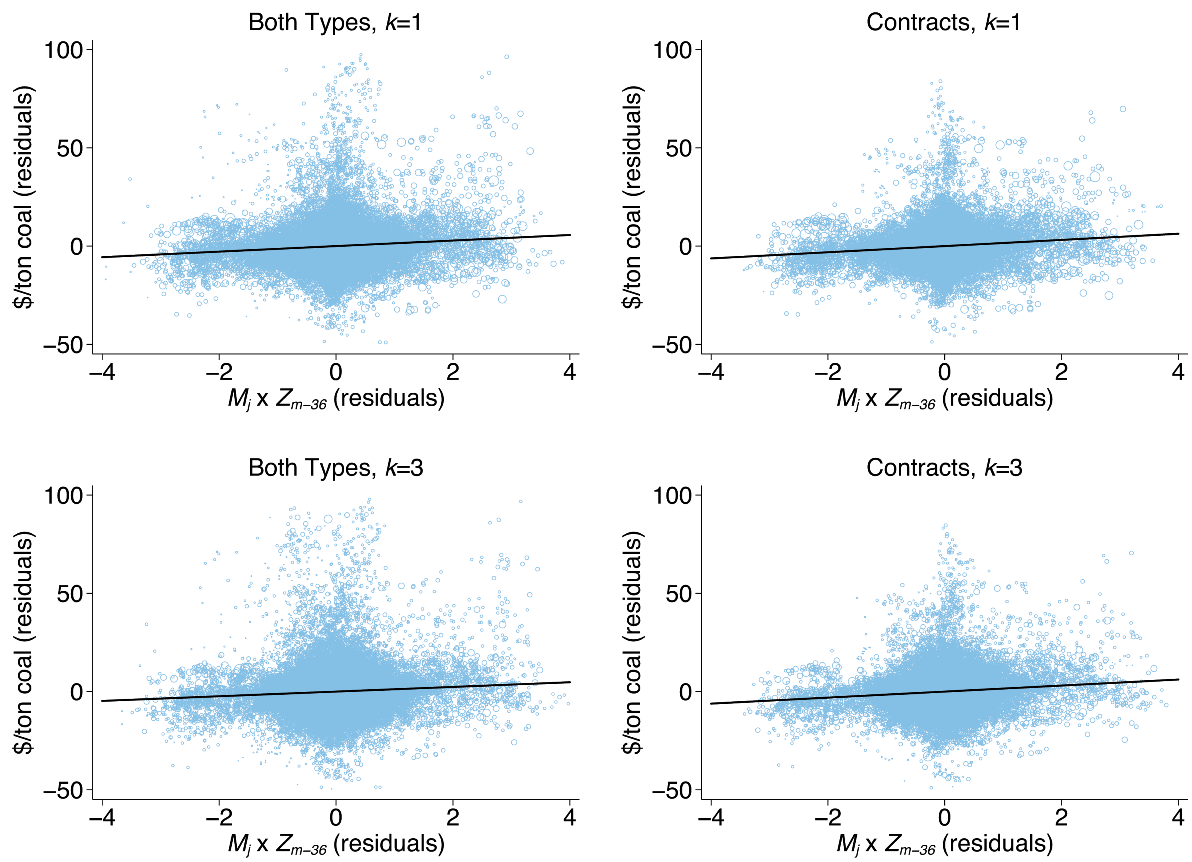
	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\Delta\ \widehat{\text{Markup}})_j \times (\text{Gas Price})_m$	0.413	0.482	0.430	0.584*	0.840	0.321
	(0.442)	(0.349)	(0.437)	(0.348)	(0.808)	(0.699)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant $\times$ county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	33.27	34.20	32.07	33.04	41.76	41.39
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: This table is identical to Table 1.6.4 from the main text, except that it includes 3 plants with extreme values of M_j (i.e. $|M_j| > 2$). Balanced panels include plants receiving at least 1 shipment in each sample year (2002–2015). I report means of the dependent variable for plants with $M_j = 0$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[96] These three plants are top-coded and bottom-coded in the bottom-right histogram of Figure 1.5.6.

Figures A.5.31 and A.5.32 illustrate how including these three outlier plants attenuates $\hat{\tau}$. Residual plots in Figure A.5.31 correspond to Columns (1)–(4) in Table 1.6.4, excluding the three outlier plants; black regression lines plot my point estimates for $\hat{\tau}$, which are each positive and statistically significant. Residual plots in Figure A.5.32 correspond to Columns (1)–(4) in Table A.5.32, including the three outlier plants; these residuals are color-coded such that blue circles denote observations that appear in Figure A.5.31, while orange circles denote observations for the three outlier plants (less than 2 percent of total weighted observations in each regression).[97] Including plants with extreme values of M_j doubles the support of right-hand-side residuals, and the horizontal axes in Figure A.5.32 extend from -8 to 8, rather than from -4 to 4.

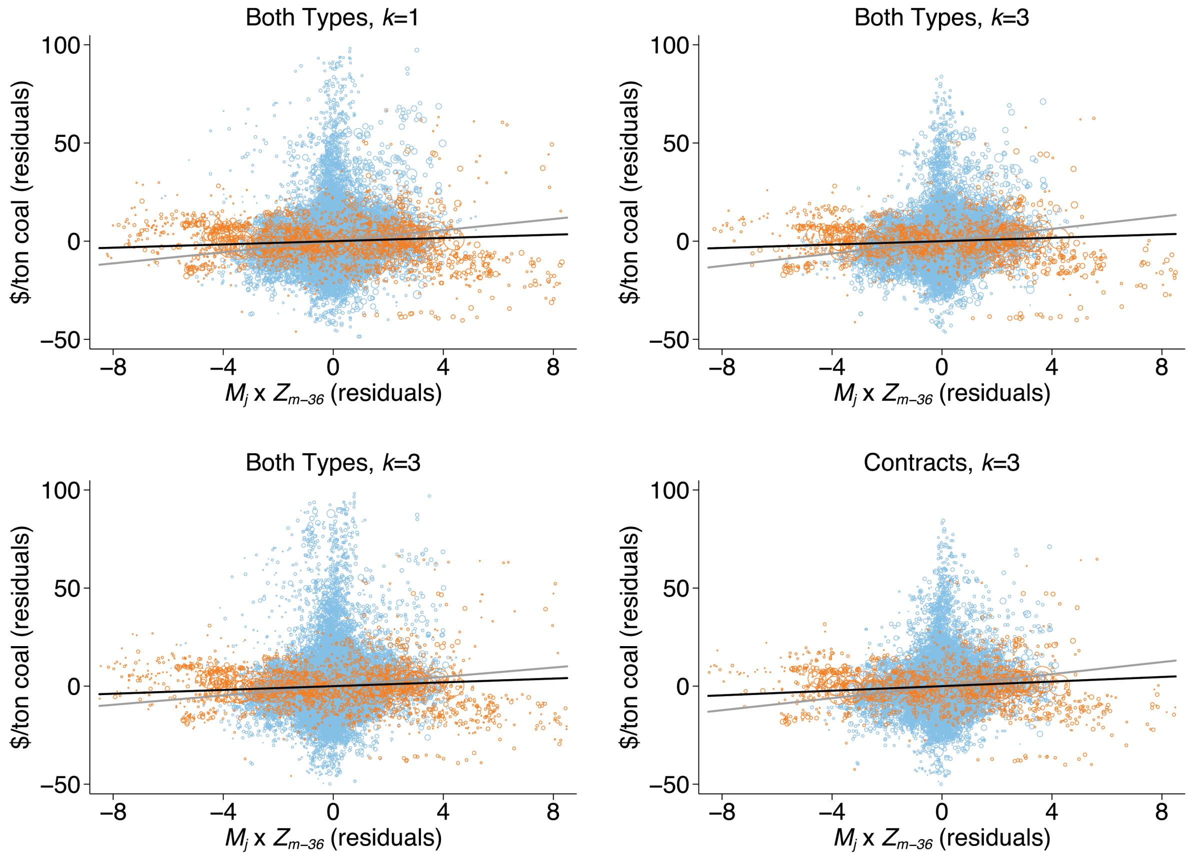
Figure A.5.31: Residual Plots, Main Specification (Table 1.6.4)



Notes: This figure presents residual plots for Columns (1)–(4) in Table 1.6.4. Blue circles plot residuals from P_{ojms} and $M_j \times Z^{HH}_{m-36}$, after partialing out all right-hand-side variables in Equation (1.9) *except* $M_j \times Z^{HH}_{m-36}$. Black lines plot the regression coefficients from Table 1.6.4. Each circle's size scales with its regression weight.

[97]The blue circles in Figures A.5.31 and A.5.32 are not identical, as they are residuals from different Frisch-Waugh regressions. They represent an identical sample of *observations* across two sets of regressions.

Figure A.5.32: Residual Plots, Including Extremely High/Low M_j (Table A.5.32)



Notes: This figure presents residual plots for Columns (1)–(4) in Table A.5.32. Blue and orange circles plot residuals from P_{ojms} and $M_j \times Z^{HH}_{m-36}$, after partialing out all right-hand-side variables in Equation (1.9) *except* $M_j \times Z^{HH}_{m-36}$. Blue circles denote observations for plants with $M_j \in [-2, 2]$, or the same sample of observations that appear in Table 1.6.4 and Figure A.5.31. Orange circles denote observations for the 3 plants with $|M_j| > 2$, which are included in Table A.5.32 only, representing less than 2 percent of total observations. Gray lines plot regression coefficients from Table 1.6.4 and Figure A.5.32. Black lines plot regression coefficients from Table A.5.32, illustrating how a few extreme values of M_j pull the slopes of the regression lines toward zero. Each circle's size scales with its regression weight.

Figure A.5.32 plots two sets of regression lines. Gray lines exclude the three outlier plants, and are identical to the regression lines in Figure A.5.31; black lines include the three outlier plants, and plot the coefficients in Table A.5.32. This illustrates how extreme, high-leverage values of the orange residuals pull the slopes of the gray regression lines towards zero.

I discretize the support of M_j into five indicator variables in Table 1.6.6 of the main text, in order to recover more easily interpretable point estimates. This strategy also has the advantage of releasing the parametric assumptions on M_j that attenuate my point estimates in Table A.5.32 and Figue A.5.32. I apply the same strategy to the sample of plants including extreme values of M_j, and I report these results in Table A.5.33. While these point estimates are attenuated, they retain (weak) statistical significance for the 4th and 5th quintiles. In other words, by treating plants with $M_j = 5$ and $M_j = 0.8$ equally (using an indicator variable), I reduce the leverage of extreme values of M_j and recover point estimates that are consistent with my main results (albeit attenuated due to additional measurement error). Table A.5.34 repeats this exercise with 48 lags (rather than 36 lags), finding slightly larger (and more statistically significant) cumulative effects including extreme values of M_j. Table A.5.35 adds two outer bins for extreme values of M_j, which further increases the magnitude and significance of my estimated cumulative effects with 48 lags.

Table A.5.33: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, Extremely Low/High M_j

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{Gas Price})_m$	0.021	−0.115	0.254	0.186	−0.870*	−1.321**
	(0.220)	(0.186)	(0.212)	(0.172)	(0.489)	(0.607)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{Gas Price})_m$	0.132	−0.035	0.378*	0.285	−0.396	−1.017*
	(0.230)	(0.200)	(0.226)	(0.197)	(0.442)	(0.534)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{Gas Price})_m$	0.462*	0.268	0.486*	0.362*	0.819	0.186
	(0.276)	(0.238)	(0.250)	(0.199)	(0.810)	(0.887)
$\mathbf{1}\Big[M_j \in (0.70, 5.07]\Big] \times (\text{Gas Price})_m$	0.781*	0.585*	0.923**	0.819***	1.035	−0.017
	(0.403)	(0.298)	(0.377)	(0.279)	(1.055)	(0.928)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	30.47	34.35	29.70	34.04	33.55	35.54
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: This table is identical to Table 1.6.6, except that each regression includes plants with very small/large values of M_j (i.e. $|M_j| > 2$). I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.5.34: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, Low/High M_j, 48 Lags

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{Gas Price})_m$	0.082	−0.032	0.206	0.097	0.302	0.117
	(0.261)	(0.209)	(0.226)	(0.185)	(0.251)	(0.275)
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{Gas Price})_m$	−0.121	−0.216	0.082	−0.012	0.243	−0.236
	(0.247)	(0.217)	(0.227)	(0.204)	(0.418)	(0.424)
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{Gas Price})_m$	0.581*	0.443	0.676**	0.552*	1.190**	0.573
	(0.342)	(0.305)	(0.333)	(0.296)	(0.496)	(0.571)
$\mathbf{1}\Big[M_j \in (0.70, 5.07]\Big] \times (\text{Gas Price})_m$	0.908*	0.674*	1.079**	0.856**	1.017	0.104
	(0.461)	(0.350)	(0.448)	(0.345)	(0.877)	(0.727)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	30.47	34.35	29.70	34.04	33.55	35.54
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: This table is identical to Table A.5.33, except that each regression estimates cumulative effects over 48 lags. Regressions include plants with very small/large values of M_j (i.e. $|M_j| > 2$). I report means of the dependent variable for the omitted group of plants, with $M_j \leq 0.22$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

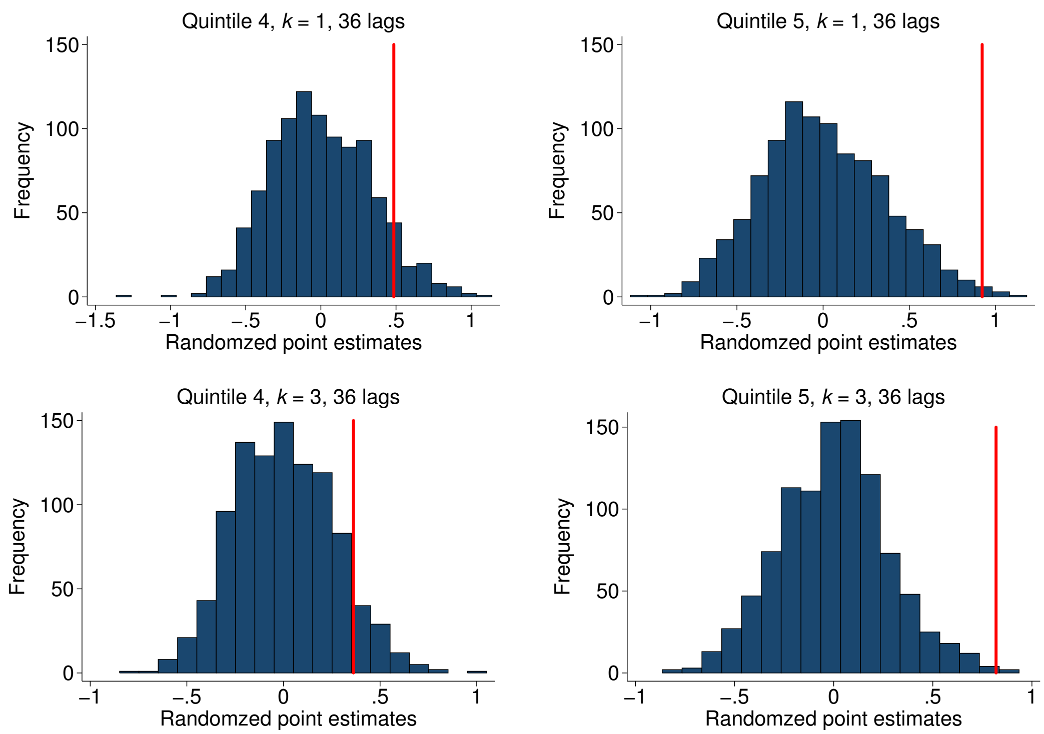
Table A.5.35: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, 48 Lags, Outer M_j Bins

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\left[M_j \in [-2.72, -2.00]\right] \times (\text{Gas Price})_m$	-0.555	-0.592	-0.402	-0.677	0.000	0.000
	(0.634)	(0.443)	(0.700)	(0.459)	(0.000)	(0.000)
$\mathbf{1}\left[M_j \in (0.22, 0.35]\right] \times (\text{Gas Price})_m$	0.081	-0.018	0.211	0.115	0.276	0.084
	(0.263)	(0.211)	(0.228)	(0.185)	(0.269)	(0.287)
$\mathbf{1}\left[M_j \in (0.35, 0.52]\right] \times (\text{Gas Price})_m$	-0.117	-0.192	0.094	0.020	0.215	-0.276
	(0.253)	(0.220)	(0.230)	(0.205)	(0.441)	(0.440)
$\mathbf{1}\left[M_j \in (0.52, 0.70]\right] \times (\text{Gas Price})_m$	0.602^{*}	0.489	0.707^{**}	0.611^{**}	1.151^{**}	0.517
	(0.361)	(0.316)	(0.346)	(0.301)	(0.531)	(0.599)
$\mathbf{1}\left[M_j \in (0.70, 2.00]\right] \times (\text{Gas Price})_m$	0.950^{*}	0.727^{*}	1.114^{**}	0.924^{**}	1.018	0.098
	(0.498)	(0.373)	(0.487)	(0.365)	(0.871)	(0.731)
$\mathbf{1}\left[M_j \in (2.00, 5.07]\right] \times (\text{Gas Price})_m$	1.167	1.310	1.588	1.760^{**}	0.496	-0.990
	(1.256)	(0.840)	(1.257)	(0.826)	(1.503)	(1.549)
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	30.47	34.35	29.70	34.04	33.55	35.54
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: This table is identical to Table A.5.34, except that it includes extra bins for plants with very small/large values of M_j. Each regression reports cumulative effects through 48 months. I report means of the dependent variable for the omitted group of plants, with $M_j \in (-2.00, 0.22]$. Standard errors are clustered by plant. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

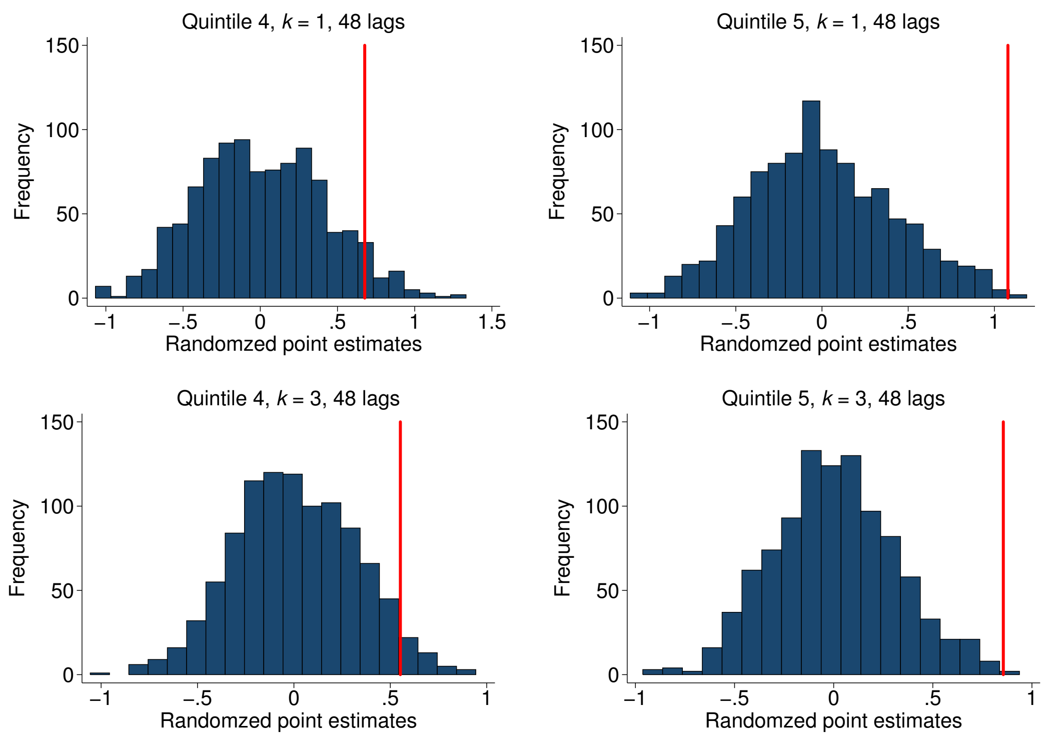
I conduct a randomization inference test to estimate the likelihood of recovering the same coefficient estimates if M_j were randomly assigned to plants. I scramble the vector of M_j values across all plants in the balanced sample, reestimate Columns (3)–(4) of Tables A.5.33–A.5.34 given these randomized values of M_j, and iterate these steps 1,000 times. Figures A.5.33–A.5.34 plot the resulting histograms of randomized coefficients for quintiles 4–5 (i.e. $M_j \in (0.52, 0.70]$ and $M_j \in (0.70.5.07]$). Red lines denote my actual point estimates, for which M_j is correctly assigned to plant j. For all four quintile 5 (4) coefficients, the actual point estimates fall above the 99th (90th) percentiles of their respective randomized distributions. This shows that it would be extremely unlikely to recover statistically significant differential changes in markups by chance, and that M_j carries a high degree of predictive power in spite of measurement error.

Figure A.5.33: Randomization Test – Contract Shipments, 36 Lags, Quintiles 4–5 of M_j



Notes: This figure presents the results of a randomization test, where I reestimate Columns (3)–(4) in Table A.5.33 after randomly scrambling the vector of M_j values across plants. I perform 1,000 independent draws, where each iteration randomly assigns 190 values of M_j to the 190 utility-owned coal plants with coal deliveries in each sample year (i.e. the 190 plants out of the 324 plants appearing in Figure 1.5.4, both filled and hollow, that do not retire or go idle). Each iteration then estimates the identical regressions from Columns (3)–(4), applying non-scrambled nearest-neighbor weights and including either 94 of 124 coal plants, and storing the point estimates for quintiles 4 and 5 (i.e. $M_j \in (0.52, 0.70]$ and $M_j \in (0.70.5.07]$, respectively). Red lines mark actual (non-randomized) point estimates from Table A.5.33 (i.e. 0.486, 0.923, 0.362, 0.819). Point estimates in the left column both fall above the 90th percentiles of their respective randomized distributions. Point estimates in the both fall above the 99th percentiles of their respective randomized distributions.

Figure A.5.34: Randomization Test – Contract Shipments, 48 Lags, Quintiles 4–5 of M_j



Notes: This figure presents the results of a randomization test, where I reestimate Columns (3)–(4) in Table A.5.34 after randomly scrambling the vector of M_j values across plants. I perform 1,000 independent draws, where each iteration randomly assigns 190 values of M_j to the 190 utility-owned coal plants with coal deliveries in each sample year (i.e. the 190 plants out of the 324 plants appearing in Figure 1.5.4, both filled and hollow, that do not retire or go idle). Each iteration then estimates the identical regressions from Columns (3)–(4), applying non-scrambled nearest-neighbor weights and including either 94 of 124 coal plants, and storing the point estimates for quintiles 4 and 5 (i.e. $M_j \in (0.52, 0.70]$ and $M_j \in (0.70.5.07]$, respectively). Red lines mark actual (non-randomized) point estimates from Table A.5.34 (i.e. 0.676, 1.079, 0.552, 0.856). Both point estimates in the left column are at the 95th percentiles of their respective randomized distributions. Point estimates in the right column both both fall above the 99th percentiles of their respective randomized distributions.

Finally, I estimate two sets of alternative standard errors for Table A.5.33, in order to account for estimation error in M_j. First, I block-bootstrap coal plants with replacement, using 100 bootstrap iterations. Second, I estimate Equation (1.9) by simulating random draws of M_j for each plant. I construct simulated distributions of M_j by iterating all post-estimation steps of my demand estimation procedure, introducing random noise to account for the prediction error from estimating $\tilde{P}_{uh}$. I report these alternative standard errors in Table A.5.36: parentheses denote standard errors clustered by plant (identical to Table A.5.33); square brackets denote standard errors that are block-bootstrapped by plant, treating M_j as fixed; curly braces denote bootstrapped standard errors from simulating M_j independently for each plant, for each bootstrap iteration. These three sets of standard errors are quite close, especially for contract shipments in quintiles 4 and 5. My results largely retain statistical significance, even after accounting for estimation error in M_j.

To be precise, I construct the standard errors in curly braces using the following steps:

1. For 100 simulations, indexed by S:
   a) Take one independent random normal draw for each plant j, in each month m.
   b) Inflate each random normal draw by the *hourly* mean and standard error of each $\tilde{P}_{uh}$ estimate, from Step 3 of my demand estimation procedure. This introduces a correlated random disturbance into $\tilde{P}_{uh}$ for each unit u within plant j, and for each hour h within month m (consistent with having clustered by month in Equation (A.29)). It also scales the disturbance for each unit-hour by the standard error from the nonlinear prediction of each $\tilde{P}_{uh}$. Let $\tilde{P}^S_{uh}$ denote the simulate values of $\tilde{P}_{uh}$.
   c) Iterate Steps 4 and 5 of my demand estimation procedure for the vector of simulated $\tilde{P}^S_{uh}$'s, to arrive a simulated values of $\hat{\lambda}^S_{0j}$, $\hat{\lambda}^S_{1j}$, and $\hat{\lambda}^S_{2j}$.
2. For 1000 bootstrap iterations, indexed by B:
   a) Construct a $[J \times 1]$ vector of bootstrapped M_j values (denoted M^B_j), by independently drawing triples $\{\hat{\lambda}^S_{0j}, \hat{\lambda}^S_{1j}, \hat{\lambda}^S_{2j}\}$ from the simulated distributions of each plant j. (S is the same for each simulated $\hat{\lambda}^S$ within each plant j, but S is independent across plants.)
   b) Estimate each regression in Table A.5.33, replacing M_j with M^B_j.
   c) Store the resulting $\hat{\tau}^B$ point estimates for quintiles 2–5.
3. Assign standard errors equal to the standard deviations of the distributions of stored $\hat{\tau}^B$ estimates.

Table A.5.36: Markup Changes – Quantiles of $\Delta\,\widehat{\text{Markup}}$, Alternative Standard Errors

	Both Types		Contracts		Spot Market	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\Big[M_j \in (0.22, 0.35]\Big] \times (\text{Gas Price})_m$	(0.220)	(0.186)	(0.212)	(0.172)	$(0.489)^{*}$	$(0.607)^{**}$
	[0.253]	[0.194]	[0.230]	[0.189]	[0.550]	$[0.712]^{*}$
	{0.217}	{0.200}	{0.188}	{0.168}	{0.914}	{0.977}
$\mathbf{1}\Big[M_j \in (0.35, 0.52]\Big] \times (\text{Gas Price})_m$	(0.230)	(0.200)	$(0.226)^{*}$	(0.197)	(0.442)	$(0.534)^{*}$
	[0.247]	[0.223]	[0.263]	[0.232]	[0.482]	[0.770]
	{0.173}	{0.158}	$\{0.149\}^{**}$	$\{0.134\}^{**}$	{0.702}	{0.750}
$\mathbf{1}\Big[M_j \in (0.52, 0.70]\Big] \times (\text{Gas Price})_m$	$(0.276)^{*}$	(0.238)	$(0.250)^{*}$	$(0.199)^{*}$	(0.810)	(0.887)
	[0.318]	[0.305]	$[0.268]^{*}$	[0.263]	[0.819]	[0.928]
	$\{0.211\}^{**}$	{0.205}	$\{0.183\}^{***}$	$\{0.176\}^{**}$	{0.612}	{0.673}
$\mathbf{1}\Big[M_j \in (0.70, 5.07]\Big] \times (\text{Gas Price})_m$	$(0.403)^{*}$	$(0.298)^{*}$	$(0.377)^{**}$	$(0.279)^{***}$	(1.055)	(0.928)
	$[0.416]^{*}$	[0.370]	$[0.430]^{**}$	$[0.437]^{*}$	[1.041]	[1.219]
	$\{0.211\}^{***}$	$\{0.210\}^{***}$	$\{0.204\}^{***}$	$\{0.200\}^{***}$	{0.624}	{0.671}
k nearest neighbors	1	3	1	3	1	3
Balanced panel	Yes	Yes	Yes	Yes	Yes	Yes
Plant × county FEs	Yes	Yes	Yes	Yes	Yes	Yes
Month-of-sample FEs	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep var	30.47	34.35	29.70	34.04	33.55	35.54
Plants	97	127	95	125	87	117
Plant-county-months	27,392	36,980	23,010	30,813	7,272	10,105
Observations	58,062	76,927	45,883	60,406	12,090	16,399

Notes: This table ireports alternative standard errors for each regression in Table A.5.33. The first row (in parentheses) are my preferred standard errors, clustered by plant, and as reported in Table A.5.33. The second row [in brackets] are standard errors that are block-bootstrapped by plant, with 100 bootstrap iterations. The third row {in braces} are bootstrapped standard errors where each bootstrap iteration independently draws a new M_j value from each plant's simulated distribution of M_j values. I construct simulated distributions of M_j using the standard errors from my demand estimation, which I describe in Appendix A.4. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.5.5 Pass-Through Estimates

In Table 1.7.7 in the main text, I report implied carbon tax pass-through rates based on my derivations from Appendix A.1.2 and my point estimates from Table A.5.19, for a balanced panel of coal plants and $k = 3$ nearest-neighbor matches. Here, I report three additional sets of pass-through estimates: (i) Table A.5.37 uses point estimates from a balanced panel with $k = 1$ nearest neighbors (from Table A.5.19); (ii) Table A.5.38 uses point estimates from an unbalanced panel with $k = 3$ nearest nearest neighbors (from Table A.5.20); and (iii) Table A.5.39 uses point estimates from an unbalanced panel with $k = 1$ nearest neighbors (from Table A.5.20). Pass-through rates are broadly consistent across all four sets of estimates.

Table A.5.37: Pass-Through of Implied Carbon Tax ($k = 1$ Nearest Neighbors)

	Plant Group (Quintile of M_j)				
	(1)	(2)	(3)	(4)	(5)
$\Delta\,\widehat{\text{Markup}}$ (M_j)	$(-2.00, 0.22]$	$(0.22, 0.35]$	$(0.35, 0.52]$	$(0.52, 0.70]$	$(0.70, 2.00]$
Share of plants	0.33	0.15	0.13	0.17	0.15
A. Fuel prices only					
ρ_j, all shipments	1.00	**0.99**	**0.94**	**0.87**	**0.76**
		[0.98, 1.00]	[0.84, 0.97]	[0.77, 0.93]	[0.27, 0.87]
		(0.87, 1.11)	*(0.85, 1.04)*	*(0.78, 0.97)*	*(0.60, 0.93)*
ρ_j, contracts only	1.00	**0.93**	**0.89**	**0.86**	**0.73**
		[0.84, 0.96]	[0.67, 0.94]	[0.75, 0.92]	[0.18, 0.85]
		(0.81, 1.05)	*(0.79, 0.98)*	*(0.77, 0.95)*	*(0.57, 0.88)*
B. Fuel + environmental costs					
ρ_j, all shipments	1.00	**0.99**	**0.95**	**0.89**	**0.79**
		[0.98, 1.00]	[0.86, 0.98]	[0.79, 0.95]	[0.39, 0.89]
		(0.88, 1.10)	*(0.87, 1.03)*	*(0.80, 0.97)*	*(0.65, 0.94)*
ρ_j, contracts only	1.00	**0.94**	**0.90**	**0.88**	**0.76**
		[0.86, 0.97]	[0.72, 0.95]	[0.77, 0.95]	[0.31, 0.87]
		(0.83, 1.04)	*(0.81, 0.99)*	*(0.80, 0.95)*	*(0.63, 0.90)*

Notes: This table is identical to Table 1.7.7 from the main text, except that I convert results from Table A.5.19 using $k = 1$ nearest neighbors. Average pass-through rates are in bold, and square brackets report the minimum and maximum pass-through rates for plants in each group. I report the 95 percent confidence intervals for the *average* (bolded) pass-through rates in parentheses and italics (calculated from the confidence interval of each $\hat{\tau}$ estimate). Whereas Panel A follows Equation (1.11) by only including fuel prices, Panel B includes environmental costs following Equation (A.21).

Table A.5.38: Pass-Through of Implied Carbon Tax ($k = 3$, Unbalanced Panel)

	Plant Group (Quintile of M_j)				
	(1)	(2)	(3)	(4)	(5)
$\Delta\,\widehat{\text{Markup}}$ (M_j)	$(-2.00, 0.22]$	$(0.22, 0.35]$	$(0.35, 0.52]$	$(0.52, 0.70]$	$(0.70, 2.00]$
Share of plants	0.43	0.15	0.14	0.13	0.16
A. Fuel prices only					
ρ_j, all shipments	1.00	**1.02**	**0.98**	**0.92**	**0.84**
		[1.01, 1.05]	[0.95, 0.99]	[0.85, 0.95]	[0.50, 0.91]
		(0.92, 1.11)	*(0.91, 1.06)*	*(0.84, 1.00)*	*(0.71, 0.96)*
ρ_j, contracts only	1.00	**0.95**	**0.92**	**0.90**	**0.78**
		[0.88, 0.98]	[0.76, 0.96]	[0.81, 0.94]	[0.35, 0.88]
		(0.86, 1.04)	*(0.84, 1.00)*	*(0.83, 0.97)*	*(0.66, 0.90)*
B. Fuel + environmental costs					
ρ_j, all shipments	1.00	**1.02**	**0.99**	**0.93**	**0.86**
		[1.01, 1.04]	[0.96, 0.99]	[0.87, 0.97]	[0.58, 0.92]
		(0.93, 1.10)	*(0.92, 1.06)*	*(0.86, 1.00)*	*(0.75, 0.96)*
ρ_j, contracts only	1.00	**0.96**	**0.93**	**0.91**	**0.81**
		[0.90, 0.99]	[0.80, 0.96]	[0.83, 0.96]	[0.46, 0.90]
		(0.88, 1.04)	*(0.86, 1.00)*	*(0.85, 0.98)*	*(0.70, 0.92)*

Notes: This table is identical to Table 1.7.7 from the main text, except that I convert results from Table A.5.20 using $k = 3$ nearest neighbors and a full (unbalanced) panel of coal plants. Average pass-through rates are in bold, and square brackets report the minimum and maximum pass-through rates for plants in each group. I report the 95 percent confidence intervals for the *average* (bolded) pass-through rates in parentheses and italics (calculated from the confidence interval of each $\hat{\tau}$ estimate). Whereas Panel A follows Equation (1.11) by only including fuel prices, Panel B includes environmental costs following Equation (A.21).

Table A.5.39: Pass-Through of Implied Carbon Tax ($k = 1$, Unbalanced Panel)

	Plant Group (Quintile of M_j)				
	(1)	(2)	(3)	(4)	(5)
$\Delta\,\widehat{\text{Markup}}$ (M_j)	$(-2.00, 0.22]$	$(0.22, 0.35]$	$(0.35, 0.52]$	$(0.52, 0.70]$	$(0.70, 2.00]$
Share of plants	0.38	0.17	0.12	0.16	0.15
A. Fuel prices only					
ρ_j, all shipments	1.00	**1.00**	**0.96**	**0.89**	**0.78**
		[1.00, 1.00]	[0.86, 0.98]	[0.79, 0.93]	[0.35, 0.89]
		(0.90, 1.11)	*(0.87, 1.04)*	*(0.80, 0.98)*	*(0.62, 0.95)*
ρ_j, contracts only	1.00	**0.95**	**0.91**	**0.88**	**0.74**
		[0.87, 0.98]	[0.71, 0.95]	[0.77, 0.92]	[0.25, 0.86]
		(0.85, 1.05)	*(0.82, 1.00)*	*(0.79, 0.96)*	*(0.58, 0.91)*
B. Fuel + environmental costs					
ρ_j, all shipments	1.00	**1.00**	**0.96**	**0.90**	**0.81**
		[1.00, 1.00]	[0.88, 0.98]	[0.82, 0.95]	[0.45, 0.90]
		(0.91, 1.09)	*(0.89, 1.04)*	*(0.82, 0.99)*	*(0.66, 0.96)*
ρ_j, contracts only	1.00	**0.96**	**0.92**	**0.89**	**0.77**
		[0.89, 0.98]	[0.75, 0.96]	[0.80, 0.95]	[0.37, 0.88]
		(0.87, 1.05)	*(0.84, 1.00)*	*(0.81, 0.97)*	*(0.63, 0.92)*

Notes: This table is identical to Table 1.7.7 from the main text, except that I convert results from Table A.5.20 using $k = 1$ nearest neighbors and a full (unbalanced) panel of coal plants. Average pass-through rates are in bold, and square brackets report the minimum and maximum pass-through rates for plants in each group. I report the 95 percent confidence intervals for the *average* (bolded) pass-through rates in parentheses and italics (calculated from the confidence interval of each $\hat{\tau}$ estimate). Whereas Panel A follows Equation (1.11) by only including fuel prices, Panel B includes environmental costs following Equation (A.21).

A.6 CO_2 Emissions Counterfactuals

This section describes my method for estimating CO_2 emissions counterfactuals, which I report in Section 1.7.4 of the main text. I begin by estimating the relationship between relative fuel costs and CO_2 emissions, for each CEMS coal generating unit. Then, I use each plant's fitted model to estimate generation under two counterfactual fuel cost ratios. First, I consider a scenario where the fracking boom did not happen, and gas prices remained high. Second, I consider a scenario where the fracking boom did happen, but coal markups remained fixed (i.e., full pass-through of relative fuel price shocks). In each counterfactual scenario, coal markups did not decrease, meaning that counterfactual coal prices would have been slightly higher than observed coal prices. Comparing across these counterfactual estimates, I am able to calculate CO_2 abatement due to short-run coal-to-gas substitution, both with and without changes to coal markups.

A.6.1 Counterfactuals Algorithm

Step 1: I construct a time-series of counterfactual gas prices in the absence of the fracking boom, using historic NYMEX monthly futures prices as of December 2008. This follows Holladay and LaRiviere (2017), who estimate a structural break in the time-series of Henry Hub gas spot prices on December 5, 2008. The top panel of Figure A.6.35 compares these futures prices to actual natural gas prices. Visually, these prices are close to pre-Recession levels, and would be roughly constant if I removed seasonal variation.[98] Given that natural gas plants also pay pipeline and distribution charges on top of the commodity spot price, I add the average difference between Henry Hub prices and delivered prices to construct counterfactual prices for each gas plant g on each day d, or Z_{gd}^{CF}.[99]

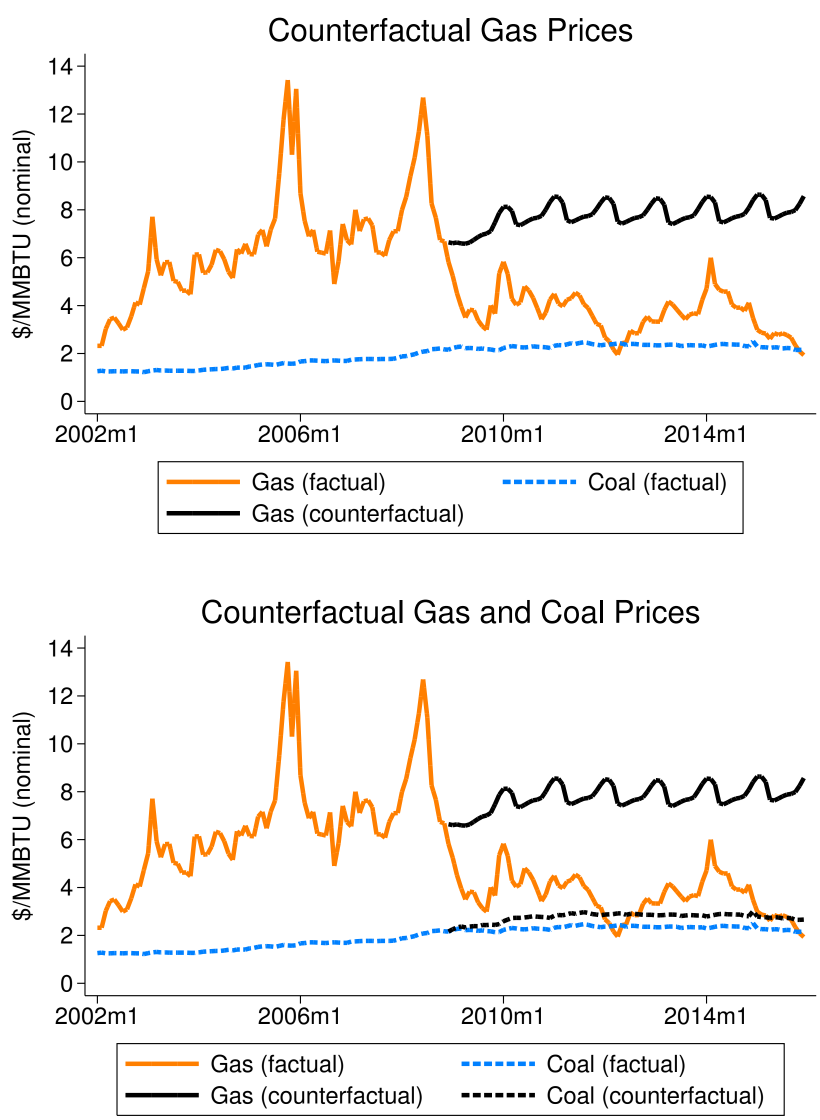
Step 2: Using these counterfactual gas prices, I predict counterfactual coal prices if markups had not changed. My empirical analysis demonstrates that the fracking boom caused coal prices to fall; hence, counterfactual coal prices in the absence of the fracking boom would have been higher. The bottom panel of Figure A.6.35 illustrates what average coal prices may have looked like if the fracking boom did not happen, or if coal markets had not changed.

I estimate Equation (1.9) using indicator variables for quintiles of M_j (as in Table 1.6.6), and converting the dependent variable (P_{ojms}) from \$/ton to \$/MMBTU (i.e., the relevant units for coal factor demand). Then, I predict counterfactual coal prices using this fitted regression model, replacing factual with counterfactual gas prices for all lagged difference-in-differences coefficients. I denote these counterfactual coal prices as P_{jm}^{CF}, averaging across all of plant j's shipments in month m.

[98] U.S. natural gas prices tend to rise in the winter, due to increased demand for space heating. I include this expected seasonal variation in counterfactual gas prices. Importantly, my counterfactual analysis begins *after* the Recession-related price spike.

[99] See Appendix A.2.5 for further detail on my method for converting Henry Hub prices to plant-specific delivered gas prices.

Figure A.6.35: Counterfactual Fuel Prices



Notes: The top panel shows factual gas and coal prices (as in Figure 1.2.2), and counterfactual gas prices in the absence of the fracking boom. These counterfactual gas prices are Henry Hub future prices as of December 2008, which reflected market expectations of gas prices in the month before the historic drop in gas prices. The bottom panel adds illustrative coal price counterfactuals, as coal markups would not have fallen if the fracking boom had not happened.

While my main analysis includes only a subset of coal plants with nearest-neighbor matches, I now want to calculate CO_2 abatement by summing across *all* coal plants. I regress counterfactual coal prices on the interaction of actual delivered coal prices (P_{ojms}) and predicted markup changes (M_j), estimating a separate coefficient for each sample month. I also include commodity controls ($\mathbf{C}_{ojms}$), shipment controls ($\mathbf{T}_{ojms}$), plant controls ($\mathbf{X}_{jm}$), coal county fixed effects (η_o), month fixed effects (δ_m), as in Equation (1.5). Taking predicted values from this regression, I am able to populate P^{CF}_{jm} for the 55 percent of plants that do not appear in my main regression analysis. For non-rail plants and non-utility plants (with withheld prices), I set $P^{CF}_{jm} = \bar{P}_{jm}$, which assumes no fracking-induced markup changes.[100]

Step 3: I construct factual and counterfactual coal-to-gas cost ratios via Step 1 of my demand estimation algorithm (see Appendix A.4.1 above). However, I now average *both* coal and gas marginal costs across all generating units of each fuel type, within each PCA, to consider a counterfactual where *many* plants' coal prices change:

$$MC^{coal}_{um} \equiv \sum_{j \in \mathrm{PCA}_u} \left(\frac{Q^{elec}_{jm} \cdot HR_{jm} \cdot (P_{jm} + MC^{env}_{jm})}{\sum_{j \in \mathrm{PCA}_u} Q^{elec}_{jm}} \right) \tag{A.40}$$

$$MC^{gas}_{ud} \equiv \sum_{g \in \mathrm{PCA}_u} \left(\frac{Q^{elec}_{gm} \cdot HR_{gm} \cdot (Z_{gd} + MC^{env}_{gm})}{\sum_{g \in \mathrm{PCA}_u} Q^{elec}_{gm}} \right) \tag{A.41}$$

Replacing factual with counterfactual prices:

$$MC^{coal,CF}_{um} \equiv \sum_{j \in \mathrm{PCA}_u} \left(\frac{Q^{elec}_{jm} \cdot HR_{jm} \cdot (P^{CF}_{jm} + MC^{env}_{jm})}{\sum_{j \in \mathrm{PCA}_u} Q^{elec}_{jm}} \right) \tag{A.42}$$

$$MC^{gas,CF}_{ud} \equiv \sum_{g \in \mathrm{PCA}_u} \left(\frac{Q^{elec}_{gm} \cdot HR_{gm} \cdot (Z^{CF}_{gd} + MC^{env}_{gm})}{\sum_{g \in \mathrm{PCA}_u} Q^{elec}_{gm}} \right) \tag{A.43}$$

Then, I construct three cost ratios:

$$CR_{ud} = \frac{MC^{coal}_{um}}{MC^{gas}_{ud}} \quad , \qquad CR^{NO\Delta Z}_{ud} = \frac{MC^{coal,CF}_{um}}{MC^{gas,CF}_{ud}} \quad , \qquad CR^{NO\Delta\mu}_{ud} = \frac{MC^{coal,CF}_{um}}{MC^{gas}_{ud}} \tag{A.44}$$

[100] Table 1.6.6 uses a balanced sample of coal plants, to account the potentially confounding effects of coal plant retirements. This 55 percent of plants includes both plants without nearest-neighbor matches and plants dropped from Table 1.6.6 for not appearing in each sample year. If coal markups similarly decreased for non-utility plants during the fracking boom, then this counterfactual exercise would understate the full extent to which reductions in coal markups eroded the potential environmental benefits of the fracking boom. I use $\bar{P}_{jm}$ to denote the BTU-weighted average delivered price (for non-rail plants with reported prices), and average delivered coal prices at the state-month level (for non-utility plants, which EIA publishes in its Electric Power Monthly reports). These (mis-measured) aggregate prices are more suited for estimating that conditional probability of a coal unit operating in a given hour, than they would be for estimating coal markups at the shipment level.

CR_{ud} uses factual gas prices and factual coal prices. $CR_{ud}^{NO\Delta Z}$ assumes that the fracking boom did not happen, using both counterfactual gas prices and counterfactual coal prices. Given that the difference between P_{jm} and P_{jm}^{CF} is the response to changes in gas prices, P_{jm}^{CF} are the appropriate coal prices for the "no-fracking" counterfactual. Finally, $CR_{ud}^{NO\Delta\mu}$ assumes that the fracking boom did happen, but coal markups did not adjust to changes in gas price. This "full-pass-through" counterfactual uses factual gas prices but counterfactual coal prices.

Step 4: For each coal unit u, I estimate the following time-series regression, for each day d, from 2002 to 2015:

(A.45)
$$MWH_{ud} = \sum_b \alpha_{ub}\mathbf{1}[G_{ud} \in b] + \sum_b \gamma_{ub}\mathbf{1}[G_{ud} \in b] \cdot CR_{ud} + \mathcal{SP}(CR_{ud}; \zeta_u) + \xi_u \mathbf{G}_{ud} + \omega_{ud}$$

This specification is similar to Equations (1.6) and (A.29) from my demand estimation algorithm, but it differs in a several key ways:

- I estimate Equation (A.45) at the daily (d) level, rather than at the hourly level, because I no longer need to convert hourly generation into hourly coal consumption.[101]

- For the same reason, I now use total daily net generation (MWH_{ud}) as the dependent variable, instead of unit u's capacity factor. This is largely a normalization, as the mapping between capacity factor and electricity production is fixed (to a first approximation) within each unit-month and largely static for most units across months.

- I use a cost ratio (CR_{ud}) that averages across all coal units in unit u's PCA, because I now want to accommodate changes in coal price across *many* plants (as opposed to idiosyncratic changes to each plants' markups).

- Following Cullen and Mansur (2017), I include a cubic spline of the cost ratio, $\mathcal{SP}(CR_{ud}; \zeta_u)$. This allows me to more flexibly model the effect of relative fuel price changes on each unit's generation.[102]

- Instead of using year and quarter-of-year fixed effects, I control for only month-of-year fixed effects. This avoids removing year-on-year variation in fuel prices, which is important for characterizing the effects of counterfactual fuel prices.

[101] For estimating coal demand, hourly observations allow me to more accurately discretize each unit's capacity factor — there is less within-hour variation in plant operations than within-day variation in plant operations. For counterfactuals, estimating Equation (A.45) at the daily level reduces computation time without meaningfully changing the relationship between fuel prices and unit u's predicted generation.

[102] I use cubic splines with 6 knots, however the number of knots does not affect the estimation results.

Step 5: I store predicted values ($\widehat{MWH}_{ud}$) from Equation (A.45), estimated using CR_{ud}. Then, I predicted counterfactual generation ($\widehat{MWH}_{ud}^{NO\Delta Z}$, $\widehat{MWH}_{ud}^{NO\Delta\mu}$), by plugging two counterfactual cost ratios ($CR_{ud}^{NO\Delta Z}$, $CR_{ud}^{NO\Delta\mu}$) into this fitted model.[103]

Step 6: I convert $\widehat{MWH}_{ud}$, $\widehat{MWH}_{ud}^{NO\Delta Z}$, and $\widehat{MWH}_{ud}^{NO\Delta\mu}$ into $\widehat{CO2}_{ud}$, $\widehat{CO2}_{ud}^{NO\Delta Z}$, and $\widehat{CO2}_{ud}^{NO\Delta\mu}$, multiplying by unit u's monthly CO_2 emissions rate.

Step 7: I sum factual and counterfactual coal generation and coal emissions across all units in each month, for all months between December 2008 and December 2015:

$$\text{(A.46)}\qquad \widehat{\mathbf{MWH}}_m^{coal} \equiv \sum_{d\in m}\sum_{u} \widehat{MWH}_{ud} \quad , \qquad \widehat{\mathbf{CO2}}_m^{coal} \equiv \sum_{d\in m}\sum_{u} \widehat{CO2}_{ud}$$

$$\text{(A.47)}\qquad \widehat{\mathbf{MWH}}_m^{coal,NO\Delta Z} \equiv \sum_{d\in m}\sum_{u} \widehat{MWH}_{ud}^{NO\Delta Z} \quad , \qquad \widehat{\mathbf{CO2}}_m^{coal,NO\Delta Z} \equiv \sum_{d\in m}\sum_{u} \widehat{CO2}_{ud}^{NO\Delta Z}$$

$$\text{(A.48)}\qquad \widehat{\mathbf{MWH}}_m^{coal,NO\Delta\mu} \equiv \sum_{d\in m}\sum_{u} \widehat{MWH}_{ud}^{NO\Delta\mu} \quad , \qquad \widehat{\mathbf{CO2}}_m^{coal,NO\Delta\mu} \equiv \sum_{d\in m}\sum_{u} \widehat{CO2}_{ud}^{NO\Delta\mu}$$

Step 8: I sum total monthly natural gas generation and emissions, across all CEMS gas generating units, for all months between December 2008 and December 2015 ($\mathbf{MWH}_m^{gas}$, $\mathbf{CO2}_m^{gas}$) . I calculate the average CO_2 emissions rate per MWh for each month:

$$\text{(A.49)}\qquad \mathbf{E}_m^{gas} = \frac{\mathbf{CO2}_m^{gas}}{\mathbf{MWH}_m^{gas}}$$

Step 9: I calculate counterfactual natural gas emissions in each month by replacing changes in coal generation with gas generation on a 1-for-1 basis, and multiplying by the average natural gas emissions rate:

$$\text{(A.50)}\qquad \widehat{\mathbf{CO2}}_m^{gas,NO\Delta Z} \equiv \mathbf{E}_m^{gas} \times \left[\mathbf{MWH}_m^{gas} - \left(\widehat{\mathbf{MWH}}_m^{coal} - \widehat{\mathbf{MWH}}_m^{coal,NO\Delta Z}\right)\right]$$

$$\text{(A.51)}\qquad \widehat{\mathbf{CO2}}_m^{gas,NO\Delta\mu} \equiv \mathbf{E}_m^{gas} \times \left[\mathbf{MWH}_m^{gas} - \left(\widehat{\mathbf{MWH}}_m^{coal} - \widehat{\mathbf{MWH}}_m^{coal,NO\Delta\mu}\right)\right]$$

[103] I generate cubic splines for each counterfactual cost ratio, forcing the same 6 knot points as the cubic spline for the factual cost ratio in Equation (A.45).

Step 10: I sum total CO_2 emissions for both coal and gas, across all months between December 2008 and December 2015:

$$\widehat{\mathbf{CO2}} \equiv \sum_m \left[\widehat{\mathbf{CO2}}_m^{coal} + \mathbf{CO2}_m^{gas}\right] \tag{A.52}$$

$$\widehat{\mathbf{CO2}}^{NO\Delta Z} \equiv \sum_m \left[\widehat{\mathbf{CO2}}_m^{coal,NO\Delta Z} + \widehat{\mathbf{CO2}}_m^{gas,NO\Delta Z}\right] \tag{A.53}$$

$$\widehat{\mathbf{CO2}}^{NO\Delta\mu} \equiv \sum_m \left[\widehat{\mathbf{CO2}}_m^{coal,NO\Delta\mu} + \widehat{\mathbf{CO2}}_m^{gas,NO\Delta\mu}\right] \tag{A.54}$$

Step 11: I calculate realized abatement under the fracking boom as the percent reduction in realized CO_2 emissions, compared to the no-fracking counterfactual:

$$\mathbf{ABATE}^{REALIZED} = \frac{\widehat{\mathbf{CO2}}^{NO\Delta Z} - \widehat{\mathbf{CO2}}}{\widehat{\mathbf{CO2}}^{NO\Delta Z}} \approx 0.045 \tag{A.55}$$

I similarly calculate potential abatement as the percent reduction in full-pass-through counterfactual CO_2 emissions, compared to the no-fracking counterfactual:

$$\mathbf{ABATE}^{POTENTIAL} = \frac{\widehat{\mathbf{CO2}}^{NO\Delta Z} - \widehat{\mathbf{CO2}}^{NO\Delta\mu}}{\widehat{\mathbf{CO2}}^{NO\Delta Z}} \approx 0.049 \tag{A.56}$$

Based on these calculations, decreasing coal markups eroded roughly 8 percent of the potential CO_2 abatement of the fracking boom:

$$1 - \frac{\mathbf{ABATE}^{REALIZED}}{\mathbf{ABATE}^{POTENTIAL}} \approx 1 - \frac{0.045}{0.049} \approx 0.075 \tag{A.57}$$

A.6.2 Sensitivities and Interpretation

I estimate several alternative specifications, which I report in Table A.6.40. To allow additional flexibility, I introduce 3 additional cubic splines in $G_d \times CR_{ud}$, G_d, and daily maximum temperature, which has little effect on my counterfactual predictions.[104] I also test quarter-of-year fixed effects, in order to match the fixed effects used in Equations (1.6)

[104]This matches Cullen and Mansur (2017)'s main specification, which includes cubic splines in the coal-to-gas cost ratio, total system load, and temperature. The third row in Table A.6.40 includes these three splines, in addition to a spline in the interaction of total generation and cost ratio.

Table A.6.40: Counterfactual Sensitivities

	Realized abatement	Potential abatement	Share eroded
Preferred specification	0.045	0.049	0.075
Spline of $G_d \times CR_{ud}$ interaction	0.047	0.050	0.073
Splines of $G_d \times CR_{ud}$, G_d, temperature	0.046	0.050	0.072
Quarter-of-year FEs	0.045	0.048	0.076
Month-of-year and year FEs	0.024	0.027	0.093
Hourly (not daily) observations	0.044	0.048	0.077

Notes: The top row reports counterfactual results using my preferred specification (Equation (A.45)). The second row replaces the interacted sum (i.e. the second term in Equation (A.45)) with a cubic spline in the interaction of G_d and CR_{ud}. The third row includes this spline and two additional cubic splines in G_d and maximum daily temperature. The fourth row replaces month-of-year fixed effects with quarter-of-year fixed effects. The fifth row uses both month-of-year and year fixed effects. Finally, the last row estimates Equation (A.45) at the hourly level, with hour-of-day fixed effects. "Realized abatement" calculates the share of counterfactual no-fracking CO_2 emissions that did not occur. "Potential abatement" calculates this share using counterfactual coal prices that hold coal markups constant. The last column reports 1 − (realized abatement)/(potential abatement).

and (A.29), which produces nearly identical counterfactual predictions. The fifth row of Table A.6.40 includes year fixed effects, which control for medium-to-long run changes in plant operations while also absorbing most of the identifying variation in natural gas prices. This yields much smaller estimates of CO_2 abatement (0.024 vs. 0.045), which is unsurprising considering that year fixed effects now control for the large drop in prices after 2009. However, this actually implies that decreasing coal markups eroded a larger share of potential CO_2 emissions reductions (0.093 vs. 0.075). Finally, Table A.6.40 includes counterfactuals estimated using hourly observations (as in my demand estimation), to ensure that my decision to aggregate to a daily temporal frequency does not affect these counterfactual predictions.

How do the magnitudes of these results compare to the previous literature on the fracking boom? Using a similar time-series estimation strategy at the interconnection level, Cullen and Mansur (2017) estimate that a tax of \$20 per ton CO_2 would yield 4.9 percent reductions in daily CO_2 emissions.[105] Taking my derived expression for the implicit tax (Equation (A.19)) and plugging in average values for P, MC^{env}_{coal}, MC^{env}_{gas}, Z, and ΔZ, this implies an average implicit tax of \$20–32 per metric ton CO_2.[106] Hence, my

[105] Cullen and Mansur (2017) denominate this tax in short tons of CO_2, while I use metric tons (a.k.a tonnes). A \$20/ton CO_2 tax is equal to a tax of \$22/tonne CO_2.

[106] I use average delivered coal prices across all plants for P; average environmental costs across all coal/gas units, converted into \$/MMBTU for MC^{env}_{coal} and MC^{env}_{gas}; average counterfactual gas prices for

predictions for CO_2 abatement are quite close to Cullen and Mansur (2017)'s results from interconnection-wide reduced-form time series regressions, even though my calculations come from plant-specific time-series regressions for coal units only. As with Cullen and Mansur (2017), my use of time fixed effects (for the sake of econometric identification) absorbs much of the time series variation in natural gas prices (i.e. implicit carbon prices). Hence, a \$20–32 carbon tax may effectively be out-of-sample, giving the variation that remains.

These estimates of CO_2 abatement from short-run fuel-switching do not capture the full extent to which the fracking boom has decreased CO_2 emissions from U.S. electricity generation. Linn, Mastrangelo, and Burtraw (2014) find evidence that coal plants increase their thermal efficiency (i.e. lower their heat rates) in response to competitive pressure; the fracking boom has likely contributed meaningful medium-run CO_2 abatement through this channel. Low gas prices have also led to medium-to-long-run abatement on the capacity margin, by incentivizing investments in new natural gas combined-cycle plants and accelerating coal plant retirements (Brehm (2017)). Finally, long-run dynamics of electric generating capacity imply that even a small carbon tax could have a very large effect on coal capacity (Cullen and Reynolds (2016)). Equation (A.45) ignores each of these sources of fracking-induced CO_2 abatement, which is likely why my counterfactual exercise finds only 4.5 percent abatement from fuel-switching.

A simple event-study analysis suggests that carbon emissions from the U.S. electricity sector have fallen by 20–25 percent during the fracking boom. While this does not establish the *causal* effect of low gas prices on total emissions, it does suggest that total abatement was potentially much greater than my 4.5 percent short-run estimate. This also suggests that the unrealized environmental benefits of the fracking boom maybe have been much larger, since decreasing coal markups likely also impacted each of the above abatement channels, *in addition to* their impact on the short-run coal-to-gas switching margin. Importantly, CO_2 is a global pollutant, and my analysis focuses on U.S. emissions only. The fracking boom also impacted global energy markets, with theoretically ambiguous implications for global CO_2 emissions (Knittel, Metaxoglou, and Trindade (2016); Wolak (2016)).

Z; and the average difference between factual and counterfactual prices for ΔZ. Annual averages for 2010, 2011, and 2012 translate to implied taxes of \$20, \$23, and \$32, respectively.

Appendix B

Appendix: Out of the Darkness and Into the Light? Development Effects of Rural Electrification

B.1 Data

B.1.1 RGGVY Program Data

The Rural Electrification Corporation maintains an online database of RGGVY implementation status, and also hosts two separate portals for "Villages Covered" and "Villages Completed" under the 10th and 11th Plans of RGGVY.[1] Each village-specific page within these portals reported that village's pre-program electrification status, proposed RGGVY implementation details (e.g. number of households to be electrified, new transformer capacity to be installed), actual progress of RGGVY implementation (e.g. number of household connections completed, new transformer capacity installed), and implementation status (e.g. whether work has been completed in this village).[2] Unfortunately, these village-level data are of very poor quality, with rampant missing information, internal inconsistencies, and program outcomes that conflict with village-level Census data.[3] As detailed in Table B.1.1, RGGVY programs outcomes are missing for 65 percent of villages

[1]We downloaded these data from the RGGVY home page, `http://rggvy.gov.in`, which has since been deactivated as RGGVY has ben subsumed into DDUGJY. The new program website is `http://www.ddugjy.in/`.

[2]Notably, neither portal recorded the date on which a village was sanctioned for electrification, nor the date on which works were begun or completed. The only timing recorded in either set of webpages describes, to the best of our knowledge, was the latest upload of data to the internet. Each of these datasets had a separate tab on the RGGVY homepage, respectively: "Villages Covered (X & XI Plan DPR)" and "Villages Completed (X & XI Plan)". We scraped these datasets between August and October 2014.

[3]The Census data seems to be of relatively high quality, with no evidence of population manipulation. Asher and Novosad (2018) shows that PCA data has a high correlation with another Indian dataset, the Socioeconomic and Caste Census; we also find high correlations between Census data and National Rural

that were eligible under the 10th Plan (our analysis focuses on this earlier wave of the RGGVY program). These data also report a greater number of covered habitations than exist for 32 percent of villages.[4]

Table B.1.1: RGGVY Microdata Irregularities

Type of Irregularity	Percent of Villages	Percent of 10th-Plan Villages
RGGVY outcomes disagree across Covered & Completed datasets	26.8	32.7
Outcomes missing from either Covered or Completed dataset	77.9	65.3
Outcomes missing from both Covered and Completed datasets	74.4	59.9
All outcomes missing from both Covered and Completed datasets	33.4	22.3
Completed dataset reports status not energised	24.4	14.1
RGGVY covers more habitations than exist in village	32.2	31.7

Notes: This table shows data irregularities across the RGGVY Covered and Completed village datasets, which we do not use in our analysis. The right two columns shows the percent of all villages and the percent of 10th-Plan villages that satisfy each irregularity criterion, where the denominator excludes missing and unmatched villages. Program outcomes considered in the first four rows include the count of household connections, aggregate transformer capacity installed, and aggregate transmission capacity installed. (The first three rows count villages where *any* outcome disagrees or is missing; the fourth row counts only villages for which *all* of these outcomes are missing.)

Even if these village-level program data were of better quality, we might worry about using them for our analysis. We would expect the RGGVY microdata to identify precisely which villages were treated under the program. However, we would still need to construct a control group from the subset of villages *not* represented in the RGGVY microdata. Because the RGGVY datasets do not include information on villages left out of the program, we would need to merge the RGGVY microdata into village-level Census data and denote villages included in the RGGVY dataset as "treated" and all other villages as "control." Any imperfect merge or missing RGGVY microdata would cause us to incorrectly categorize a village. We might also worry about manipulation of RGGVY village-level outcomes, if implementing agencies (or data tabulators) had an incentive to overstate the extent of electrification attributable to RGGVY.[5]

For these reasons, we have chosen to exclude the RGGVY village-level data from our analysis entirely. Instead, we rely on district-level RGGVY summary reports to determine the Five-Year Plan under which each electrification project was sanctioned, the implementing agency responsible for implementation, and the approximate timing of

Drinking Water Programme's census of habitations (described further in Section B.1.5). For these reasons, we are inclined to trust the Census data over the RGGVY data.

[4]We base this total number of habitations on our merge of the Habitation Census to the village-level Census panel, considering only those RGGVY villages that match to a *matched* village from the habitation census merge (described below).

[5]We are not the only researchers to find inconsistencies in microdata from Indian programs. Asher and Novosad (2018) document striking irregularities in PMGSY population data.

electrification. Since the program was implemented based on Detailed Project Reports (DPR) at the district level, we can use these aggregate data to link villages to DPRs, and, importantly, to the 10th or 11th Plan.[6] Our analysis only relies on these district-level reports to assign districts to Five-Year Plans. As this is a matter of public record and involves large transfers of public funds from the Rural Electrification Corporation to decentralized implementing agencies, we are much more confident in the accuracy of this aggregate information.

Table B.1.2 summarizes these district-level progress reports.[7] We see that under both the 10th and 11th Plans, the majority of district-level DPRs were implemented by local electricity distribution companies. However, there were also many districts whose implementing agencies included large public sector undertakings, state departments of power, and state electricity boards. This table also shows that the majority of villages covered under both Plans were categorized as "electrified" villages, meaning that there was a minimum of power access prior to RGGVY.[8] For these villages, RGGVY sought to provide "more intensive electrification," including bringing energy access to below-poverty-line households that still lacked connections. In early 2015, the program reported that over 97 percent of 10th-Plan projects had been completed, and that over 90 percent of 11th-Plan projects had been completed.

B.1.2 Geospatial Data

Our main source of geospatial information is ML Infomap's VillageMap.[9] This dataset includes village boundary shapefiles for nearly every village in India, as defined by the 2001 village-level Census. We take the 2001 boundaries as fixed, which is consistent with our decision to use the 2001 village as our unit of analysis.[10] Every square meter in India belongs to a village or city/town; the only "blank spaces" in the village maps are forests, bodies of water, and "towns" (urban areas). Village boundaries are set by the Census Organization of India (Census of India (2011)). Figure B.1.1 shows an example of the level of detail included in the village boundary dataset.

[6] These reports were available at `http://rggvy.gov.in` under the "Progress Reports" menu, and can now be found at `http://ddugjy.gov.in`. Most districts appear in exactly one DPR. For districts with multiple DPRs, we have aggregated DPRs up to the district-level, in order to create a one-to-one mapping between districts and DPRs. Of the 530 RGGVY districts, only 30 districts had DPRs aggregated across both the 10th and 11th Plans.

[7] The state of Goa was excluded from RGGVY along with all 7 union territories, because 100 percent of their villages were electrified prior to 2005 (Ministry of Power (2012)). We treat Telangana as part of Andhra Pradesh, since its 2014 split from Andhra Pradesh occurred after our period of analysis.

[8] The definition of "electrification" has changed over time in India, but as long as "one bulb was burning" anywhere in the village, a village was considered electrified.

[9] These data are also used by Min (2011) and Asher and Novosad (2017), among others, to map villages in India.

[10] In the 2 percent of cases where a 2001 village matches to multiple 2011 villages in the 2001/2011 Census concordance, we aggregate 2011 data up to the 2001 village definition.

Table B.1.2: Summary Statistics – RGGVY Implementation and Scope

Type of Implementing Agency	States	Districts	Award Dates	Unelectrified Villages	Electrified Villages	BPL Connections
A. 10th Plan						
Public Sector Undertakings	11	57	2005–07	32,638	20,126	2,386,042
State Depts. of Power	5	10	2007–10	542	1,088	45,921
State Electricity Boards	3	6	2006–07	4,482	2,604	441,639
Distribution Companies	14	156	2005–08	26,429	76,495	4,649,733
Total	25	229		64,091	100,313	7,523,335
B. 11th Plan						
Public Sector Undertakings	9	78	2008–11	30,298	62,705	6,768,765
State Depts. of Power	5	33	2008–10	2,710	3,850	212,887
State Electricity Boards	3	34	2008–11	80	18,166	149,882
Distribution Companies	15	183	2008–10	13,118	135,038	7,363,814
Rural Electricity Coops	2	5	2008–11	0	755	79,220
Total	25	331		46,206	220,514	14,574,568

Notes: Data summarize district-level progress reports, previously available at `http://rggvy.gov.in`. Public sector undertakings include government-owned generating companies, such as Power Grid Corporation of India and National Hydroelectric Power Corporation. The right three columns show the number of previously unelectrified and previously electrified villages covered by the program, as well as the the number of below poverty line households to receive electric connections. Villages classified as electrified had basic electricity infrastructure with at least 10 percent of households electrified prior to RGGVY implementation. 23 (of 27) states contain both 10th and 11th Plan districts, while 30 (of 530) individual districts were targeted under both Plans. For a few districts, we correct financial award dates reported to have occurred *before* their respective project sanction dates or before the official announcement of the program.

We were unable to acquire village shapefiles for Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, and Sikkim. This forces us to drop these 5 states from our geospatial dataset (and, subsequently, from our analysis of nighttime lights), though we are still able to include them for the remainder of our analysis. Fortunately, these small states comprise only 2 percent of all Indian villages and less than 1 percent of India's rural population. They also represent less than 3 percent of villages covered under RGGVY, and only 1 percent of villages covered by the 10th Plan.

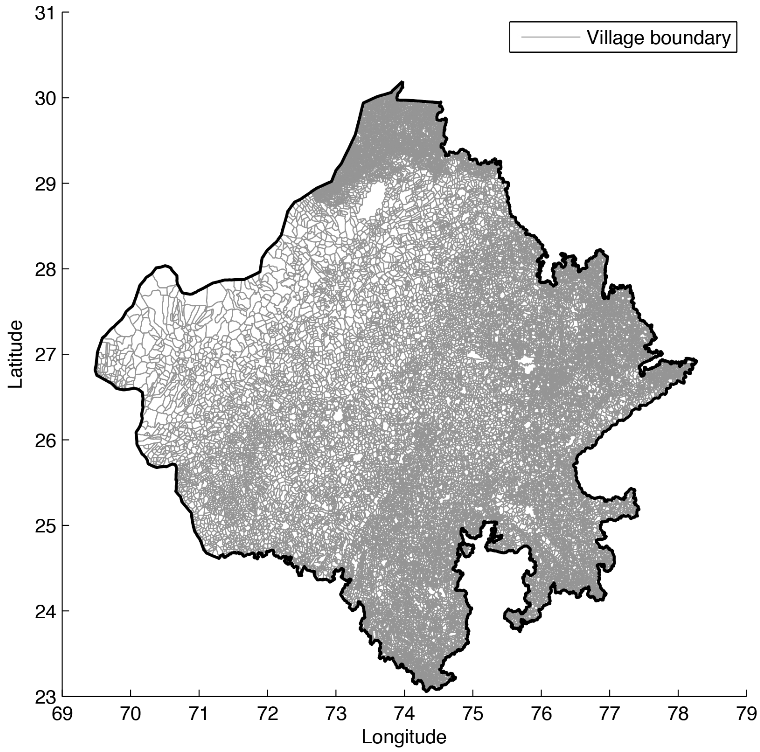
We also exclude Assam, Himachal Pradesh, Jammu and Kashmir, Uttar Pradesh, and Uttarakhand from our geospatial dataset, because we believe that the shapefiles for these states are extremely poor quality. For these 5 states, the correlation between the village area measurement implied by the shapefiles and the village area reported in the Indian Census (the official body in charge of defining village boundaries) is below 0.35. Table B.1.3 shows these correlations by state, demonstrating a clear gap between the 12 correlated and 5 uncorrelated states. These uncorrelated states represent 39 percent of villages in RGGVY 10th-Plan districts, most of which are in Uttar Pradesh. While we restrict our analysis of nighttime brightness to these 12 correlated states, our analysis of Census outcomes includes villages in all 22 states in Table B.1.3.

Table B.1.3: Correlation of Shapefiles with Village Areas

State	Area Correlation	Percent of Total Villages	Percent of 10th-Plan Villages
Jharkhand	0.978	5.0	5.9
West Bengal	0.954	6.4	10.8
Bihar	0.932	6.6	9.8
Gujarat	0.896	3.1	0.8
Haryana	0.873	1.1	0.4
Karnataka	0.806	4.6	6.5
Maharashtra	0.781	7.0	1.5
Andhra Pradesh	0.772	4.5	7.2
Rajasthan	0.714	6.8	9.9
Orissa	0.680	8.1	2.5
Madhya Pradesh	0.638	8.9	3.4
Chhattisgarh	0.605	3.3	1.2
Uttarakhand	0.326	2.7	5.3
Uttar Pradesh	0.281	16.6	31.8
Himachal Pradesh	0.138	3.0	0.4
Assam	0.106	4.1	1.0
Jammu and Kashmir	0.002	1.1	0.5
Arunachal Pradesh	missing	0.6	0.3
Meghalaya	missing	0.9	0.3
Mizoram	missing	0.1	0.1
Nagaland	missing	0.2	0.1
Sikkim	missing	0.1	0.1
States with correlation > 0.35		68.3	59.9
States with correlation < 0.35		28.8	39.0
States with missing shapefiles		2.9	1.0

Notes: This table shows the correlation between areas calculated from village shapefiles and village areas reported in the Census's 2001 Village Directory. Our spatial dataset includes only the 12 states for which this correlation is at least 0.35. We omit the 5 states with shapefile areas that are uncorrelated with reported village areas, as we take this as a sign of low quality shapefiles. The middle column reports the percent of Indian villages contained in each state, while the right column shows the percent of villages in districts eligible for RGGVY under the 10th Plan. Omitted from this table are 3 states which were not eligible under RGGVY's 10th Plan (Goa, Punjab, and Tamil Nadu), as well as 2 states which were eligible under RGGVY's 10th Plan but contain no single-habitation 10th-Plan villages in our RD bandwidth (Kerala and Tripura).

Figure B.1.1: Rajasthani Village Boundaries



Notes: This figure shows the approximately 41,575 villages in Rajasthan. The 1st and 99th percentiles of Rajasthani village area are 0.2 km^2 and 29.5 km^2, respectively.

Our full geospatial dataset includes village boundaries for 172,013 villages across the 12 RGGVY states remaining after the sample restrictions detailed above.[11] Each boundary shapefile is identified by its 2001 Census code, as well as village attributes from the 2001 Primary Census Abstract. We use this identifying information to link geospatial information into our administrative datasets.

B.1.3 Nighttime Lights

In order to credibly measure electrification, we use the National Oceanic and Atmospheric Administration (NOAA)'s Defense Meteorological Satellite Program-Operational Line Scan (DMSP-OLS) Nighttime Lights data[12]. Descriptions of these data can be found in Elvidge et al. (1997) and Doll (2008).[13] Data are publicly available from 1992 to 2013;

[11]This number excludes village with null or missing populations, which appear to have been miscoded.

[12]The data are available for download here: `http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`

[13]Researchers have also used these nighttime lights data as proxies for economic activity, including Noor et al. (2008), Bleakley and Lin (2012), Henderson, Storeygard, and Weil (2012), Li, Ge, and Chen

we use images from 1999 to 2013 in this paper. These images are compiled from nightly satellite photographs taken between 8:30 and 10:00 PM local time, and they are extremely high resolution: pixels are 30 arc-second squares.[14] The Indian subcontinent alone contains 417,876 pixels. Each pixel is assigned a digital number (DN) indicating brightness, ranging from 0 to 63. This DN is approximately proportional to average luminosity.[15] While the images often top-code very bright locations such as urban centers (Henderson, Storeygard, and Weil (2012)), we focus our attention on rural areas with very low risk of top-coding.

We might instead worry that the DMSP-OLS sensors are not sensitive enough to detect the subtle changes in brightness associated with rural electrification. However, a substantial body of evidence suggests otherwise. Elvidge et al. (1997) and Elvidge et al. (2001) use DMSP-OLS data to estimate electrification rates around the world at the national level. A variety of papers have also mapped nighttime lights to electrification rates at the sub-national level, including Ebener et al. (2005), Doll, Muller, and Morley (2006), Chand et al. (2009), and Townsend and Bruce (2010). Three studies are particularly relevant to our work: Min et al. (2013) use original survey data from Senegal and Mali to show that electrified rural villages are significantly brighter in the DMSP-OLS data than their unelectrified counterparts. Min and Gaba (2014) find a strong correlation between village-level ground-based electricity usage survey data and DMSP-OLS nighttime lights imagery in rural Vietnam, showing that both streetlights and electrified homes are correlated with higher DMSP-OLS DN readings. Finally, Min (2011) shows a strong correlation between power distribution and nighttime lights and administrative data on electrification and nighttime lights in the Indian state of Uttar Pradesh.[16] Based on these findings, we are confident that the DMSP-OLS data are capable of accurately measuring rural electrification — any activity in rural India bright enough to be visible from space is likely to require electricity.[17] Furthermore, any bottom coding in our data will lead us to underestimate the effects of RGGVY on nighttime brightness.[18]

(2013), Michalopoulos and Papaioannou (2013), and Michalopoulos and Papaioannou (2014). Given that we are studying electrification directly, we refrain from using nighttime lighting as a proxy for GDP.

[14]These pixels are squares with approximately 500 meter sides at the equator.

[15]See Chen and Nordhaus (2011) for details.

[16]Note that this analysis was done at the district level, obviating the need for village-level shapefiles.

[17]There are two obvious exceptions: the first is agricultural fires; the second is car headlamps. Because of the yearly nature of the DMSP-OLS data, ephemeral fires do not appear in the final stable lights averaged datasets (details can be found in NOAA's DMSP-OLS data description). Roads are sparse in the villages we are looking at; it is extremely unlikely that there is enough road traffic to appear in the DMSP-OLS dataset. Even if there were enough road traffic to be detected in any given flyover of the satellite, this road traffic is likely to be erratic enough, and the satellite images are taken at variable enough times of night, that it is unlikely that we mistake vehicle traffic for consistent village electrification (Min and Gaba (2014)).

[18]Suppose, for example, that the satellite can detect brightnesses of λ or greater only, and that at baseline, villages A and B both have brightness of $\lambda - 5$. These are both coded as λ. Village B is electrified under RGGVY, and now has brightness of $\lambda + 5$. We observe a difference of 5 between A and B, when in fact, the true difference was 10.

NOAA releases three different DMSP-OLS lights products: average visible lights; stable lights; and average visible × percent lights. The average visible lights dataset contains the average DN over the year's observations for each pixel. The stable lights dataset is a more heavily processed version of the average visible lights: it contains the lighting from persistently lit places, and does not include the light from fires and other sporadic events. Finally, the average visible × percent lights takes the average visible DN in a pixel and multiplies it by the frequency with which it is observed.[19] Our analysis focuses on the average visible lights data, which are best equipped to detect the low levels of lighting associated with electrification of rural villages (Min et al. (2013)). Section B.2.1 performs a robustness check using the stable lights dataset.

We construct village-level nighttime lights values by combining village shapefiles with the nighttime brightness data in ArcGIS, overlaying the 2001 village boundaries on top of lights images for each year. Figure B.1.2 shows an example of this overlay from a region of Andhra Pradesh. For each village, we calculate the maximum DN value over all pixels whose centroids are contained within a village boundary. In other words, a village's brightness for a given year is equal to the brightness of its brightest pixel.[20] For villages too small to contain a pixel's centroid, we assign the value of the pixel at the village centroid as the maximum lights value.[21] Between 1999 and 2007, NOAA had two satellites operating DMSP-OLS equipment.[22] For these years, we calculate the mean and maximum lights values for each satellite separately, and then we take an unweighted average across satellites to obtain village-year DN statistics.

One concern about the DMSP-OLS dataset is that the sensors used to calculate the DN tend to degrade and become dirty over time. The satellites do not contain on-board calibration equipment, and the sensors are only adjusted before being loaded onto the satellite platform. NOAA's Earth Observation Group, which manages DMSP-OLS data, has done some ex-post calibration in order to bring different satellite-years in line with one another. However, this algorithm must assume that brightness in one region (usually the island of Sicily is used) remains fixed over time, thereby calibrating the sensors to that region's DN values. This process is imperfect and not fully transparent. Most economists who use these data in a panel setting include satellite or year fixed effects to control for inconsistencies in sensors over time (e.g. Henderson, Storeygard, and Weil (2012)). In our empirical analysis of nighttime brightness, we use an estimator that relies primarily on cross-sectional variation in brightness, which should assuage these concerns.

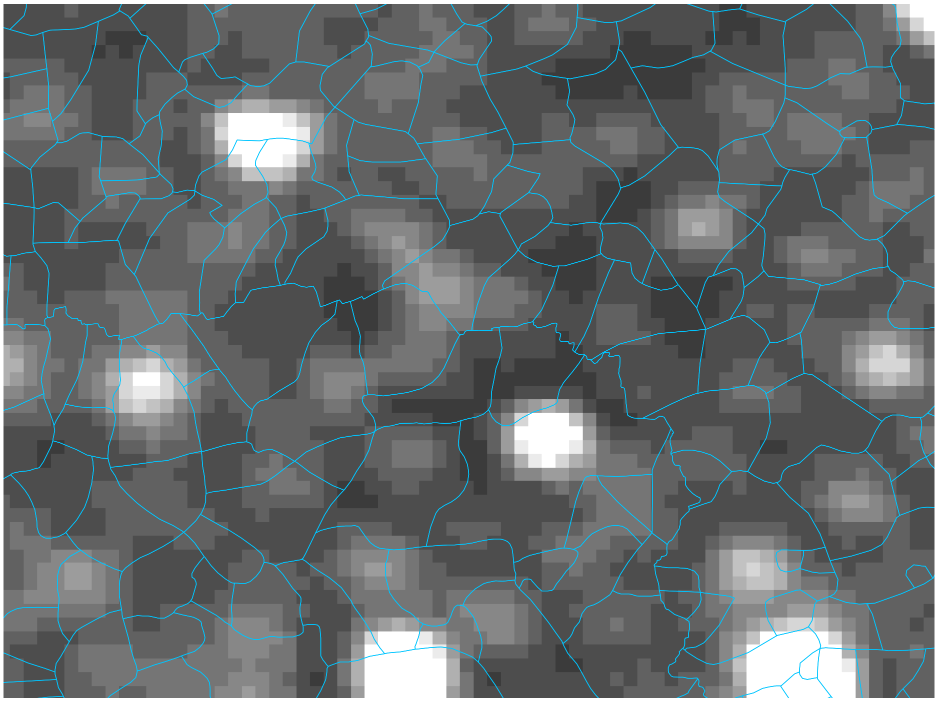
[19] These lights have been used by Alam (2013) to examine power quality in India, but they are much less frequently used in the economics literature.

[20] We also conduct sensitivity analysis on assigning each village the average brightness across all its pixels. We use the standard `Zonal Statistics as Table` operation in ArcGIS to calculate both the maximum and mean brightness.

[21] All of the villages that did not contain a pixel centroid only overlapped with one pixel, so this is the correct operation for these very small villages.

[22] In 1999, F12 and F14 were active; from 2000 to 2003, F14 and F15 were active; from 2004 to 2007, F15 and F16 were active; from 2008 to 2009, only F16 was active; and from 2010 to 2013, only F18 was active.

Figure B.1.2: Example of Nighttime Lights with Village Boundaries



Notes: This image shows a close-up of averge visible nighttime brightness overlaid with village boundaries, for an area in Rajasthan. The $\approx$1km^2 pixels in this image range in brightness values from 3 to 38.

We undertake an additional procedure to remove measurement error from the lights data. After constructing a village-year panel of the maximum nighttime brightness, we linearly project lights values from 2001 and 2011 on the values of the two years before and after. We run the following OLS regressions, weighting by village area[23]:

$$\begin{aligned} L_v^{2001} &= \alpha_0 + \alpha_1 L_v^{1999} + \alpha_2 L_v^{2000} + \alpha_3 L_v^{2002} + \alpha_4 L_v^{2003} + \varepsilon_v \\ L_v^{2011} &= \kappa_0 + \kappa_1 L_v^{2009} + \kappa_2 L_v^{2010} + \kappa_3 L_v^{2012} + \kappa_4 L_v^{2013} + \upsilon_v \ , \end{aligned} \tag{B.1}$$

where L_v^y is the maximum brightness of village v in year y. We use the estimated $\hat{\alpha}_i$s and $\hat{\kappa}_i$s (for $i = \{0, 1, 2, 3, 4\}$) to construct the nighttime lights outcome variables we use in our regressions, $\hat{L}_v^{2001}$ and $\hat{L}_v^{2011}$. This removes random year-to-year noise in the 2001 and 2011 lights data that cannot be explained by the brightness observed in adjacent years. It also isolates the more stable year-to-year changes in brightness that we would associate with new electricity infrastructure. In Section B.2.1, we include a sensitivity analysis that varies the number of adjacent years in these linear projections, while also considering simple unweighted averages of adjacent years.

[23] The results are not sensitive to the decision to weight. Unweighted regressions produce nearly identical results.

B.1.4 Census Data

We construct a village-level panel dataset using three datasets from the Census of India's 2001 and 2011 decennial Censuses. These datasets are all available for download from the Census of India's website.[24] Below, we describe each dataset separately, along with the process we used to construct our 2001–2011 village panel.

B.1.4.1 Primary Census Abstract

The Primary Census Abstract (PCA) reports village population and employment information for all geographic units across India's 28 states and 8 union territories.[25] This dataset includes the total number of households in each village, along with village population broken down by gender, the 0–6 age cohort, scheduled caste, and scheduled tribe.[26] Village-level literacy rates are also included. The 2001 PCA dataset contains 593,643 villages, while the 2011 dataset contains 597,483 villages.

The PCA reports employment counts by gender, for three disjoint subsets of the village population: "main workers" who participate in any economically productive activity (with or without compensation) for at least 6 months of the year; "marginal workers" who do so for less than 6 months of the year; and "non-workers" who do not participate in economically productive activity. Within each of these groups, workers of each gender are separately categorized as either cultivators, agricultural laborers, household industry workers, or other. The distinction between cultivators and agricultural laborers is that agricultural laborers work for wages, while taking on no risk in cultivation and owning no right to cultivate land. "Other" workers include all workers not covered by the other three categories, such as government workers, teachers, doctors, and factory workers, and includes all non-farm, non-household employment.

B.1.4.2 Houselisting Primary Census Abstract

The Houselisting Primary Census Abstract (HPCA) reports on a variety of household attributes and amenities at the village level. These include physical housing stock characteristics such as type of floor/wall/roof and number of rooms; drinking water source; type of latrine; primary cooking fuel; and main source of in-home lighting (e.g., electricity vs. kerosene). This dataset also includes information on household inhabitants, including the number of members; number of married couples; and whether houses are owned or rented. Finally, the HPCA includes data on household asset ownership, including whether houses own mobile phones, televisions, motorcycles, radios, and other durable goods.

[24] `http://censusindia.gov.in`. We downloaded these data between September 2014 and August 2015.

[25] India currently has 29 states, but the Andhra Pradesh-Telangana split occurred in 2014, after the 2011 data were collected and published.

[26] Scheduled castes (SC) and scheduled tribes (ST) are official designations for castes and ethnic groups that have been historically disadvantaged. Since its independence, India has targeted SC and ST communities for affirmative action in social programs and political representation.

For each of the above categories, the HPCA reports the proportion of households within each administrative unit that satisfy each respective criterion. This allows us to treat each variable as continuous, with considerable variation across villages. The 2011 HPCA is publicly available at the village level, and it reports on 597,502 villages. The 2001 HPCA is only available at the Census block level (i.e. the administrative unit between district and village), and it reports block-specific values for most variables across all 5,415 blocks. However, a few variables are only reported at the district level, including indicators of physical upkeep, ownership status, number of rooms, and number of married couples per household.

B.1.4.3 Village Directory

The Village Directory (VD) provides detailed village-level information on a variety of amenities.[27] These data are analogous to a community survey that is often included with household survey data. Unlike the amenities featured in the HPCA, the VD reports public, community-level characteristics that are not specific to individual households. These include the number of primary/middle/secondary schools and other educational facilities; the number of hospitals, community health centers, and other health facilities; community drinking water sources; phone, post office, and other communication services; bus/rail service and road quality; and the presence of banking facilities and credit societies. The dataset also includes 1/0 indicators for the availability of electric power services, broken out by agricultural, domestic, and commercial end-use sectors; the 2011 VD additionally reports average hours per day of electric power received by each sector. For most villages, the VD lists the most important manufacturing and agricultural commodities (the latter in 2011 only). Finally, the VD contains several geographic variables, including village area, area of cropland irrigated (by water source), distance to the nearest road and navigable waterway, and distance to the nearest town. After removing villages with populations that are either zero or missing, the 2001 and 2011 VD datasets contain 593,643 and 596,615 villages, respectively.

B.1.4.4 Census panel dataset

We are able to match villages across the above six datasets using their official census codes. Within each Census year (2001 and 2011), state, district, and village codes are coded consistently across PCA, HPCA, and VD datasets, so merging these data is straightforward.[28] In order to link villages across Census years, we take advantage of the Census's 2001–2011 concordance. We treat the 2001 PCA village as our master cross-sectional unit,

[27]The 2001 VD was a standalone product, while the 2011 VD was distributed as part of the District Census Handbook (DCHB).

[28]Subdistrict, tehsil, and block codes are not consistently coded across datasets, which reflects different administrative conventions across states. For example, while the PCA and HPCA assigns a single tehsil code to each village, the VD assigns separate tehsil and block codes. In our merges, we match on only state, district, and village code, ignoring block code. Because village code is *virtually* unique within each district, this does not affect the accuracy of this merge.

re-aggregating any 2001 villages that split into multiple villages by 2011.[29] Our final panel includes only those villages that match to all 5 village level datasets – 2001/2011 PCA, 2011 HPCA, and 2001/2011 VD. Since village-level data do not exist for the 2001 HPCA, we instead assign each village the values from its parent block (or district, when block-level data is unavailable).[30]

In many cases, the variables reported in the 2011 Census datasets differ from those published in the corresponding 2001 datasets. This is especially true for the VD and certain sections of the HPCA. Because our preferred specification includes a control for the 2001 level of the outcome variable wherever possible, we combine and redefine variables such that our final panel contains only consistent variables.[31] The full panel dataset contains 580,643 villages, with over 200 matching pre/post variables. Tables B.1.4, B.1.5, and B.1.6 provide summary statistics for our Census panel dataset, for three subsets of variables originating from PCA, HPCA, and VD, respectively.

B.1.5 Habitation Merge

Because the RGGVY program determines eligibility based on habitation population (only villages with constituent habitations of at least 300 people were eligible for electrification under the 10th Plan), the 2001 village population as reported in the PCA provides an imperfect indicator of eligibility status.[32] Any village with a population below 300 cannot contain a habitation that is eligible under the 10th Plan. However, a village with a population above 300 may or may not be eligible, depending on the number and size of its constituent habitations. In order to accurately assign eligibility status, we supplement our Census panel dataset with information about the habitations within each village.

To the best our knowledge, there exists only one comprehensive nationwide habitation-level data source: the National Rural Drinking Water Programme conducted a census of habitations in 2003 and 2009 with the purpose of assessing the drinking water availability for all rural habitations in India.[33] The two waves of the census list the habitation names

[29] There are many 2001 villages that match to more than one 2011 village, based on the 2001–2011 concordance. We interpret these as administrative splits, and re-aggregation of all 2011 variables affords us a consistent comparison across years. We drop the *very* few cases (i.e. < 0.01 percent) where multiple 2001 villages appear to have merged into a single 2011 village.

[30] While this is imperfect, the alternative would be to ignore any block-level information on the 2001 levels of 2011 HPCA variables. Our RD analysis of 2011 village-level outcomes greatly benefits from the use of 2001 controls to increase precision.

[31] For example, the 2001 VD lists multiple types of tubewells, while the 2011 VD lists only a single tubewell indicator. We construct a single tubewell indicator from the 2001 VD, such that this indicator is coded identically across VD years.

[32] Recall that a habitation is a sub-village administrative unit, similar to a neighborhood. Habitations were not official census administrative units, but are frequently used in making policy that affects rural village in India.

[33] The data, along with more information on the National Rural Drinking Water Programme are available from `http://indiawater.gov.in/imisreports/nrdwpmain.aspx`. In fact, the RGGVY program documentation lists this habitation census as a reference to be used by implementing agencies (Ministry of Power (2014b)).

Table B.1.4: Summary Statistics – Primary Census Abstract

Village Characteristics	All Districts		10th-Plan Districts		10th-Plan Districts 150–450 Population	
	2001	2011	2001	2011	2001	2011
Village population	1222.19	1416.42	1234.40	1442.17	297.60	359.46
	(1713.62)	(1963.73)	(1607.80)	(1879.55)	(86.60)	(154.96)
Number of households	226.02	286.50	218.46	277.65	54.08	70.55
	(337.69)	(418.29)	(303.64)	(382.86)	(18.50)	(31.10)
Share of pop SC or ST	0.36	0.37	0.33	0.33	0.39	0.40
	(0.32)	(0.32)	(0.30)	(0.30)	(0.36)	(0.36)
Literacy rate	0.46	0.57	0.44	0.56	0.44	0.55
	(0.17)	(0.15)	(0.16)	(0.13)	(0.17)	(0.15)
Employment rate	0.44	0.45	0.42	0.42	0.44	0.45
	(0.13)	(0.14)	(0.13)	(0.14)	(0.14)	(0.15)
Male employment rate	0.53	0.54	0.52	0.52	0.52	0.53
	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)
Female employment rate	0.35	0.35	0.32	0.32	0.36	0.36
	(0.21)	(0.22)	(0.21)	(0.22)	(0.23)	(0.23)
Male ag workers/male pop	0.41	0.40	0.40	0.40	0.42	0.42
	(0.14)	(0.15)	(0.14)	(0.15)	(0.15)	(0.16)
Male hhold workers/male pop	0.01	0.01	0.01	0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Male oth workers/male pop	0.11	0.12	0.10	0.11	0.09	0.10
	(0.11)	(0.12)	(0.10)	(0.11)	(0.10)	(0.12)
Female ag workers/female pop	0.30	0.28	0.27	0.25	0.32	0.30
	(0.22)	(0.22)	(0.21)	(0.21)	(0.23)	(0.23)
Female hhld workers/female pop	0.01	0.01	0.01	0.01	0.01	0.01
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)
Female oth workers/female pop	0.04	0.05	0.03	0.05	0.03	0.05
	(0.08)	(0.08)	(0.07)	(0.08)	(0.07)	(0.09)
Male main workers/male pop	0.44	0.41	0.43	0.38	0.42	0.37
	(0.13)	(0.16)	(0.13)	(0.17)	(0.15)	(0.19)
Female main workers/female pop	0.18	0.18	0.16	0.16	0.17	0.17
	(0.19)	(0.19)	(0.18)	(0.18)	(0.20)	(0.20)
Male marg workers/male pop	0.09	0.13	0.09	0.14	0.10	0.16
	(0.11)	(0.14)	(0.10)	(0.14)	(0.12)	(0.17)
Female marg workers/female pop	0.17	0.17	0.16	0.16	0.19	0.19
	(0.18)	(0.19)	(0.17)	(0.18)	(0.20)	(0.21)
Number of villages	580,643	580,643	290,067	290,067	62,647	62,647

Notes: This table reports means and standard deviations from the 2001 and 2011 Primary Census Abstract. The left two columns include all villages that match across Census datasets, for all 27 RGGVY states. The middle two columns include the subset of those villages located in districts eligible for RGGVY under the 10th Plan. The right two columns include only 10th-Plan villages with 2001 populations between 150 and 450 (i.e. our RD bandwidth). Worker by sector is presented as the fraction of total workers in the village (main + marginal for each gender, repsectively). Agricultural workers include both cultivators and agricultural laborers; other workers are classified as non-agricultural, non-household workers. By definition, main workers work at least 6 months per year, while marginal worker work less than 6 months per year. The employment rate (by gender) divides the sum of main and marginal workers by the village population (of that gender). SC and ST refer to Schedule Caste and Scheduled Tribe designations.

Table B.1.5: Summary Statistics – Houselisting Primary Census Abstract

Household Characteristics	All Districts		10th-Plan Districts		10th-Plan Districts 150–450 Population	
	2001	2011	2001	2011	2001	2011
Average household size	5.42	5.03	5.60	5.24	5.53	5.16
	(0.61)‡	(0.82)	(0.64)‡	(0.86)	(0.61)‡	(0.86)
Average number of rooms	2.11	2.03	2.18	2.04	2.20	2.06
	(0.52)‡	(0.73)	(0.54)‡	(0.73)	(0.54)‡	(0.81)
Good condition (share HH)	0.43	0.43	0.42	0.43	0.43	0.43
	(0.14)‡	(0.31)	(0.11)‡	(0.30)	(0.12)‡	(0.35)
Livable condition (share HH)	0.51	0.50	0.51	0.51	0.50	0.51
	(0.12)‡	(0.29)	(0.10)‡	(0.28)	(0.11)‡	(0.33)
Dilapidated condition (share HH)	0.06	0.06	0.06	0.07	0.06	0.07
	(0.04)‡	(0.10)	(0.03)‡	(0.11)	(0.03)‡	(0.12)
Owns phone (share HH)	0.03	0.51	0.02	0.54	0.02	0.51
	(0.04)†	(0.27)	(0.02)†	(0.26)	(0.02)†	(0.29)
Owns TV (share HH)	0.17	0.28	0.15	0.24	0.14	0.23
	(0.14)†	(0.25)	(0.10)†	(0.21)	(0.10)†	(0.22)
Owns bicycle (share HH)	0.43	0.47	0.48	0.51	0.45	0.50
	(0.23)†	(0.29)	(0.25)†	(0.30)	(0.26)†	(0.33)
Owns motorcycle/scooter (share HH)	0.06	0.13	0.05	0.12	0.05	0.11
	(0.05)†	(0.13)	(0.03)†	(0.12)	(0.03)†	(0.13)
Owns car (share HH)	0.01	0.02	0.01	0.02	0.01	0.02
	(0.01)†	(0.04)	(0.01)†	(0.04)	(0.01)†	(0.04)
Electric/gas cooking (share HH)	0.05	0.08	0.04	0.06	0.05	0.06
	(0.07)†	(0.16)	(0.06)†	(0.13)	(0.07)†	(0.13)
Non-electric/gas cooking (share HH)	0.94	0.92	0.96	0.93	0.95	0.94
	(0.08)†	(0.16)	(0.06)†	(0.13)	(0.07)†	(0.14)
Thatched roof (share HH)	0.27	0.21	0.30	0.23	0.27	0.22
	(0.26)†	(0.26)	(0.24)†	(0.25)	(0.24)†	(0.27)
Mud floor (share HH)	0.76	0.70	0.78	0.73	0.79	0.76
	(0.18)†	(0.29)	(0.17)†	(0.28)	(0.16)†	(0.28)
Electric/solar lighting (share HH)	0.40	0.52	0.31	0.45	0.32	0.49
	(0.29)†	(0.37)	(0.25)†	(0.36)	(0.26)†	(0.38)
Non-electric/solar lighting (share HH)	0.59	0.48	0.68	0.54	0.68	0.51
	(0.30)†	(0.37)	(0.25)†	(0.36)	(0.26)†	(0.38)
Indoor water (share HH)	0.25	0.29	0.25	0.29	0.21	0.23
	(0.19)†	(0.30)	(0.19)†	(0.29)	(0.17)†	(0.28)
Number of villages		580,643		290,067		62,647

Notes: This table reports means and standard deviations from the 2001 and 2011 Houselisting Primary Census Abstract. The left two columns include all villages that match across Census datasets, for all 27 RGGVY states. The middle two columns include the subset of those villages located in districts eligible for RGGVY under the 10th Plan. The right two columns include only 10th-Plan villages with 2001 populations between 150 and 450 (i.e. our RD bandwidth). The 2001 HPCA reports data at higher levels of aggregation than village (i.e. either block or district), hence 2001 columns report district- or block-level averages. A ‡ indicates between-district standard deviations, while a † indicates between-block standard deviations. Share variables are reported in the HPCA as the share of households satisfying each condition. Households are categorized as good, livable, or dilapidated based on their physical structure. Phone ownership includes both landline and mobile phones; non-electric/gas cooking includes households that cook with kerosene, charcoal, crop residue, cowdung, and firewood; non-electric/solar lighting sources include kerosene and other oil.

Table B.1.6: Summary Statistics – Village Directory

Village Characteristics	All Districts		10th-Plan Districts		10th-Plan Districts 150–450 Population	
	2001	2011	2001	2011	2001	2011
Village area (ha)	432.89	413.25	402.66	378.37	178.09	185.39
	(1369.13)	(972.28)	(1721.84)	(1163.87)	(562.11)	(1664.20)
Proportion of area irrigated	0.43	0.40	0.63	0.56	0.51	0.45
	(10.28)	(14.46)	(14.48)	(20.28)	(6.77)	(9.32)
Drinking water facilities (0/1)	0.99	0.99	0.99	0.99	0.99	0.99
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
Educational facilities (0/1)	0.80	0.84	0.76	0.80	0.58	0.64
	(0.40)	(0.36)	(0.43)	(0.40)	(0.49)	(0.48)
# of primary schools	1.12	1.41	1.05	1.31	0.59	0.71
	(1.19)	(1.50)	(1.18)	(1.51)	(0.56)	(0.63)
# of secondary schools	0.12	0.24	0.09	0.22	0.01	0.05
	(0.37)	(0.60)	(0.33)	(0.60)	(0.13)	(0.24)
Medical facilities (0/1)	0.32	0.43	0.29	0.45	0.12	0.28
	(0.47)	(0.50)	(0.45)	(0.50)	(0.32)	(0.45)
# of primary health centers	0.03	0.10	0.03	0.16	0.00	0.01
	(0.18)	(39.55)	(0.17)	(55.95)	(0.07)	(0.11)
# of dispensaries	0.08	0.07	0.07	0.04	0.01	0.01
	(0.50)	(8.48)	(0.48)	(1.24)	(0.14)	(0.14)
Ag credit societies (0/1)	0.14	0.14	0.11	0.10	0.03	0.03
	(0.34)	(0.34)	(0.31)	(0.30)	(0.16)	(0.16)
Banking facilities (0/1)	0.06	0.07	0.06	0.07	0.01	0.02
	(0.24)	(0.26)	(0.23)	(0.25)	(0.11)	(0.13)
Elec power (0/1)	0.77	0.90	0.70	0.91	0.62	0.88
	(0.42)	(0.31)	(0.46)	(0.29)	(0.49)	(0.33)
Elec power, agriculture (0/1)	0.60	0.68	0.52	0.67	0.41	0.59
	(0.49)	(0.47)	(0.50)	(0.47)	(0.49)	(0.49)
Elec power, domestic use (0/1)	0.75	0.89	0.68	0.90	0.59	0.88
	(0.43)	(0.31)	(0.47)	(0.30)	(0.49)	(0.33)
Elec power, all end uses (0/1)	0.49	0.55	0.38	0.52	0.29	0.43
	(0.50)	(0.50)	(0.49)	(0.50)	(0.46)	(0.49)
Post office (0/1)	0.23	0.11	0.20	0.10	0.04	0.03
	(0.42)	(0.32)	(0.40)	(0.30)	(0.20)	(0.18)
Bus service (0/1)	0.75	0.45	0.63	0.35	0.43	0.22
	(0.43)	(0.50)	(0.48)	(0.48)	(0.49)	(0.41)
Rail service (0/1)	0.05	0.02	0.04	0.03	0.01	0.01
	(0.21)	(0.15)	(0.19)	(0.16)	(0.11)	(0.11)
Number of villages	580,643	580,643	290,067	290,067	62,647	62,647

Notes: This table reports means and standard deviations from the 2001 and 2011 Village Directory. The left two columns include all villages that match across Census datasets, for all 27 RGGVY states. The middle two columns include the subset of those villages located in districts eligible for RGGVY under the 10th Plan. The right two columns include only 10th-Plan villages with 2001 populations between 150 and 450 (i.e. our RD bandwidth). Educational facilities include schools and adult training centers. Medical facilities include hospitals, family or maternity welfare centers, clincs, and dispensaries (i.e. small outpatient facilities). Electric power end uses include agriculture, domestic use, and commercial use (the latter is not reported separately in the 2001 Village Directory).

and populations for 483,510 and 567,406 villages, respectively, and include over 1.3 and 1.6 million individual habitations, respectively. Together, these datasets cover over 95 percent of India's villages. Unlike the Census products described above, the habitation census datasets do not include village census codes, meaning that the only village identifier present is the village name. Linking these datasets to our master Census village panel is therefore not straightforward.

We apply a multi-step string matching algorithm in order to merge the 2003 and 2009 habitation census data into our Census panel. First, for each village census code, we construct a list of the various transliterations of that village's name that appear in each of our datasets.[34] Second, we search for exact string matches between villages with census codes and village names that appear in the Water Habitation Census. We repeat this procedure for each version of the village name. Because transliterations of names from Hindi, Bengali, and other Indian languages into English are not standardized from Hindi, Bengali, and other Indian languages, merging on multiple spellings increases the likelihood of an exact string match. We save the matches in a separate file, and then remove them from subsequent steps. Third, we repeat this process using the `reclink` fuzzy string match function in Stata. Finally, for the remaining unmatched villages, we apply the `Masala merge` fuzzy match routine developed by Asher and Novosad (2018). This algorithm computes the Levensthein distance between strings, while accounting for letter substitutions and interpolations common to Hindi transliterations, and produces a set of village name matches.[35]

After completing this matching procedure separately for the 2003 and 2009 habitation datasets, we combine the results into a single village-level dataset to merge with the Census panel. This dataset includes indicators for villages having matched to the 2003 and/or 2009 habitation census; the 2003 and 2009 counts of the number of habitations per village; the 2003 and 2009 populations of the largest habitation in each village; and the 2003 and 2009 populations summed over all constituent habitations. Overall, 86.1 percent of villages match to either the 2003 or 2009 habitation census, and 50.6 percent of matched villages have exactly 1 habitation.[36]

Table B.1.7 reports match rates after each step in this fuzzy merge process. For 91.3 percent of matches (and 88.6 percent of matches with 2001 village populations between 150 and 450), the village population implied by the habitation census (i.e. the sum of all habitation populations in a village) deviates from the village population reported by

[34] These include the 2001 and 2011 PCA, the 2001 VD, the Census 2001–2011 village concordance, and the RGGVY microdata. The latter simply provide another set of variant spellings, in order to increase the chances of an exact string match.

[35] `Masala merge` is more accurate and flexible than standard fuzzy merging routines, such as `reclink`. However we use `reclink` to remove close matches before applying `Masala merge`, as this significantly reduces the time required to execute the computationally intensive `Masala merge` program. We thank the authors for sharing this algorithm with us. A longer description of `Masala merge`, as well as the code, can be found here: `http://www.dartmouth.edu/~novosad/code.html`.

[36] These match rates are very close to those achieved by Asher and Novosad (2018). For villages that match to both the 2003 and 2009 habitation census, there is a correlation of 0.98 between 2003 and 2009 habitation counts.

the Census by less than 20 percent.[37] We suspect large population disparities to indicate erroneous matches, hence we exclude the 11.4 percent of matched villages with disparities greater than 20 percent from all RD specifications.[38] Figure B.1.3 shows habitation matches across the support of 2001 village populations, with lower match rates for smaller villages closer to our RD bandwidth.

Table B.1.7: Summary of Habitation Census Merge Results

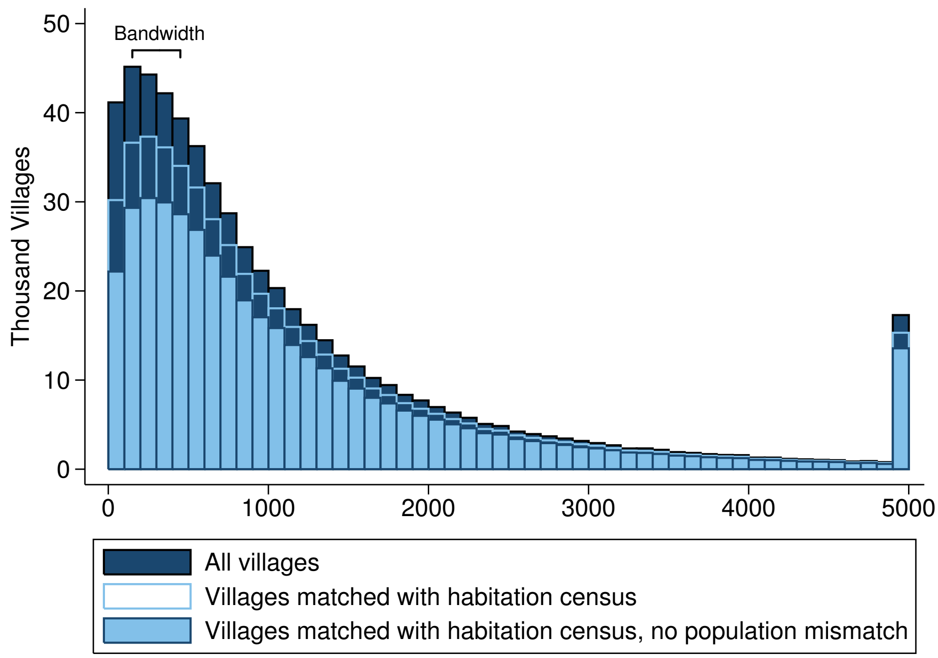
Habitation census match	2003 and 2009	2003 only	2009 only	Unmatched
A. Match rates (all villages)				
Exact matches	0.486	0.065	0.132	0.317
+ `reclink`	0.639	0.050	0.126	0.185
+ `Masala merge`	0.651	0.085	0.125	0.139
B. Match rates (150–450 population)				
Exact matches	0.471	0.046	0.140	0.343
+ `reclink`	0.627	0.048	0.139	0.186
+ `Masala merge`	0.641	0.069	0.137	0.153
C. Summary statistics (all villages)				
Average habitations per village	2.672	2.397	3.352	
Share single-habitation villages	0.571	0.402	0.517	
Share with population mismatch > 20%	0.087	0.478	0.095	
D. Summary statistics (150–450 population)				
Average habitations per village	1.875	1.776	2.057	
Share single-habitation villages	0.654	0.517	0.607	
Share with population mismatch > 20%	0.114	0.545	0.106	

Notes: This table shows results from the habitation merge algorithm described above. Panels A and B report the share of villages that have merged after each step of the algorithm. Panels C and D calculate summary statistics on the subset of Census panel villages that successfully merge to the habitation dataset. Panels A and C report match counts and summary statistics for all 580,643 villages, while Panels B and D report only the 129,453 villages with 2001 populations between 150 and 450. We define a population mismatch over 20 percent to be a matched village for which the sum of constituent habitation populations deviates from both 2001 and 2011 Census populations by at least 20 percent.

[37] Since the habitation censuses were conducted two years after the 2001 Census and two years before the 2011 Census, we should not expect 2003/2009 habitation population to correspond exactly with 2001/2011 village populations — even with 100 percent match accuracy.

[38] Mismatched villages would lead to measurement error in our RD threshold, which relies on correctly identifying villages with a single habitation. Tables B.2.3 and B.2.12 include these villages as sensitivity analysis.

Figure B.1.3: Habitation Merge Results, by 2001 Village Population



Notes: This figure shows a histogram of Indian villages by 2001 population (solid navy), and the subset of villages that we successfully matched with the habitation census (hollow blue). The solid light blue bars show the subset of matched villages with population disparities of less than 20 percent, which we include in our RD analysis. Match rates are lower for smaller villages, yet we still match 84.7 percent of villages with 2001 populations between 150 and 450. Exclude villages with population disparities, this leaves us with 69.6 percent of villages with 2001 populations between 150 and 450.

B.1.6 Socio-Economic Caste Census Dataset

We use microdata from the Socioeconomic and Caste Census (SECC) to examine the effects of RGGVY on income. These data were collected between 2011 and 2012, with the intention of recording data on the socioeconomic status of every single Indian.[39] The SECC is a follow-up to the 2002 Below Poverty Line Census, which only included households that were likely to be below the poverty line.[40] The 2011 SECC was expanded to the entire population, and while it used the Enumeration Blocks from the 2011 Census, it was collected separately, using an electronic tablet-based data collection platform. We obtained a subset of these data from the Ministry of Petroleum and Natural Gas, whose liquid petroleum gas subsidy program, Pradhan Mantri Ujjwala Yojana, uses SECC data

[39]See `http://www.secc.gov.in/aboutusReport` for further details.

[40]We do not use the 2002 data in our analysis. It does not comprehensively survey the entire population, and we have been unable to gain access to the data.

to determine eligibility.[41] As a result, we observe the universe of rural individuals that are eligible for this fuel subsidy program. This includes "households having one of the Deprivations [in the SECC]"[42], where a "deprivation" is a household poverty indicator. In the SECC, to be considered for poverty status, a household must not be automatically deemed either too wealthy or too destitute. After removing these affluent and destitute households, the remaining households are considered for poverty status. That is, only the subset of households without one (or more) affluence indicator(s) and without one (or more) destitution indicator(s) were tested for poverty indicators. Among the tested households, the households found to display at least one poverty indicator were eligible for the program, and therefore included in our dataset.

Specifically, our sample of fuel-subsidy-eligible households was constructed by first removing all households that satisfied at least one affluence indicator, or "exclusion."[43] In particular, households with motorized 2/3/4 wheelers or fishing boats; with mechanized 3-4 wheeler agricultural equipment; with a Kisan credit card (issued by the government to assist farmers) with a credit limit over Rs. 50,000; with a government employee member; operating a non-agricultural enterprise registered with the government; with a member earning more than Rs. 10,000 per month; paying income tax; paying professional tax; with 3 or more rooms with "pucca" (essentially permanent) walls and roof; with a refrigerator; with a landline phone; with more than 2.5 acres of irrigated land and irrigation equipment; with 5 or more acres of irrigated land for two or more crop seasons; **or** owning at least 7.5 acres of land and irrigation equipment were all excluded from our SECC dataset. Next, destitute households were "automatically" included if they were without shelter; were destitute or living on alms; earned income from manually scavenging; belonged to a primitive tribal group; or were engaged in legally released bonded labor.

Our SECC dataset includes all remaining households that satisfy **at least one** poverty (or "deprivation") critertion. These criteria are: households with one or fewer rooms, "kuccha" (non-pucca) walls and roof; households with no adult members between the age of 18 and 59; female-headed households with no adult male member between 16 and 59; households with a "differently-able" member with no other able-bodied member; scheduled caste and scheduled tribe households; households with no literate adults above age 25; or landless households deriving a majority of their income from manual labor. These households yield a dataset the 332 million individuals from 81 million households with no affluence indicators (auto-exclusions), no destitution indicators (auto-inclusions), and at least one poverty indicator (deprivation).[44] This represents roughly half of all households in rural India.

[41] The Ministry of Rural Development, who collected the SECC, are in the process of making the full dataset publicly available. As of now, only district-level aggregates are posted at `http://secc.gov.in/welcome`. We downloaded our data in Excel format from `http://lpgdedupe.nic.in/secc/secc_data.html`.

[42] See `http://www.pmujjwalayojana.com/faq.html`.

[43] See `http://secc.gov.in/reportlistContent`.

[44] We were unable to download SECC data from several districts. While most of these are urban districts with virtually no rural villages, six of these districts contain a nontrivial number of rural villages. These missing districts are: Chamoli, Uttarakhand (1,246 villages); Jalor, Rajasthan (802 villages);

This SECC datset contains individual-level data on age, gender, employment, caste, and marital status; and household-level data on the housing stock, land ownership, asset ownership, income sources, and the household head. Importantly for our analysis, the SECC includes data on income, in the form of an indicator for whether the main income earner in each household receives less than 5,000 rupees per month, or between 5,000 and 10,000 rupees per month.[45] Ideally, we would like to have a continuous measure of income, yet we believe this to be the best measurement of income available in any Indian Census product. We cannot use income data included in other surveys, such as the more comprehensive NSS or ICRISAT's VDSA, because these datasets do not contain a large enough sample of villages near the 300-person RGGVY cutoff to be useful for our identification strategy. The SECC also contains information on the main source of household income, and whether at least one household member has a salaried job. We use these variables to test for the effects of RGGVY on incomes. The SECC also contains information on land and asset ownership and household head characteristics. Because the SECC contains records at the individual level, we can also use it to corroborate overall employment results from the PCA, and to test for heterogeneity by age group.

In particular, we use data each individual supplies on occupation to test for the effects of RGGVY on sectoral changes. Unlike the other data in the SECC, the occupation field in the SECC was not implemented via a drop-down menu in the survey program, but rather was left open-ended. This results in a large number of employment categories, rampant misspellings, and transliterations from Indian languages to English. We attempt to consolidate these to sectors that match the PCA: "agriculture," "household," and "other" labor. To do this, for each district, we use the `strgroup` command in `Stata` to group similar words. We then replace each group of words with the word in the group that appears in the largest number of entries, under the assumption that it is most likely the correct spelling. From this cleaned subset of words, we use regular expressions to categorize agriculture, students, dependents, shopkeepers, retirees, drivers, household laborers, children, and generic workers. We deploy this grouping algorithm a second time, again keeping the most-represented word in each group. Finally, we sort the remaining occupations into three mutually-exclusive worker categories: "agriculture," "household," and "other" (which are as consistent as possible with the PCA coding of these variables). We also create three mutually exclusive non-labor categories: "students," "dependents," and "none" (the latter includes both unemployed individuals and individuals that do not report an occupation).

We merge these SECC data with our village-level Census dataset using a fuzzy match algorithm similar to that described in Section B.1.5. While the SECC data include Census codes for state, district, and block, they do not include the same village codes that are used by the Census. However, upon inspection we discovered that in nearly all districts,

---

Jalpaiguri, West Bengal (768 villages); Dhanbad, Jharkhand (1,760 vilages); Dindigul, Tamil Nadu (481 villages); and Thanjavur, Tamil Nadu (516 villages). Together, these districts contain 5,573 villages, representing less than 1 percent of the total number of villages in our Census dataset.

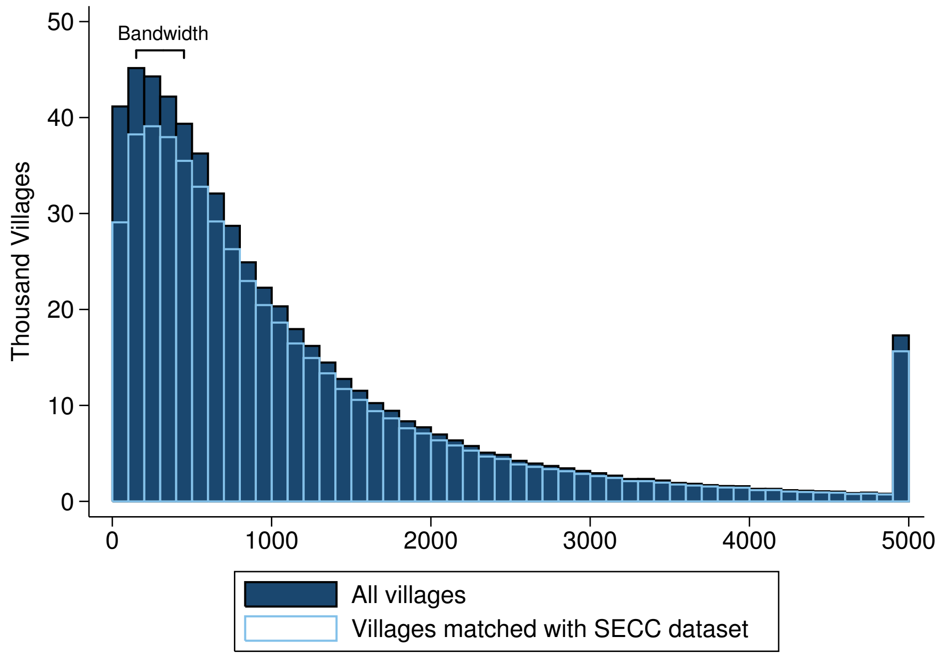
[45] As described above, households whose primary income earner earns above 10,000 rupees per month were automatically excluded from our sample.

Table B.1.8: Summary Statistics – SECC Village-Level Dataset

2011 Village Characteristics	All Districts	All Districts	10th-Plan Districts	10th-Plan Districts 150–450 Population
	Raw Data	Matched	Matched	Matched
A. Share of households				
Monthly income greater than Rs 5,000	0.08	0.08	0.09	0.08
	(0.17)	(0.16)	(0.17)	(0.19)
Main source of income: cultivation	0.27	0.27	0.28	0.34
	(0.31)	(0.30)	(0.30)	(0.35)
Main source of income: non-farm enterprise	0.01	0.01	0.01	0.01
	(0.05)	(0.05)	(0.04)	(0.05)
Main source of income: manual/casual labor	0.64	0.65	0.63	0.57
	(0.33)	(0.33)	(0.33)	(0.37)
Main source of income: domestic service	0.02	0.02	0.02	0.02
	(0.07)	(0.07)	(0.07)	(0.08)
Main source of income: other sources	0.06	0.06	0.06	0.06
	(0.16)	(0.16)	(0.16)	(0.18)
At least one member has salaried job	0.02	0.02	0.02	0.02
	(0.09)	(0.09)	(0.09)	(0.10)
Owning any land	0.36	0.37	0.39	0.47
	(0.32)	(0.31)	(0.31)	(0.35)
Owning irrigation equipment	0.03	0.03	0.04	0.04
	(0.09)	(0.09)	(0.10)	(0.11)
Household head is female	0.16	0.16	0.16	0.18
	(0.14)	(0.14)	(0.15)	(0.18)
Household head is literate	0.45	0.45	0.43	0.43
	(0.24)	(0.23)	(0.23)	(0.26)
Household head has middle school education	0.19	0.19	0.18	0.18
	(0.16)	(0.16)	(0.16)	(0.18)
B. Share of each subpopulation				
Male adult employment rate	0.73	0.73	0.71	0.71
	(0.20)	(0.19)	(0.20)	(0.23)
Female adult employment rate	0.75	0.76	0.73	0.74
	(0.20)	(0.19)	(0.20)	(0.22)
Male youth employment rate	0.09	0.09	0.08	0.09
	(0.15)	(0.15)	(0.16)	(0.18)
Female youth employment rate	0.09	0.09	0.09	0.09
	(0.16)	(0.15)	(0.16)	(0.18)
Adult literacy rate	0.50	0.50	0.48	0.49
	(0.22)	(0.21)	(0.21)	(0.23)
Number of villages	548,489	516,456	255,989	54,481
Number of households	80.6M	74.6M	36.6M	2M
Number of individuals	331.8M	308.5M	158.7M	8.5M
Share of village households included		0.51	0.49	0.50
		(0.26)	(0.25)	(0.28)

Notes: This table reports means and standard deviations from the SECC village-level dataset, which includes all individuals residing in households with at least one poverty indicator in 2011. Variables in Panel A are coded as the share of households in each village, while variables in Panel B are calculated by aggregating up from the individual level. For all summary statistics, the denominator is the total number of households (individuals of a given subpopulation) included in this SECC dataset for each village. The left column reports raw means and standard deviations for all villages in the SECC dataset, while the remaining columns include only SECC villages that match to villages in our Census dataset. We define "adult" to include all individuals at least 16 years old. The last row reports the average fraction of each village's total number of households (per the 2011 Census) that is reported in the SECC dataset, with standard deviations in parentheses.

Figure B.1.4: SECC Merge Results, by 2001 Village Population



Notes: This figure shows a histogram of Indian villages by 2001 population, and the subset of villages that we successfully matched with a village in the SECC dataset (hollow blue). Overall, we match 88.9 percent of Census villages to the SECC dataset, and for 94.1 percent of these matches, at least 10 percent of total households are included in our SECC dataset (because they have at least one poverty indicator). Within our 150–450 population bandwidth, we match 88.7 percent of Census villages to the SECC dataset, and for 91.9 percent of these matches, our SECC dataset includes at least 10 percent of total village households.

SECC village codes simply reindex Census village codes and largely preserve the relative order of village codes within a block. Hence, the first step of our algorithm searches for exact matches on (reindexed) villages codes, allowing for discrepancies between village names of up to 2 characters. Second, for the remaining unmatched villages, we search for exact matches on villages name. Third, for the few remaining unmatched villages, we apply the `Masala merge` algorithm discussed above. Ultimately, we are able to match 94.2 percent of SECC villages to a village in our Census dataset, with very few (i.e., less than 5 percent) of matches relying on the fuzzy `Masala merge` algorithm.[46] However, because our subset of SECC data does not include the full population of Indian villages, only 89.7 percent of villages in our Census dataset match to an SECC village. Figure

[46]This excludes the 2.7 percent of SECC-Census matches for which the SECC data include either a village population or a village household count over 10 percent larger than those reported by the 2011 Census. We also omit this 2.7 of villages from our analysis, as these population and household count discrepancies make these SECC-Census matches suspect.

B.1.4 summarizes the proportion of total villages that match to our SECC dataset in a histogram by population.

Table B.1.8 displays summary statistics from our SECC dataset. The first column includes all villages in our raw subset of the SECC data, while the remaining columns report on SECC villages that match to the Census. They reveal no systematic differences induced by our matching algorithm. The bottom row reveals that proportion of village households included in our SECC dataset (i.e. the proportion of households with at least one poverty indicator) is very similar after restricting the matched sample to only villages in RGGVY 10th-Plan districts and within our 150–450 RD bandwidth. Consistent with Table 2.4.1, villages in our main RD analysis sample are slightly more agricultural, and households in these villages are more likely to own land, more likely to rely on cultivation as a source of income, and less likely to earn income from manual labor. Panel B reports employment rates separately by gender and age, and we see that both male and female youth unemployment rates are quite low across all subsets of the SECC data.

B.1.7 DISE Schools Dataset

In order to include educational outcomes in our analysis, we construct a panel dataset containing the universe of primary (grades 1–5) and upper primary (grades 6–8) schools in India for the 2005–2006 through 2014–2015 school years. These data are publicly available online through the District Information System for Education (DISE)'s School Report Cards website.[47] In total, these data cover 1.68 million schools across over 600,000 villages. However, this yearly school panel is quite unbalanced, with the average school appearing in only 7 out of 10 total years.[48]

The DISE data were originally intended to inform policymakers about the effectiveness of the District Primary Education Programme (DPEP), and data collection for DPEP districts began in 1995. In the early 2000s, DISE was extended to cover the rest of the country in the early 2000s. The information in this dataset is collected by school headmasters or head teachers, and submitted to district- and subsequently state-level authorities before entering the official national system. Quality control is performed by the cluster resource coordinator, and again at the district level.

This dataset includes a large number of variables, although all variables do not appear in every year of the data. We restrict our analysis to student enrollment counts, which appear consistently throughout the full panel. DISE records the number of students, broken

[47] These data can be found here: `http://schoolreportcards.in/SRC-New/`.

[48] While this partly reflects new school construction between 2005 and 2014, we encountered a number of errors when downloading data from the DISE website that contributed to the unbalanced nature of this panel. This means that 7 percent of schools in our DISE panel are missing data for at least 1 year between their first and last reporting years. We have no reason to believe that these errors are anything other than random.

Figure B.1.5: Sample DISE data, 2012–2013

Change Language - Hindi, Marathi, Kannada, Malayalam, Tamil, Telugu, Gujarati, Punjabi

Click to print this page   Previous year report card - 2011-12, 2010-11, 2009-10, 2008-09   Click Here to see **Detailed Report**   **Click here to see RTE Report Card**

SCHOOL REPORT CARD:2012-13*

State RAJASTHAN	**District Name BUNDI**	Grade*
School Code 08230302902	**School Name** GOVT. PS GUJRO KA JHOPRA	6/10
Block Name NAINWA	**Cluster Name** BAMANGAON, GGUPS	
Village Name BAMANGAON	**Name of Head Master** KAMLESH (Post Graduate)	

General Information (PINCODE:323801)

Rural / Urban	Rural	Distance from BRC (Km.)	17	Distance From CRC (Km.)	0
Type of Residential School	NA	Residential School	No	Approachable by all weather roads	Yes
School Category	Primary only	Lowest Class in school	1	Highest class in school	5
Pre-primary Section	No	Total Students (Pre-primary)	0	Total Teachers (Pre-primary)	0
Year of Establishment	2001	Year of recognition	2001	Year of Upgradation from Pri. to U.Pri.	
Management	Local Body	Academic Inspections	13	**School Funds (In Rs.)**	**Recd.** **Expd.**
Type of School	Co-Educational	Shift School	No	Teaching Learning Material fund	500 500
Special School for CWSN	No	No.of visits by Resource teacher for CWSN	0	School Development Fund	5000 5000
No. of Visits by BRC Coordinator	10	No. of Visits by CRC Coordinator	3	Collection from Students	0 0

Staff Category (Primary & Upper Primary only)

Teaching Staff	**Pri.**	**U.Pri.**	Teacher(s) Male	1	Teacher(s) Female	0
Sanctioned	2	0	Part-time instructor (Upper Primary only)	0	Non-teaching Staff	0
In Position	1	0	Teachers Involved in Non-teaching assignments	0	Head Master/Head Teacher	Yes
Contract Teachers	0	0	Avg. working days spent on Non-tch assignments	0	Teachers Received in service Training	1
Graduate & Above	1		Teachers with Professional Qualification	1	Teachers Aged above 55	

School Building, Equipment & Facilities

Number of Building Blocks	**Pucca**	1	**Partially Pucca**	0	**Kuccha**	0	**Tent**	0

Classrooms Require Major Repairs	0	# of Classrooms for Teaching	2	Number of Other Rooms	1
Classrooms Require Minor Repairs	0	Status of School Building	Government	Separate Room for Head Master	No
# of Classrooms in Good Condition	2	Playground	No	Land available for playground	No
Ramp for Disabled Children Needed	Yes	Ramps for Disabled Children Available	Yes	Hand rails for Ramp	No
Medical check-up of Students	Yes	Electricity	No	Computer Aided Learning Lab	No
Furniture for Students	No	# of Computers Available	0	# of Computers Functional	0

Toilets	**Boys**	**Girls**	Library	Yes	# of Books in School Library	158
Total	0	1	Drinking Water Facility	Handpump	Drinking Water Functional	No
Functional	0	1	Measured campus plan prepared	Yes	Boundary Wall	No

Enrolment & Repeaters

Enrolment	2011-12	Total			SC		ST		OBC		Repeaters		CWSN	Muslim	
Grade	All	All	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls		Boys	Girls
I	5	10	5	5	0	0	0	0	3	4	0	0	0	0	0
II	8	7	3	4	0	0	0	0	1	4	0	0	0	0	0
III	5	6	2	4	0	0	0	0	1	4	0	0	0	0	0
IV	1	5	1	4	0	0	0	0	1	4	0	0	0	0	0
V	2	2	1	1	0	0	0	0	1	0	0	0	0	0	0
VI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VII	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
VIII	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total	21	30	12	18	0	0	0	0	7	16	0	0	0	0	0

Incentives (Previous Academic Year)

	Primary only										**Upper Primary only**									
	General		SC Students		ST Students		OBC Students		Muslim Minority		General		SC Students		ST Students		OBC Students		Muslim Minority	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Textbooks	1	0	0	1	0	0	5	14	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Stationary	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Uniform	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Scholarship	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Transport Facility	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Residential Facility	0	0	0	0	0	0	0	0	0	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Key Indicators

Pupil : Teacher Ratio	30	Student : Classroom Ratio	15	% Change in Enr. Over Prev. Year	42.86
% SC Students	0.00	%SC Girls to SC Enrolment	0.00	% Girls Enrolment	60.00
% ST Students	0.00	%ST Girls to ST Enrolment	0.00	% CWSN Enrolment	0.00
%Muslim Students	0.00	% Muslim Girls to Muslim Enrolment	0.00	%Classrooms Require Major Repair	0.00
% OBC Enrolment	76.67	% Repeaters to Total Enrolment	0.00	Teachers with Prof. Qualification	100.00

** Based on availability of Ramp, Playground, Boundary Wall, Drinking Water, Boys Toilet, Girls Toilet, Library, PTR≤30 at Primary Schools, PTR ≤35 at Upper Primary Level, SCR≤30 at Primary Schools, SCR ≤35 at Upper Primary Level and Classroom-Teacher Ratio≥1.*

CWSN : Children with Special Needs, BRC : Block Resource Center, CRC : Cluster Resource Center , NA : Not Applicable, na : Not Available

    *As on 30th September 2012

Notes: This figure displays the variables available in the 2012–2013 DISE dataset, from the Government Private School Gujro Ka Jhopra in Rajasthan.

down by grade and gender, for each school-year pair.[49] Table B.1.9 summarizes enrollment counts from this dataset at the village level. Across India, there are approximately 200 children in primary school (grades 1–5) and approximately 70 children in upper primary school (grades 6–8) on average per village. There are slightly more boys on average than girls in primary school, and this gap grows among upper primary students. On average, we observe between 1.9 and 2.3 schools per village in the DISE data; we observe between 536,330 and and 559,442 unique villages in each year, reflecting the unbalanced panel nature of these data.

DISE also reports information on various school facility characteristics. The DISE dataset includes, for example, what building materials each school is made out of, whether a school has toilets, whether a school is private or public, the funding received by the school, and the number of teachers employed by each school. Figure B.1.5 provides an example of the data included in the 2012–2013 school year's DISE dataset for a randomly selected school, Government Private School Gujro Ka Jhopra in Rajasthan.

In constructing our DISE panel dataset, we link schools across years based on their unique numerical school codes. DISE also reports the state, district, block, and village name of each school. We use this information to match schools to villages in our main analysis dataset. DISE does not report village census codes, so we are forced to undertake a fuzzy match procedure similar to that used to merge habitations into our village dataset (see Section B.1.5). We begin by building a concordance between school codes and village names as reported in the DISE dataset. In many cases, one school will be associated with multiple village names in different years — this can occur both when villages are split or combined, and when single villages are spelled differently across datasets.[50] We conservatively allow for multiple village spellings for a single school, in order to maximize our chances of matching to a village name in the census.

We search for exact string matches between villages with census codes and village names that appear in the DISE data, following the exact string match, `reclink`, and `Masala merge` algorithm described in Section B.1.5. After establishing a list of census villages matched to DISE villages, we remove duplicates by reintroducing the school code data. In cases where one school code matches to multiple census villages, we select the match which occurs most frequently.[51] We use the results of this matching procedure to construct two datasets: first, a dataset of schools matched to census villages, which we use to generate our school-level RD results in the main text and in Appendix B.2.8; and second, a village-level dataset that sums enrollment across schools in a village to a single village-level observation, which we use for sensitivity analysis below.

[49]In some years, these variables are also broken down by students who belong to Scheduled Castes and Scheduled Tribes, as well as students with special needs.

[50]We only observe each school once per school year. These different linkages occur when we observe schools across different years.

[51]For example: if school A matches to village 1 in 2005, village 1 in 2006, and village 2 in 2007, we keep the match that links school A with village 1.

Table B.1.9: Summary Statistics – DISE Schools Dataset

	School year beginning in									
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
A. Students enrolled per village										
Grades 1–5, male	100.3	104.8	104.2	103.3	100.8	101.0	100.9	97.5	69.0	72.8
	(144.6)	(152.2)	(154.5)	(153.8)	(157.7)	(158.6)	(162.9)	(163.7)	(149.3)	(142.4)
Grades 1–5, female	91.8	97.3	97.3	97.3	95.4	95.6	95.3	92.2	64.4	68.4
	(129.1)	(136.7)	(139.6)	(137.8)	(142.5)	(142.4)	(145.9)	(146.6)	(131.2)	(127.7)
Grades 6–8, male	33.7	36.3	37.5	38.2	39.0	41.0	43.2	44.0	33.7	37.5
	(85.7)	(88.6)	(90.9)	(89.9)	(93.4)	(94.7)	(97.9)	(101.8)	(94.2)	(93.7)
Grades 6–8, female	27.8	31.1	32.9	34.5	36.3	38.8	41.4	42.6	31.9	35.8
	(74.0)	(78.5)	(81.8)	(82.5)	(87.2)	(89.4)	(92.1)	(96.7)	(87.1)	(89.2)
Total	253.6	269.4	271.9	273.4	271.5	276.4	280.9	276.3	198.9	214.6
	(391.3)	(412.4)	(424.5)	(421.8)	(440.2)	(445.9)	(460.4)	(469.8)	(431.6)	(425.1)
B. Observations per year										
Number of villages	536, 330	542, 135	552, 028	526, 765	558, 708	562, 035	564, 904	567, 565	568, 382	559, 442
Number of schools	0.99M	1.06M	1.11M	1.09M	1.16M	1.21M	1.24M	1.26M	1.32M	1.28M
Schools per village	1.9	1.9	2.0	2.1	2.1	2.1	2.2	2.2	2.3	2.3
	(1.8)	(2.0)	(2.1)	(2.1)	(2.1)	(2.2)	(2.3)	(2.3)	(2.5)	(2.4)

	Unique villages	Years per village	Unique schools	Years per school
C. Observation counts	667, 514	8.30	1, 678, 545	6.98
		(2.86)		(3.32)

Notes: This table reports means and standard deviations for school enrollment, by village by year. Panel A sums enrollment across all schools within a village, for boys/girls and for primary/upper primary levels. Panel B reports the number of unique villages and schools in each annual cross-section of this unbalanced panel. Panel C counts the number of unique villages and schools present across all ten years, as well as the average number of observations per cross-sectional unit. Note that a "village" in these data is defined by DISE school code identifiers, and this does not exactly correspond to the Census definition of a village.

Table B.1.10 reports the proportion of census villages that we successfully match to the DISE dataset, after each stage of the matching algorithm. We achieve an overall match rate of 67 percent across all villages, and 58 percent for villages with populations within our main RD bandwidth of 150–450 people. Note that unlike the habitation match described above, we should not expect to be able to match 100 percent of villages *ex ante*: not every village in India is home to a school. According to the Village Directory, only 84 percent of villages actually had schools in 2011. Of these villages, we are able to successfully match 74 percent of villages to schools, and 66 percent of villages with populations between 150 and 450. Figure B.1.6 displays our match results graphically, demonstrating both that small villages are less likely to have schools than large villages, and that our matching algorithm is more successful in larger villages than in smaller villages. Our final match rates for villages *containing schools* (according to the Village Directory) within our preferred RD bandwidth are comparable to those achieved with habitation merges. Our RD analysis on school enrollment ultimately includes only 41 percent of the villages in our RD regressions for Census outcomes, because only 51 percent of single-habitation villages in our main RD sample match to the DISE data. We also must exclude the 10 percent of school-village matches with missing enrollment data in either 2011–2012 or 2005–2006, as our main RD specification uses 2011–2012 enrollment as an outcome and 2005–2006 enrollment as a control variable.[52]

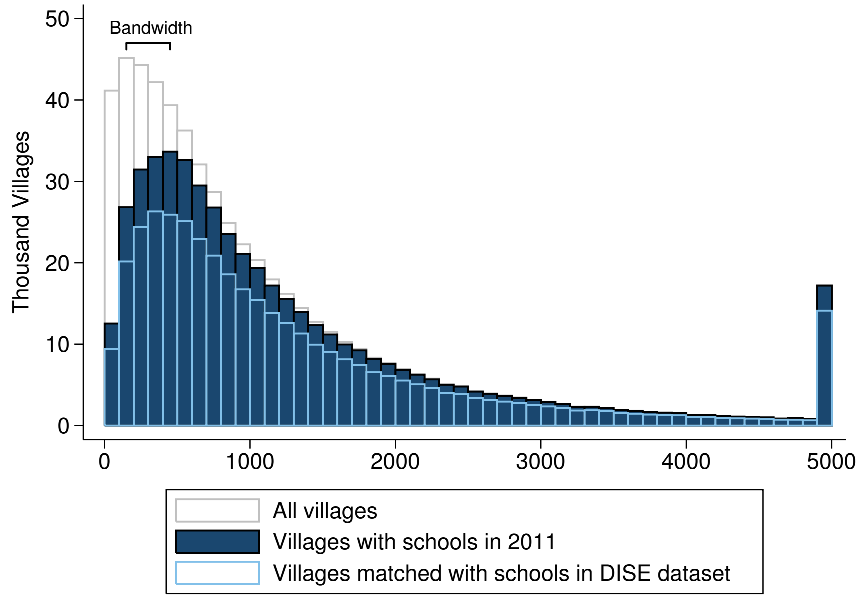
B.1.8 Village Counts by Dataset

Our master dataset includes RGGVY district-level implementation details, village characteristics from the 2001 and 2011 Census, and a count of habitations per village as determined by the fuzzy merge to the habitation census. Each merge between different data sources is imperfect, and Table B.1.11 shows the number of villages present after each step in this merging process. We focus on single-habitation villages in RGGVY 10th-Plan districts (in the middle and right columns), in order to ensure the internal validity of our RD design. Our analysis of nighttime brightness requires an additional merge between this panel dataset and village-specific brightness for the restricted 12-state geospatial sample (as indicated by the fourth row of Table B.1.11).[53] Likewise, our analyses of SECC outcomes and school enrollment each require an additional (fuzzy) merge with the SECC and DISE datasets, respectively (as indicated by rows 4 and 6 of Table B.1.11). We do not enforce the lights match when running regressions on non-spatial outcomes, nor we do not enforce SECC or schools matches when running regressions on non-SECC/non-enrollment outcomes.

[52]In many cases, these data appears "missing" because schools likely did not exist during the 2005–2006 school year. We perform sensitivities on the choice of school years in Section B.2.8.

[53]Village shapefiles include attribute tables with 2001 Census codes, allowing a straightforward merge between Census datasets and shapefiles.

Figure B.1.6: School Merge Results, by 2001 Village Population



Notes: This figure shows a histogram of Indian villages by 2001 population (hollow white), the subset of villages with schools 2011 (as reported by the 2011 Village Directory; solid navy), and the subset of villages that we successfully matched with a school in the the DISE dataset (hollow blue). After adjusting for the share of small villages without schools, we achieve match rates closer to those shown in Figure B.1.3.

Table B.1.10: Summary of School Merge Results

	All Villages	150–450 Population
A. Match rates		
Exact matches	0.371	0.324
+ `reclink`	0.624	0.541
+ `Masala merge`	0.667	0.578
B. Summary statistics		
Average number of schools per matched village	3.594	1.972
Share of villages with school in 2011	0.844	0.741
Match rates for villages with school in 2011	0.734	0.654

Notes: This table shows results from the school merge algorithm described above, and it is analogous to Table B.1.7. Panel A reports the share of villages that have merged after each step of the algorithm. Panel B calculates summary statistics on the subset of Census panel villages that successfully merge to the DISE schools dataset. The first column reports match counts and summary statistics for all 580,643 villages, while the second column considers only the 129,453 villages with 2001 populations between 150 and 450. Our merge algorithm does not restrict matches to only the subset of villages reported to have schools in the 2011 Census.

Table B.1.11: Count of Villages by Merged Dataset

Number of Villages	Total	RGGVY 10th-Plan	RGGVY 10th-Plan Single-Hab.	RGGVY 10th-Plan Single-Hab. 150–450
Raw Census datasets (village-level)	$>593{,}000$			
2001–2011 Census panel	580,643	290,067		
2001–2011 Census panel + habitations	499,799	218,841	115,444	29,765
2001–2011 Census panel + habitations + SECC	380,528	193,493	100,827	25,942
2001–2011 Census panel + habitations + lights	309,105	137,519	70,790	18,686
2001–2011 Census panel + habitations + schools	297,330	140,196	68,872	15,215†

Notes: All village counts exclude Goa and the 7 Union Territories, which were not covered under RGGVY. Rows labeled "habitations" count only villages that we can successfully match to the 2003 or 2009 census of habitations, with population disparities of less than 20 percent. The row labeled "SECC" includes only villages that we successfully link to our SECC dataset. The row labeled "lights" counts only villages that we successfully match to village shapefiles. This excludes 5 states for which we do not have shapefiles (Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, Sikkim) and 5 states for which the shapefiles are not correlated with reported village areas (Assam, Himachal Pradesh, Jammu and Kashmir, Uttar Pradesh, Uttarakhand). The row labeled "schools" includes only villages that we successfully link to a school in the DISE dataset. We conduct our RD analysis on villages in the right-most column, which restricts the middle column to villages with 2001 populations between 150 and 450.

† Only 81 percent of village-school matches report nonmissing enrollment data for both 2011 and 2005, further reducing our sample size for regressions on enrollment outcomes.

B.2 Empirics

In this appendix, we provide a variety of robustness and falsification exercises to complement the results in the main text. We first discuss sensitivity of our nighttime brightness results, and then move to the economic outcomes of interest.

B.2.1 Nighttime Brightness: RD Robustness

B.2.1.1 Sample and outcome variable definition

As discussed above in Section B.1.2, because of missing or low-quality shapefiles, we are forced to drop ten states from our nighttime lighting analysis sample. This leaves us with twelve states with 10th-Plan RGGVY districts, which contain over 18,000 single-habitation villages within our RD bandwidth. Our main specification assigns village brightness based on each village's brightest pixel. We focus on the brightest pixel due to the typical organizational structure of South Asian villages, which have concentrated inhabited regions surrounded by agricultural lands. Since RGGVY is targeted at electrifying public places and homes, the brightest pixel will best reflect brightness that is attributable to RGGVY infrastructure upgrades.

We also linearly project village brightness on the values in adjacent years in order to remove year-to-year measurement error and focus on more permanent year-to-year changes in electricity use (see Equation (B.1) in Appendix B.1.3). Table B.2.1 demonstrates the degree to which this cross-year calibration refines our estimates, while Table B.2.2 compares these weight-averaged linear projections to unweighted 3- and 5-year averages. These unweighted averages yield very similar point estimates, however we see that the linear projections provide greater precision. In addition, Table B.2.2 shows that our results are almost identical when we average village brightness across all pixels (i.e., Columns (2), (4), and (6)), as opposed to using the maximum brightness.

We have chosen to use NOAA's average lights product, as opposed to the more processed stable nighttime lights. While the latter images exclude fires and other sporadic lights, they are also less likely to detect the low levels of lighting we might expect from small recently electrified villages. Table B.2.3 compares RD estimates using each data product, and we see that the stable lights actually yield larger RD coefficients than the (preferred) average visible lights. Columns (2) and (4) of Table B.2.3 display the results of an additional sensitivity test, by including 2,373 single-habitation villages whose official village Census populations differ from their matched habitation populations by over 20 percent. Such substantial population disparities suggest that these Census villages may be wrongly matched to single-habitation villages, hence we have excluded them from all other specifications. As expected, including these potentially erroneous matches attenuates our RD estimates, by having introduced measurement error in the RD indicator variable.

Table B.2.1: RD Sensitivity – Raw vs. Projected Lights

2011 village brightness	Raw Lights	Projected (2010–2012)	Projected (2009–2013)
	(1)	(2)	(3)
1[2001 pop $\geq$ 300]	0.0788	0.1408**	0.1493**
	(0.0698)	(0.0680)	(0.0603)
2001 population	−0.0002	−0.0006	−0.0008
	(0.0009)	(0.0007)	(0.0007)
1[2001 pop $\geq$ 300] $\times$ 2001 pop	0.0012	0.0007	0.0008
	(0.0011)	(0.0009)	(0.0008)
2001 Control	Yes	Yes	Yes
State FEs	Yes	Yes	Yes
RD bandwidth	150	150	150
Number of observations	18,686	18,686	18,686
Number of districts	130	130	130
Mean of dependent variable	6.244	6.368	6.370
R^2	0.673	0.746	0.766

Notes: This table shows results from estimating Equation (2.1), using raw and projected 2011 brightness. Column (1) uses raw 2011 lights as the dependent variable. Column (2) uses a linear projection of 2011 lights on 2010 and 2012 values. Column (3) reproduces our preferred specificaton from Table 2.5.2, using a linear projection of 2011 lights on 2009, 2010, 2012, and 2013 values. For each column, the 2001 lights control uses the analogous projection, for consistency within each regression. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.2: RD Sensitivity – Alternative Lights Variables

2011 village brightness	Projected		3-Year Average		5-Year Average	
	Max	Mean	Max	Mean	Max	Mean
	(1)	(2)	(3)	(4)	(5)	(6)
1[2001 pop $\geq$ 300]	0.1493**	0.1386**	0.1163*	0.1146*	0.1219*	0.1188**
	(0.0603)	(0.0556)	(0.0671)	(0.0622)	(0.0631)	(0.0592)
2001 population	−0.0008	−0.0008	−0.0004	−0.0004	−0.0004	−0.0004
	(0.0007)	(0.0007)	(0.0008)	(0.0008)	(0.0008)	(0.0007)
1[2001 pop $\geq$ 300]$\times$ 2001pop	0.0008	0.0006	0.0007	0.0004	0.0006	0.0004
	(0.0008)	(0.0007)	(0.0009)	(0.0009)	(0.0009)	(0.0008)
2001 Control	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150	150	150
Number of observations	18,686	18,686	18,686	18,686	18,686	18,686
Number of districts	130	130	130	130	130	130
Mean of dependent variable	6.370	6.077	7.047	6.711	6.735	6.411
R^2	0.766	0.762	0.753	0.751	0.761	0.760

Notes: This table shows results from estimating Equation (2.1), using alternative definitions of 2011 brightness as the outcome variable. Columns (1)–(2) use a linear projection of 2011 lights on 2009, 2010, 2012, and 2013 values, with Column (1) reproduces our preferred specification from Table 2.5.2. Columns (3)–(4) use unweighted averages of 2010–2012 values, while Columns (5)–(6) use unweighted averages of 2009–2013 values. Columns (1), (3), and (5) assign village brightness based on the brightest pixel, whereas Columns (2), (4), and (6) average village brightness across all pixels contained in the village boundary. We construct 2001 lights controls to be analogous to their respective outcome variables, for consistency within each regression. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.3: RD Sensitivity – NOAA DMSP–OLS Datasets

2011 village brightness	Average Visible Lights		Stable Lights	
	(1)	(2)	(3)	(4)
1[2001 pop $\geq$ 300]	0.1493**	0.1170**	0.1810**	0.1419**
	(0.0603)	(0.0530)	(0.0726)	(0.0680)
2001 population	−0.0008	−0.0004	−0.0012	−0.0008
	(0.0007)	(0.0006)	(0.0008)	(0.0007)
1[2001 pop $\geq$ 300] $\times$ 2001 pop	0.0008	0.0004	0.0008	0.0005
	(0.0008)	(0.0007)	(0.0010)	(0.0009)
Forced population match	Yes	No	Yes	No
2001 Control	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150
Number of observations	18,686	21,059	18,686	21,059
Number of districts	130	130	130	130
Mean of dependent variable	6.370	6.333	4.873	4.825
R^2	0.766	0.775	0.782	0.789

Notes: This table shows results from estimating Equation (2.1), using alternative NOAA DMSP–OLS lights data. Columns (1)–(2) show the average visible lights data, which is our preferred measure of nighttime brightness (Column (1) is reproduced from Table 2.5.2). Columns (3)–(4) show results for NOAA's more processed stable lights product. Columns (2) and (4) include villages that match to the 2003 and/or 2009 habitation census datasets, but have population disparities of greater than 20 percent (indicating potentially erroneous matches). For each specification, 2011 and 2001 brightness values are constructed using a linear projection on the brightness values of adjacent years, using their respective NOAA data products. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2.1.2 Bandwidths

Because RD designs rely on data close to the threshold to estimate a local average treatment effect (LATE) *at* the threshold, any sensitivity of RD estimates to bandwidth selection could undermine their internal validity. Our preferred specification uses a 150-person bandwidth on either side of the 300-person cutoff, including villages of between 150 and 450 people. Figure B.2.1 estimates Equation (2.1) using bandwidths ranging from 50 to 300 people. This demonstrates that our RD point estimate is not sensitive to bandwidth selection, as our point estimates are quite stable. At the same time, they lose significance at narrower bandwidths, which reduce the RD sample size.

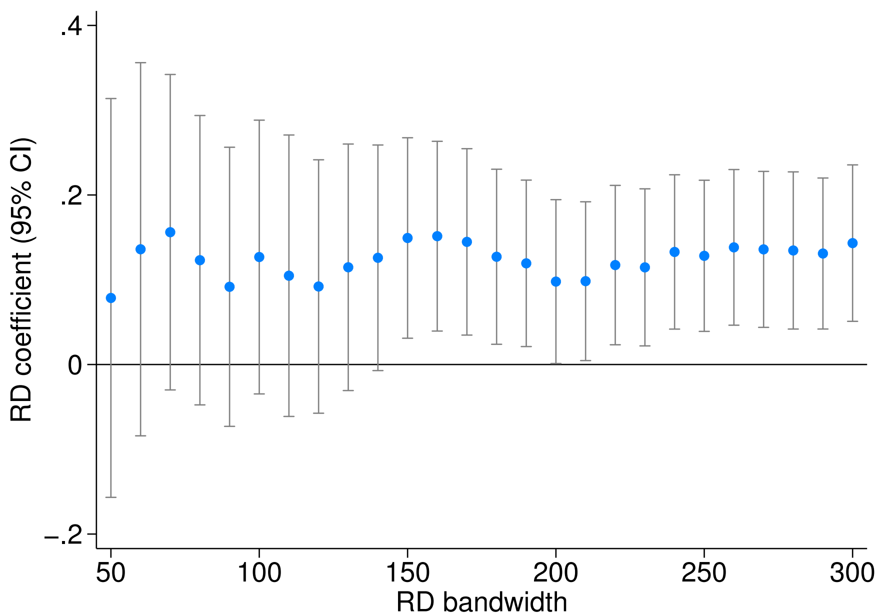
We choose a 150-person bandwidth for two primary reasons. First, we want to ensure that our estimation window does not overlap with two other program eligibility cutoffs — the 100-person cutoff implemented under the 11th Plan of RGGVY, and the 500-person cutoff used in the PMGSY road-building program. Both thresholds may have led to increases in nighttime brightness with the potential to confound our estimates at the 300-person cutoff, either directly through electrification or indirectly due to the economic benefits of road infrastructure. Second, we want to exclude very small villages (i.e. population less than 50) in our RD sample. These villages are likely quite different than villages of 200–400 inhabitants.

As an alternative strategy, we implement on the optimal RD bandwidth procedure formalized by Imbens and Kalyanaraman (2012). Using this technique, we derive an optimal bandwidth of 137 using a uniform kernel, 162 using an Epanechnikov kernel, and 174 using a triangular kernel. Figure B.2.1 shows that had we chosen any of these three bandwidths, our results would have be nearly identical.

B.2.1.3 Functional form

Our preferred RD specification excludes higher-order polynomials, following Gelman and Imbens (2017). We control for only a linear function of the running variable, allowing the slope to differ on either side of the RD threshold. While this has become standard practice for implementing RD designs, we also test for sensitivity of our estimates to higher-order polynomials. Table B.2.4 compares our preferred specification (in Column (1)) to specifications with 2nd- and 3rd-order terms. Our RD estimate is robust to the inclusion of a quadratic function of the running variable, but we lose precision when we include higher-order terms. Figure B.2.2 presents these results graphically, and we see that the 3rd-order polynomial appears to be affected by observations far from the RD threshold (as Gelman and Imbens (2017) warn might occur with higher-order polynomials). The right-hand panels of Figure B.2.2 present the same three specifications using a 174-person bandwidth, also reported in Columns (2), (4), and (6) of Table B.2.4. This is our largest Imbens-Kalyanaraman optimal bandwidth (which we calculated using uniform, Epanechnikov, and triangular kernels). We see that extending the bandwidth affects the curvature of the 3rd-order polynomial estimates, while the fitted linear and quadratic curves appear mostly unchanged.

Figure B.2.1: RD Sensitivity – Nighttime Brightness, Bandwidth



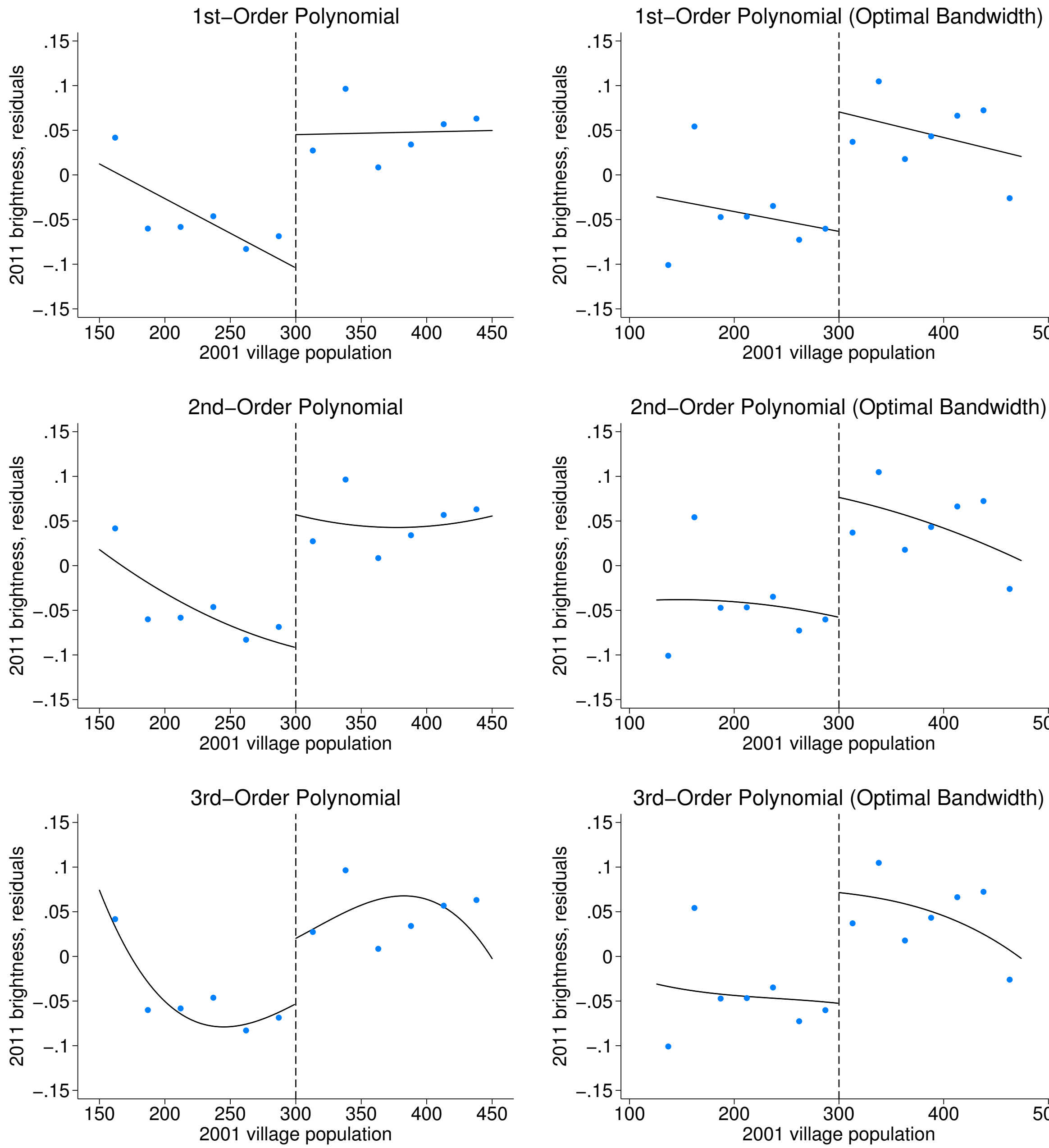
Notes: This figure presents our bandwidth sensitivity analysis for Equation (2.1), estimated separately on bandwidths ranging from 50 (i.e., 250–350 people) to 300 (i.e., 0–600 people). Each dot represents the point estimate on the RD discontinuity at a given bandwidth around the 300-person cutoff, with 95 percent confidence intervals clustered at the district level. Our chosen bandwidth of 150 includes villages with populations between 150 and 450. The optimal bandwidth, calculated using the algorithm proposed by Imbens and Kalyanaraman (2012), is 137 using a uniform kernel, 162 using an Epanechnikov kernel, and 174 using a triangular kernel.

Table B.2.4: RD Sensitivity – Higher Order Polynomials

2011 village brightness	Linear		Quadratic		Cubic	
	(1)	(2)	(3)	(4)	(5)	(6)
1[2001 pop ≥ 300]	0.1493**	0.1337**	0.1497**	0.1372**	0.0742	0.1261
	(0.0603)	(0.0549)	(0.0607)	(0.0550)	(0.0987)	(0.0864)
2001 population	−0.0008	−0.0002	−0.0011	−0.0021	0.0001	−0.0019
	(0.0007)	(0.0005)	(0.0016)	(0.0013)	(0.0016)	(0.0015)
1[2001 pop ≥ 300] × 2001 pop	0.0008	−0.0001	0.0014	0.0036	0.0014	0.0036
	(0.0008)	(0.0006)	(0.0029)	(0.0023)	(0.0029)	(0.0024)
$(2001\ \text{population})^2$			−0.0000	−0.0000	−0.0000	−0.0000
			(0.0000)	(0.0000)	(0.0000)	(0.0000)
$(2001\ \text{population})^3$					−0.0000	−0.0000
					(0.0000)	(0.0000)
2001 Control	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
RD bandwidth	150	174	150	174	150	174
Number of observations	18,686	21,551	18,686	21,551	18,686	21,551
Number of districts	130	130	130	130	130	130
Mean of dependent var	6.370	6.344	6.370	6.344	6.370	6.344
R^2	0.766	0.775	0.766	0.775	0.766	0.775

Notes: This table compares our main RD specification to two specificaions with higher-order polynomials, as presented graphically in Figure B.2.2. Columns (1)–(2) estimate our main specification using a linear function of the running variable, 2001 population. Columns (3)–(4) use a quadratic function of population, while Columns (5)–(6) use a cubic function of population. Each regression includes all villages meeting the above sample criteria with 2001 populations in the RD bandwidth (either our preferred bandwidth of 150 or our largest optimal bandwidth of 174), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.2: RD Sensitivity – Nighttime Brightness, Higher Order Polynomials



Notes: This figure presents our RD for 2011 nighttime brightness, estimated using 1st-, 2nd-, and 3rd-order polynomials. The three figures in the left-hand column correspond to the regressions in Table B.2.4, and the top-left panel reproduces Figure 2.5.4. The three figures in the right-hand column use the Imbens-Kalyanaraman optimal bandwidth of 174 (which contains 21,551 villages). Blue dots show average residuals from regressing 2011 nighttime brightness on 2001 nighttime brightness and state fixed effects. Each dot contains approximately 1,600 villages, averaged in 25-person population bins. Linear terms are estimated separately on each side of the 300-person threshold, and higher-order terms are restricted to be the same on each side of the threshold. Plots include all within-bandwidth, single-habitation, 10th-Plan villages, for the 12 states with available village shapefiles that correspond to Census village areas (with a correlation above 0.35).

B.2.1.4 Fixed effects and 2001 controls

RD designs do not rely on fixed effects or controls for identification, but their inclusion can greatly improve efficiency (Lee and Lemieux (2010)). Our main specification controls for both state fixed effects and the 2001 level of nighttime brightness. Table B.2.5 estimates this specification with and without this 2001 control, as well as with no fixed effects, state fixed effects, and district fixed effects. We see that the RD estimates without the 2001 control are noisy and imprecise, meaning that our RD requires this baseline control to find a statistically detectable effect. Stated differently, we find that RGGVY eligibility has led to a statistically significant increase in brightness *conditional* on 2001 levels of brightness, but we are unable to statistically detect the *unconditional* effect of RGGVY eligibility on 2011 brightness. This makes sense: nighttime brightness prior to RGGVY is very heterogeneous; conditioning on the pre-period level dramatically improves the signal-to-noise ratio. This is not an uncommon practice in papers using remotely sensed data, as Jayachandran et al. (2017) also rely on pre-period controls in order to detect treatment effects. Table B.2.5 also reveals that our estimates are not sensitive to the inclusion of fixed effects. We have chosen to include state fixed effects in our main specification, in order to control for large differences in nighttime brightness across our 12 sample states (see Table B.2.6).

Table B.2.7 introduces additional 2001 village controls to our main RD specification. Our RD point estimates are not affected by the inclusion of these pre-RGGVY controls, even though two (literacy rate and presence of road) have statistically significant associations with 2011 nighttime brightness. Interestingly, 2001 electric power indicator variables are poor predictors of 2011 brightness. This is not surprising, considering that these Census variables do not capture the intensity of electrification within the village.

B.2.1.5 Pre-RGGVY lights

Covariate smoothness across the threshold is a key identifying assumption in RD designs. If pre-existing factors were to jump discontinuously at the threshold, this would have the potential to confound our RD estimates. Figure B.2.3 and Table B.2.8 show that lights from 2001–2004 (i.e., *before* the announcement of RGGVY) do not exhibit any significant breaks at the 300-person cutoff. This supports our assumption that selection into eligibility around the 300-person cutoff was as-good-as-random.

B.2.1.6 Standard errors

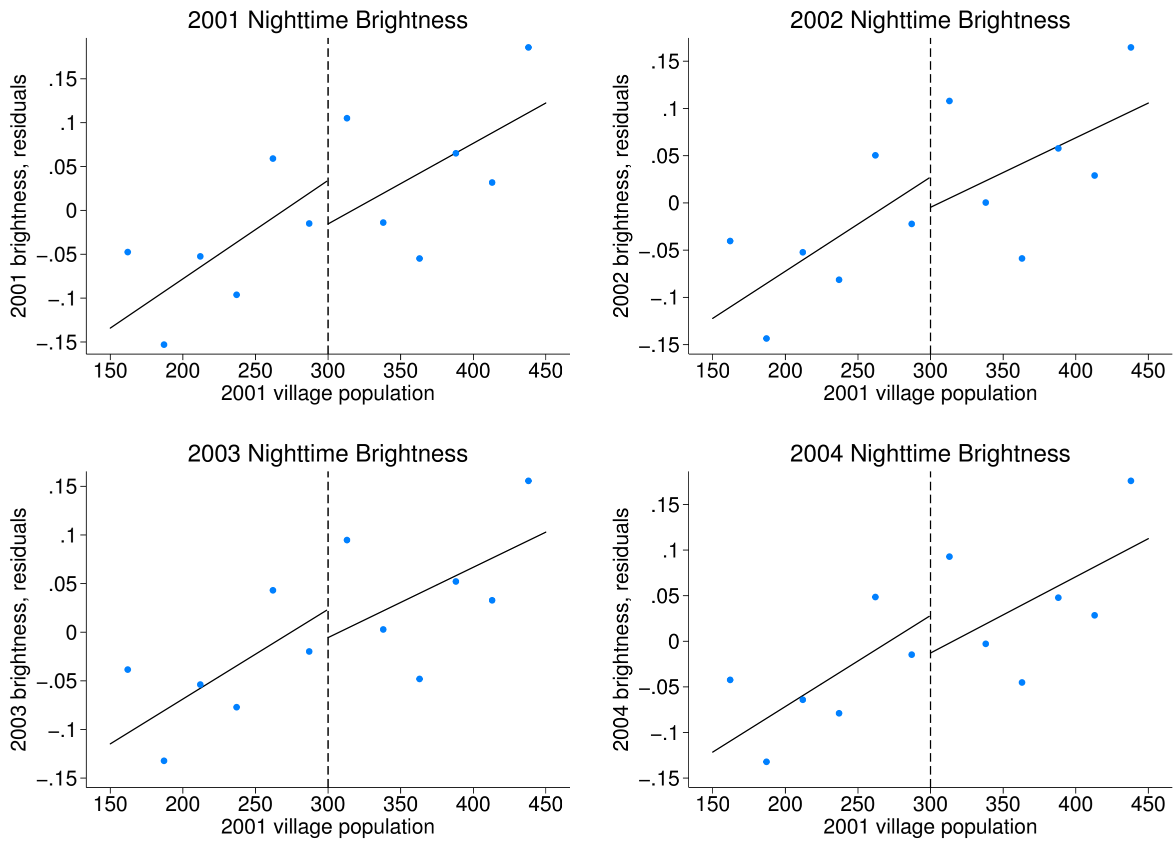
Finally, our main specification clusters standard errors at the district level. This allows for arbitrary dependence in the error structure between any two villages in the same district. Because RGGVY projects were approved based on district-specific implementation plans (or DPRs), and funds were awarded to district implementing agencies, we allow for within-district dependence to control for any unobserved factors affecting RGGVY implementation. Since districts are geographically contiguous area, this also accounts for spatial correlations between nearby villages.

Table B.2.5: RD Sensitivity – Fixed Effects and 2001 Control

	2011 village brightness					
	(1)	(2)	(3)	(4)	(5)	(6)
1[2001 pop $\geq$ 300]	0.1002	0.0836	0.0745	0.1451**	0.1493**	0.1292**
	(0.1201)	(0.1149)	(0.1144)	(0.0635)	(0.0603)	(0.0568)
2001 population	0.0006	0.0007	0.0009	−0.0012*	−0.0008	−0.0005
	(0.0013)	(0.0013)	(0.0012)	(0.0007)	(0.0007)	(0.0006)
1[2001 pop $\geq$ 300] $\times$ 2001 pop	0.0008	0.0005	−0.0001	0.0012	0.0008	0.0005
	(0.0017)	(0.0016)	(0.0015)	(0.0008)	(0.0008)	(0.0008)
2001 Control	No	No	No	Yes	Yes	Yes
State FEs	No	Yes	No	No	Yes	No
District FEs	No	No	Yes	No	No	Yes
RD bandwidth	150	150	150	150	150	150
Number of observations	18,686	18,686	18,686	18,686	18,686	18,686
Number of districts	130	130	130	130	130	130
Mean of dependent var	6.370	6.370	6.370	6.370	6.370	6.370
R^2	0.001	0.062	0.169	0.748	0.766	0.791

Notes: This table shows results from estimating Equation (2.1), using with and without fixed effects and the 2001 control. For each regression, the dependent variable is the maximum village brightness for 2011. Column (5) reproduces our preferred specification from Table 2.5.2. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.3: RD Sensitivity – Pre-RGGVY Brightness



Notes: These four figures estimate our preferred RD specification using nighttime brightness for 2001–2004, corresponding to the regressions in Table B.2.8. Each graph uses a different year's nighttime brightness as the dependent variable, after having linearly projected these values on adjacent years. Blue dots show average residuals from regressing that year's nighttime brightness on state fixed effects (without any other controls). Each dot contains approximately 1,600 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 18,686 single-habitation villages between 150–450 people, in 10th-Plan districts, for the 12 states with available village shapefiles that correspond to Census village areas (with a correlation above 0.35).

Table B.2.6: Nighttime Brightness by State

RD Sample State	2001 Average Village Brightness	2011 Average Village Brightness	Villages in RD Sample
Andhra Pradesh	5.021	7.049	1,020
Bihar	3.347	4.559	2,555
Chhattisgarh	3.999	6.807	364
Gujarat	5.111	6.007	333
Haryana	12.329	17.625	31
Jharkhand	4.211	6.537	549
Karnataka	5.495	8.002	2,747
Madhya Pradesh	4.774	5.191	1,568
Maharashtra	4.946	5.563	315
Orissa	4.258	5.681	471
Rajasthan	4.757	7.039	4,138
West Bengal	4.413	6.568	4,595
Total	4.583	6.492	18,686

Notes: This table shows the average 2001 and 2011 brightness by state, for villages in our main RD sample. The 2001 and 2011 brightness variables used in this table are the same linear projections used in our main specification.

Table B.2.9 reports standard errors on our RD point estimate for alternative assumptions about the error structure. We see that clustering by census block yields slightly larger standard errors than clustering by district.[54] We also calculate Conley "spatial HAC" standard errors, which are robust to spatial dependencies between villages within a given geographic radius, as well as heteroscedasticity and autocorrelation.[55] Table B.2.9 shows that spatial standard errors estimated with a 50-km and 250-km bandwidth are smaller than our preferred standard error estimates.

[54]Census block is the administrative unit that is smaller than district but larger than village. The only administrative unit larger than district is state. Because our nightlights regressions only include 12 states, we do not cluster at the state level for fear of bias resulting from having too few clusters (Cameron and Miller (2015)).

[55]We use code from Fetzer (2014) to implement this Conley HAC procedure. This code can be found online: `http://www.trfetzer.com/conley-spatial-hac-errors-with-fixed-effects/`. It is in turn based on code from Hsiang (2010), itself based on theory from Conley (1999) and Conley (2008). Because our RD regression is cross-sectional and does not include multiple observations for each village, the autocorrelation ("AC") component of the spatial HAC estimator is not relevant for our purposes. Our spatial standard errors apply a uniform kernel, which yields very similar estimates to those generated using a linear Bartlett kernel.

Table B.2.7: RD Sensitivity – 2001 Village Controls

	2011 village brightness					
	(1)	(2)	(3)	(4)	(5)	(6)
1[2001 pop ≥ 300]	0.1489**	0.1484**	0.1510**	0.1431**	0.1468**	0.1808***
	(0.0603)	(0.0605)	(0.0606)	(0.0615)	(0.0604)	(0.0573)
2001 population	−0.0009	−0.0008	−0.0008	−0.0008	−0.0008	−0.0013**
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0006)
1[2001 pop ≥ 300] × 2001pop	0.0009	0.0008	0.0008	0.0009	0.0009	0.0016**
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0007)
2001 literacy rate	0.9232***					
	(0.2937)					
2001 share SC/ST		−0.1558				
		(0.1181)				
2001 share of area irrigated			0.2959			
			(0.2699)			
2001 road prsent (0/1)				0.2690***		
				(0.0630)		
2001 elec, any use (0/1)					0.0072	
					(0.1343)	
2001 elec, all uses (0/1)						0.1593
						(0.1333)
2001 Brightness	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150	150	150
Number of observations	18,686	18,686	18,669	18,358	18,647	16,907
Number of districts	130	130	130	130	130	130
Mean of dependent var	6.370	6.370	6.370	6.398	6.365	6.451
R^2	0.767	0.766	0.766	0.766	0.766	0.773

Notes: This table introduces 2001 village controls to our main RD specification. SC/ST refer to official Scheduled Caste and Scheduled Tribe designations, while the 2001 presence of a road indicator includes both paved and mud roads. 2001 electric power indicators consider three distinct end uses: domestic, agricultural, and commercial. Electricity for any use (all uses) indicates whether *any one* (*all*) of these three end-use sectors had electric power in 2001. Each regression includes all villages meeting the above sample criteria with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.8: RD Sensitivity – Pre-RGGVY Brightness

Village brightness	2001	2002	2003	2004
	(1)	(2)	(3)	(4)
1[2001 pop $\geq$ 300]	-0.0496	-0.0319	-0.0290	-0.0411
	(0.0736)	(0.0721)	(0.0649)	(0.0648)
2001 population	0.0011	0.0010	0.0009	0.0010
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
1[2001 pop $\geq$ 300] $\times$ 2001 pop	-0.0002	-0.0003	-0.0002	-0.0002
	(0.0011)	(0.0011)	(0.0010)	(0.0010)
State FEs	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150
Number of observations	18,686	18,686	18,686	18,686
Number of districts	130	130	130	130
Mean of dependent variable	4.512	4.448	3.603	3.918
R^2	0.053	0.055	0.058	0.062

Notes: This table shows results from estimating our RD specification using brightness outcomes for years prior to the announcement of RGGVY. For each regression, the dependent variable is the maximum village brightness for a given year, after applying a linear projection onto brightness in adjacent years. Unlike Equation (2.1), these specifications do not control for pre-RGGVY brightness. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.9: RD Sensitivity – Alternative Standard Errors

	Clustered by Block	Clustered by District	Spatial HAC (50 km)	Spatial HAC (250 km)
Standard Error on RD coefficient	(0.0680)**	(0.0603)**	(0.0602)**	(0.0542)***

Notes: This table shows robustness of our RD point estimate ($\hat{\beta}_1$ from Table 2.5.2) to alternative standard error assumptions. Column (1) clusters standard errors by census block, which is the administrative unit between district and village. Column (2) clusters by district, which is our preferred method. Columns (3)–(4) apply standard errors that are robust to heteroscedasticity and spatial correlation, with bandwidths of 50 and 250 km. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2.2 Nighttime Brightness: Validity Tests

In addition to the robustness checks above, we also conduct placebo, randomization, and falsification tests to corroborate our results for nighttime brightness. Section 2.5 in the main text discuss our placebo test and randomization inference procedures, and Figure 2.5.5 reveals that our observed increase in nighttime brightness is very unlikely to result from spurious correlations in the relationship between brightness and village populations.

Section 2.5 describes an additional falsification exercise, where we demonstrate that nighttime brightness only increased for the subset of villages within 10th-Plan eligible districts containing a single habitation. We find no increase in brightness at the 300-person threshold for villages in 11th-Plan districts (for which the eligibility cutoff was reduced from 300 to 100 people) or for villages with more than one habitation (for which the total village population does not correspond to the population used to determine eligibility). Figure 2.5.6 graphically compares our main RD results to three "falsification" samples: multi-habitation villages in 10th-Plan districts, single-habitation villages in 11th-Plan districts, and multi-habitation villages in 11th-Plan districts. Table B.2.10 presents the same results in regression format. As expected, none of these alternative samples exhibits evidence of a positive discontinuity at the 300-person cutoff. This provides further evidence that the RGGVY program is what is driving the increase in nighttime brightness for 10th-Plan, single-habitation villages.

B.2.3 Nighttime Brightness: Timing

As a final validity test, we look at changes in nighttime brightness over time. Since we are attributing the increase in nighttime brightness to RGGVY eligibility, year-on-year changes in differential brightness at the RD cutoff should be consistent with the rollout of the RGGVY program. We should expect our RD estimates to increase incrementally over time, for two reasons. First, RGGVY project funds under the 10th Plan were disbursed gradually between 2005 and 2010. Because we estimate the average effect on brightness across 128 eligible districts, our RD estimate should increase in magnitude as more districts received RGGVY funding. (There is an additional lag between the date that a district received RGGVY funding and the rollout of project implementation across its constituent villages). Second, the effects of infrastructure improvements on observed brightness are likely not immediate. For a village that received transformer upgrades and additional household electric connections in 2009, we might expect observed brightness to increase incrementally between 2010–2011, as villages invest in appliances that use electricity and emit light.

We test for these gradual rollout and investment effects in Figure B.2.4 and Table B.2.11. As expected, we see that the RD point estimate increases monotonically from 2006 to 2011. Hence, the RD effect that we detect from NOAA's satellite images is consistent with the incremental rollout and takeup of electricity use due to the RGGVY program.

Table B.2.10: RD Sensitivity – Falsification Tests

2011 village brightness	10th-Plan Single-Hab.	10th-Plan Multi-Hab.	11th-Plan Single-Hab.	11th-Plan Multi-Hab.
	(1)	(2)	(3)	(4)
1[2001 pop ≥ 300]	0.1493**	−0.0326	−0.0109	−0.0781
	(0.0603)	(0.0751)	(0.0595)	(0.0986)
2001 population	−0.0008	−0.0001	−0.0001	0.0007
	(0.0007)	(0.0007)	(0.0005)	(0.0009)
1[2001 pop ≥ 300] × 2001 pop	0.0008	0.0011	0.0005	−0.0005
	(0.0008)	(0.0008)	(0.0008)	(0.0008)
2001 Control	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150
Number of observations	18,686	10,304	17,385	11,382
Number of districts	130	116	174	149
Mean of dependent variable	6.370	5.398	5.587	5.010
R^2	0.766	0.793	0.756	0.611

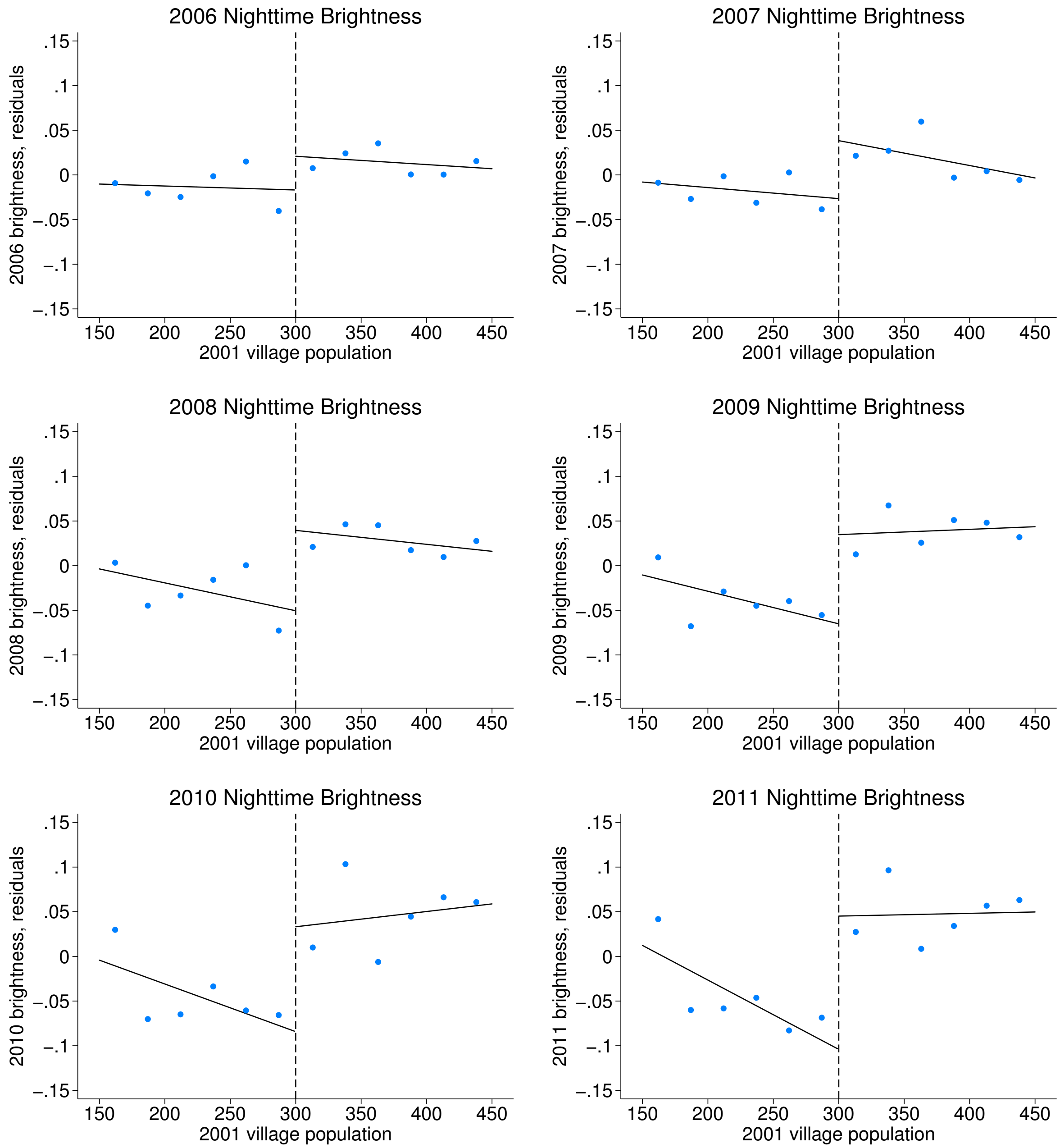
Notes: This table compares our main RD specification in Column (1) to three separate samples for which RGGVY's 300-person eligibility threshold should not be relevant. Columns (2) and (4) estimate Equation (2.1) using villages with multiple habitations, for which the running variable (village population) does not correspond to the habitation populations that determined village eligibility. Columns (3) and (4) estimate Equation (2.1) using villages that were eligibile for RGGVY under the 11th Plan, which moved the eligibility cutoff from 300 to 100 people. Figure 2.5.6 presents these three falsification tests graphically. Each regression includes all villages meeting the above sample criteria with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.11: RD Sensitivity – Brightness by Year

Village brightness	2006	2007	2008	2009	2010	2011
	(1)	(2)	(3)	(4)	(5)	(6)
1[2001 pop $\geq$ 300]	0.0377	0.0649**	0.0900**	0.0999**	0.1175*	0.1493**
	(0.0251)	(0.0321)	(0.0352)	(0.0427)	(0.0632)	(0.0603)
2001 population	−0.0000	−0.0001	−0.0003	−0.0004	−0.0005	−0.0008
	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.0008)	(0.0007)
1[2001 pop $\geq$ 300] $\times$ 2001 pop	−0.0000	−0.0002	0.0002	0.0004	0.0007	0.0008
	(0.0003)	(0.0003)	(0.0004)	(0.0006)	(0.0009)	(0.0008)
2001 Control	Yes	Yes	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
RD bandwidth	150	150	150	150	150	150
Number of observations	18,686	18,686	18,686	18,686	18,686	18,686
Number of districts	130	130	130	130	130	130
Mean of dependent var	3.844	4.344	5.063	5.123	7.542	6.370
R^2	0.914	0.867	0.854	0.801	0.758	0.766

Notes: This table shows results from estimating Equation (2.1), using brightness outcomes from varying years. For each regression, the dependent variable is the maximum village brightness for a given year, after applying a linear projection onto brightness in adjacent years. Each regression includes all single-habitation villages in 10th-Plan districts with 2001 populations in the RD bandwidth (a 150-person bandwidth includes villages with 2001 populations between 150 and 450), for the 12 states with available village shapefiles that match to Census village areas with a correlation above 0.35. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.4: RD on Nighttime Brightness Over Time



Notes: These six figures estimate our preferred RD specification using nighttime brightness for 2006–2011, corresponding to the regressions in Table B.2.11. Each graph uses a different year's nighttime brightness as the dependent variable, after having linearly projected these values on adjacent years. The lower-right graph reproduces Figure 2.5.4. Blue dots show average residuals from regressing that year's nighttime brightness on 2001 nighttime brightness and state fixed effects. Each dot contains approximately 1,600 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 18,686 single-habitation villages between 150–450 people, in 10th-Plan districts, for the 12 states with available village shapefiles that correspond to Census village areas (with a correlation above 0.35).

B.2.4 Census Outcome Results: RD Robustness

Next, we run a series of analogous RD robustness tests for the range of Census outcomes results that we report in the main text.

B.2.4.1 Sample and outcome variables

We normalize our main village-level results in Table 2.5.3 by either village population (by gender) or by the share of village households. This is because our RD specification is cross-sectional, and most village-level outcomes will vary mechanically with the running variable (e.g., number of 2011 female agricultural workers will increase mechanically in 2001 village population). However, this normalization relies on the assumption that village-level demographics were unaffected by RGGVY program eligibility. Figure B.2.5 shows RD plots for 2011 village population and the size of the 0–6 age cohort, as reported in Panel A of Table 2.5.3. We see no evidence that RGGVY eligibility changed either the total village population or the fertility rate.

As an additional sensitivity, Table B.2.12 reproduces our main RD results from Table 2.5.3, including villages with population disparities in the habitation census of over 20 percent. After adding nearly 3,962 villages that were potentially miscategorized as having a single habitation, our RD estimates are nearly identical, although slightly more precise due to a greater sample size. In contrast, including villages with population disparities attenuates our nighttime brightness results (see Table B.2.3). As we see no corresponding attenuation in Table B.2.12, this underscores the lack of evidence of any economically meaningful impacts of RGGVY electrification for villages close to the cutoff.

B.2.4.2 Bandwidths

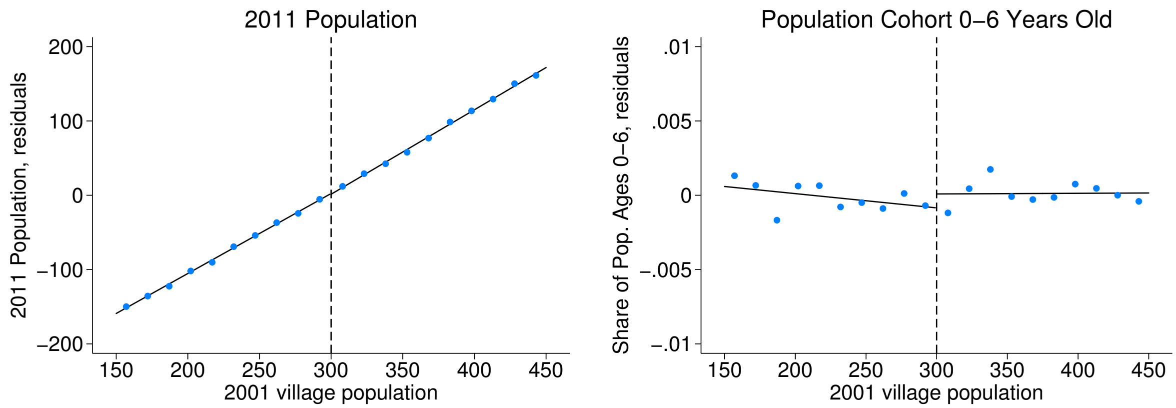
Section B.2.1.2 discusses our choice of a 150-person RD bandwidth around RGGVY's 300-person 10th-Plan cutoff, which avoids overlapping with the 100-person 11th-Plan cutoff and the 500-person PMGSY eligibility cutoff. However, we also demonstrate that our RD results for nighttime brightness are not sensitive to our choice of bandwidth (see Figure B.2.1). We present analogous bandwidth sensitivities in Figure B.2.6, for eight census outcomes reported in Table 2.5.3. This demonstrates that our RD results for these Census outcomes (male and female labor shares; mud floors; asset ownership; share of village area planted and irrigated) are not overly sensitive to our choice of bandwidth, for bandwidths between 100 and 250 (above which includes very small villages). Figure B.2.6 shows the range of feasible symmetric RD bandwidths, which stops at 300 due to our 300-person RD threshold. The optimal bandwidth calculations (following Imbens and Kalyanaraman (2012)) vary across each outcome, ranging from 130 to 353.

Table B.2.12: RD Sensitivity – Census Outcomes, No Forced Population Match

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	−0.2473	(2.335)	[−4.824, 4.329]	273.01
0–6 cohort / total population	0.0010	(0.001)	[−0.001, 0.003]	0.15
Average household size	−0.0063	(0.012)	[−0.029, 0.016]	5.12
Literacy rate	−0.0025	(0.002)	[−0.006, 0.001]	0.56
B. Labor outcomes				
Male agricultural workers / male pop	−0.0069***	(0.002)	[−0.012, −0.002]	0.42
Female agri. workers / female pop	−0.0055	(0.004)	[−0.013, 0.002]	0.29
Male household workers / male pop	−0.0008	(0.001)	[−0.002, 0.000]	0.01
Female household workers / female pop	−0.0015	(0.001)	[−0.003, 0.001]	0.01
Male other workers / male pop	0.0035*	(0.002)	[−0.000, 0.007]	0.10
Female other workers / female pop	−0.0016	(0.002)	[−0.005, 0.002]	0.05
C. Asset ownership				
Share of households with telephone	0.0025	(0.005)	[−0.007, 0.012]	0.53
Share of households with TV	0.0015	(0.004)	[−0.006, 0.009]	0.25
Share of households with bicycle	0.0007	(0.004)	[−0.007, 0.008]	0.49
Share of households with motorcycle	−0.0008	(0.003)	[−0.006, 0.004]	0.13
Share of households without assets	0.0016	(0.004)	[−0.006, 0.009]	0.23
D. Housing stock				
Share of households w/ elec/gas cooking	0.0008	(0.002)	[−0.004, 0.005]	0.06
Share of households w/ kerosene lighting	0.0023	(0.006)	[−0.009, 0.014]	0.48
Share of households with mud floors	0.0056	(0.004)	[−0.002, 0.013]	0.74
Share of households with thatched roof	−0.0037	(0.005)	[−0.013, 0.005]	0.23
Share of households dilapidated	−0.0024	(0.003)	[−0.008, 0.003]	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	0.0051	(0.010)	[−0.015, 0.025]	0.74
1/0 Post office in village	0.0017	(0.003)	[−0.005, 0.008]	0.03
1/0 Ag credit societies in village	0.0005	(0.003)	[−0.006, 0.007]	0.02
1/0 Water from tubewell in village	−0.0033	(0.011)	[−0.024, 0.017]	0.45
Share of village area irrigated	−0.0075*	(0.005)	[−0.016, 0.001]	0.35
Share of village area planted	0.0018	(0.006)	[−0.009, 0.013]	0.58

Notes: This table includes the identical set of regressions in Table 2.5.3, including villages with large population disparities with the Habitation Census (i.e., villages whose matched habitation populations disagree with official Census populations by over 20 percent). The RD bandwidth now includes 33,727 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. All specifications also include state fixed effects. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.5: RD Reduced Form – 2011 Village Population



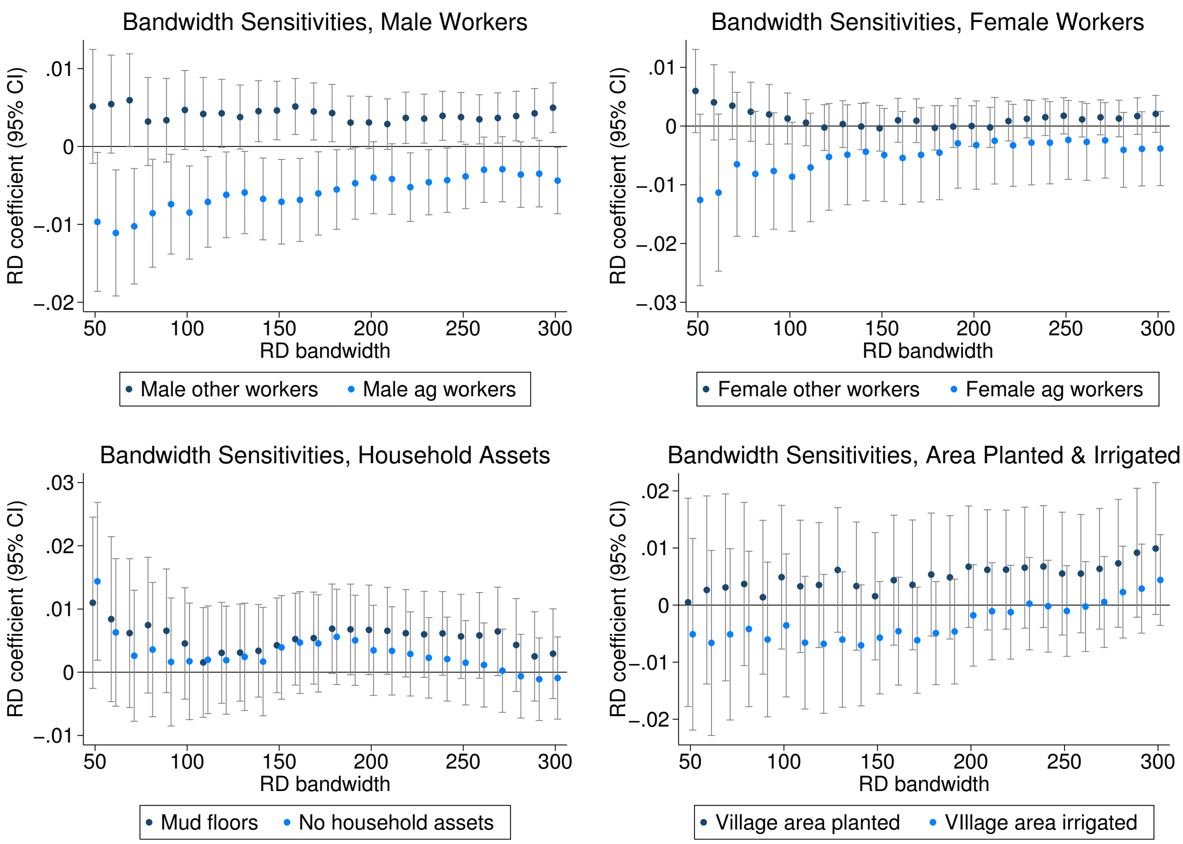
Notes: This figure presents results for RD regressions of 2011 village population (left) and the share of the 2011 population less than 7 years old (right). These correspond to the first two rows of Table 2.5.3. Blue dots show average residuals from regressing the 2011 outcome on and state fixed effects and the 2001 level of the outcome (except for the left panel, where 2001 population is the running variable and hence not an additional control). Each dot contains approximately 1,500 villages, averaged in 15-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

B.2.4.3 Functional form

Section B.2.1.3 discusses our choice to exclude higher-order polynomials from our preferred RD specification (following Gelman and Imbens (2017)). Table B.2.13 tests for sensitivity of our RD regressions on village-level census outcomes to the inclusion of a quadratic function of the running variable. This yields point estimates and confidence intervals that are very close to the linear RD specification in Table 2.5.3. Figure B.2.7 presents a subset of these results graphically, where the share of male agricultural workers remains the census outcome that is most obviously discontinuous at the 300-person threshold.

Table B.2.14 re-estimates Equation (2.1) using weighted least squares, weighting observations by their distance from the RD threshold. We define village weights as $w_v \equiv 1 - \frac{|P_v - 300|}{150}$, where P_v is the 2001 village population. These results confirm that upweighting villages close to the RD threshold does not significantly change our results.

Figure B.2.6: RD Sensitivity – Census Outcomes, Bandwidths



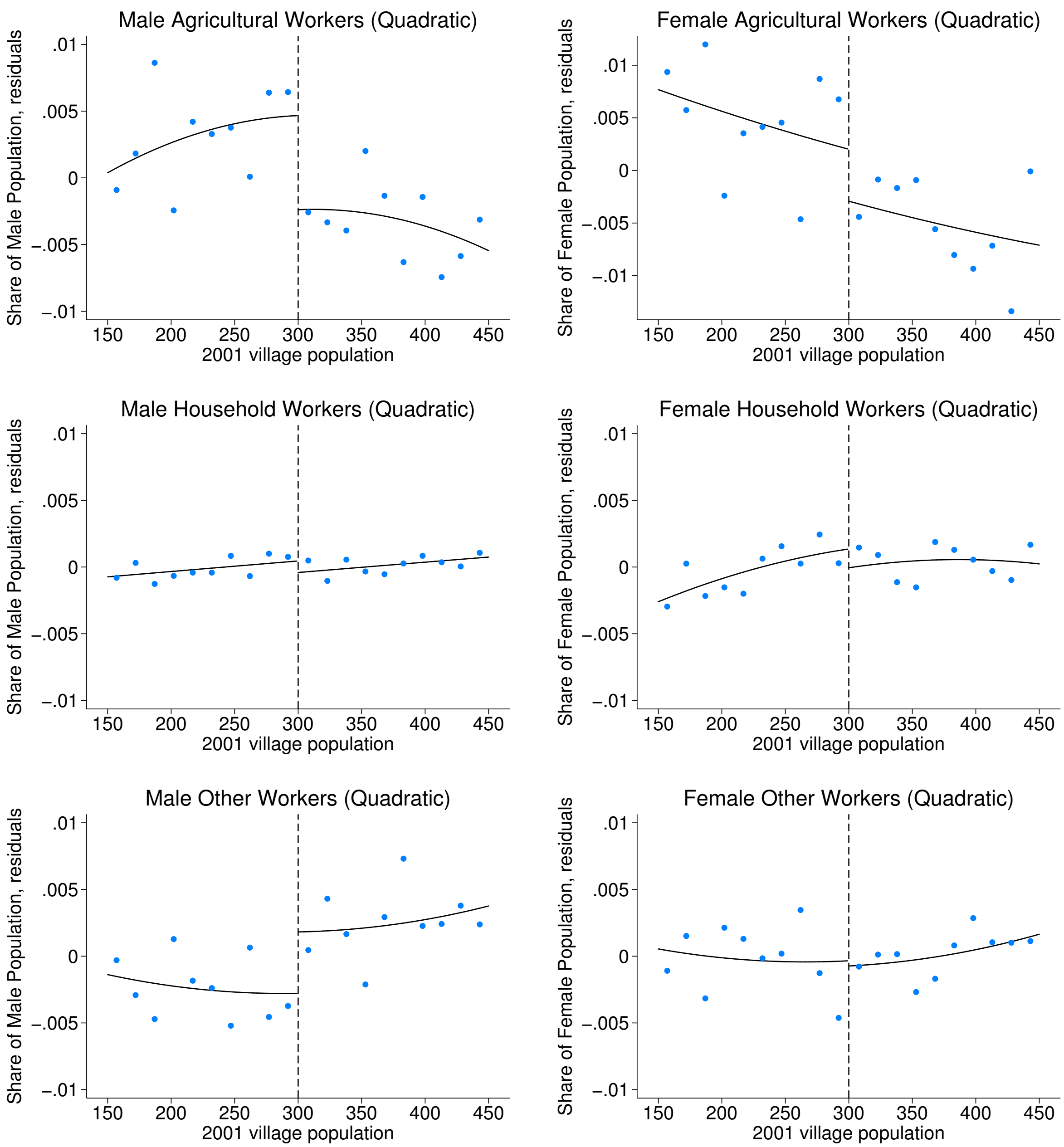
Notes: This figure presents our bandwidth sensitivity analysis for eight separate outcomes in Table 2.5.3. For each outcome, we estimate Equation (2.1) separately on bandwidths ranging from 50 (i.e., 250–350 people) to 300 (i.e., 0–600 people). Each dot represents the point estimate on the RD discontinuity at a given bandwidth around the 300-person cutoff, with 95 percent confidence intervals clustered at the district level. Our chosen bandwidth of 150 includes villages with populations between 150 and 450. The optimal RD bandwidth varies for each outcome, ranging from 130 to 354 for the eight outcomes shown here (calculated using the algorithm proposed by Imbens and Kalyanaraman (2012), using uniform, Epanechnikov, and triangular kernels).

Table B.2.13: RD Sensitivity – Census Outcomes, Quadratic in Population

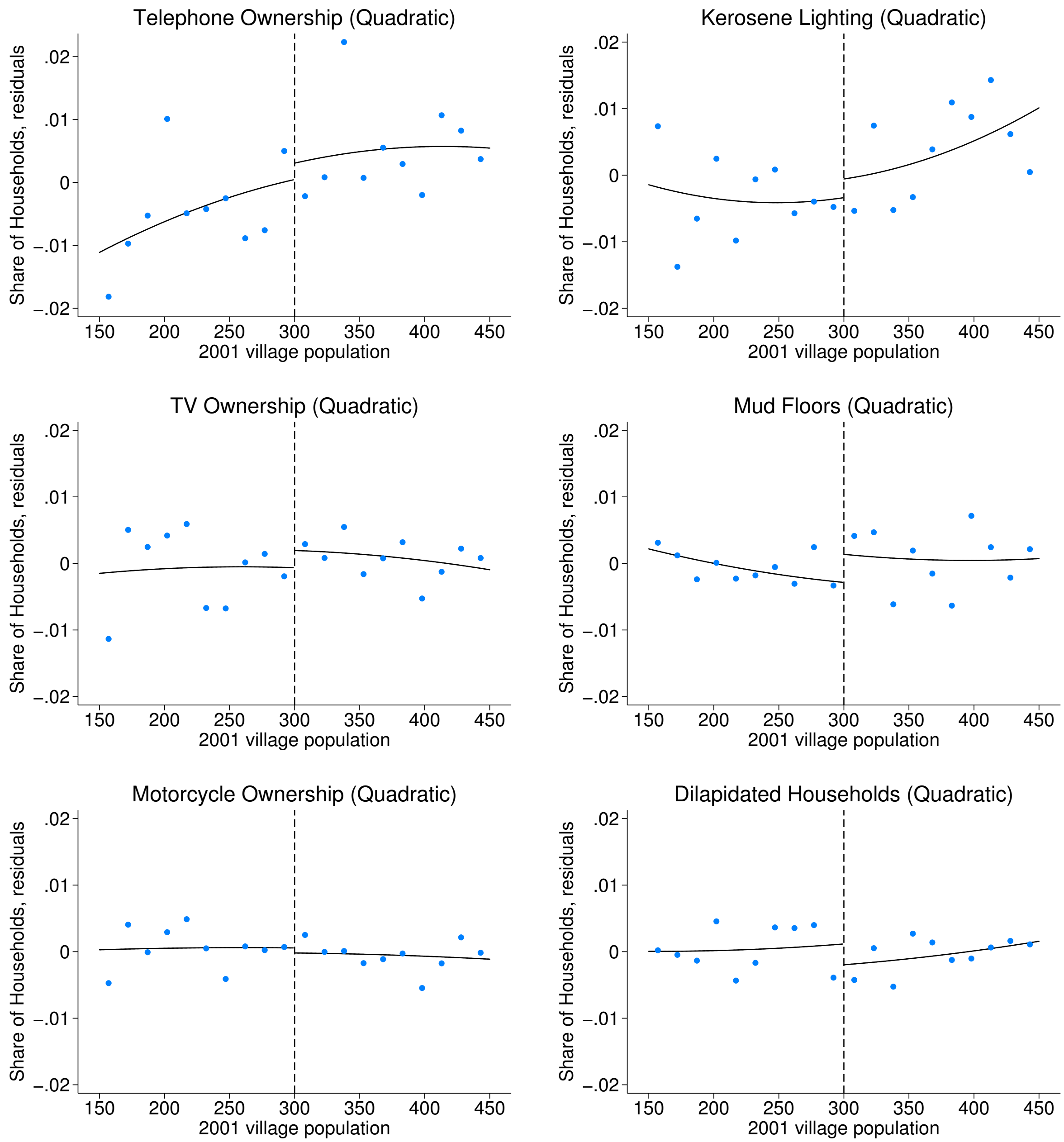
2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	−0.8933	(2.510)	[−5.813, 4.026]	271.09
0–6 cohort / total population	0.0009	(0.001)	[−0.001, 0.002]	0.14
Average household size	−0.0054	(0.013)	[−0.030, 0.019]	5.13
Literacy rate	−0.0025	(0.002)	[−0.006, 0.002]	0.57
B. Labor outcomes				
Male agricultural workers / male pop	−0.0071**	(0.003)	[−0.012, −0.002]	0.42
Female agri. workers / female pop	−0.0051	(0.004)	[−0.013, 0.003]	0.29
Male household workers / male pop	−0.0009	(0.001)	[−0.002, 0.000]	0.01
Female household workers / female pop	−0.0015	(0.001)	[−0.004, 0.001]	0.01
Male other workers / male pop	0.0047**	(0.002)	[0.001, 0.008]	0.10
Female other workers / female pop	−0.0003	(0.002)	[−0.004, 0.004]	0.05
C. Asset ownership				
Share of households with telephone	0.0028	(0.005)	[−0.008, 0.014]	0.54
Share of households with TV	0.0026	(0.004)	[−0.005, 0.010]	0.26
Share of households with bicycle	−0.0015	(0.004)	[−0.010, 0.007]	0.50
Share of households with motorcycle	−0.0008	(0.003)	[−0.006, 0.004]	0.13
Share of households without assets	0.0037	(0.004)	[−0.004, 0.012]	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	0.0005	(0.003)	[−0.005, 0.006]	0.07
Share of households w/ kerosene lighting	0.0029	(0.006)	[−0.009, 0.015]	0.48
Share of households with mud floors	0.0041	(0.004)	[−0.003, 0.012]	0.73
Share of households with thatched roof	−0.0034	(0.005)	[−0.013, 0.006]	0.23
Share of households dilapidated	−0.0031	(0.003)	[−0.009, 0.003]	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	−0.0002	(0.011)	[−0.022, 0.022]	0.75
1/0 Post office in village	0.0015	(0.004)	[−0.006, 0.009]	0.03
1/0 Ag credit societies in village	0.0015	(0.004)	[−0.006, 0.009]	0.02
1/0 Water from tubewell in village	−0.0025	(0.011)	[−0.024, 0.019]	0.44
Share of village area irrigated	−0.0056	(0.005)	[−0.015, 0.004]	0.34
Share of village area planted	0.0019	(0.006)	[−0.009, 0.013]	0.58

Notes: This table includes the identical set of regressions in Table 2.5.3, except controlling for a quadratic function of 2001 village population. The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for a 2nd-order polynomial in the running variable and state fixed effects. All specifications also control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.7: RD Sensitivity – Census Outcomes, Second-Order Polynomials



(Figured continued on next page.)



Notes: This two-page figure shows reduced form results estimating Equation (2.1) including a second-order polynomial in population, as reported in Table B.2.13. Blue dots show average residuals form regressing the 2011 outcomes on the 2001 control and state fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Table B.2.14: RD Sensitivity – Census Outcomes, Weighting Inverse Distance from Cutoff

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	-0.5747	(2.536)	$[-5.545, 4.396]$	301.22
0–6 cohort / total population	0.0005	(0.001)	$[-0.001, 0.002]$	0.14
Average household size	-0.0003	(0.014)	$[-0.027, 0.027]$	5.14
Literacy rate	-0.0009	(0.002)	$[-0.005, 0.003]$	0.57
B. Labor outcomes				
Male agricultural workers / male pop	-0.0075^{***}	(0.003)	$[-0.013, -0.002]$	0.43
Female agri. workers / female pop	-0.0063	(0.004)	$[-0.015, 0.002]$	0.28
Male household workers / male pop	-0.0008	(0.001)	$[-0.002, 0.000]$	0.01
Female household workers / female pop	-0.0010	(0.001)	$[-0.003, 0.001]$	0.01
Male other workers / male pop	0.0044^{**}	(0.002)	$[0.000, 0.009]$	0.10
Female other workers / female pop	0.0010	(0.002)	$[-0.003, 0.005]$	0.05
C. Asset ownership				
Share of households with telephone	0.0037	(0.006)	$[-0.008, 0.016]$	0.54
Share of households with TV	0.0050	(0.004)	$[-0.003, 0.013]$	0.25
Share of households with bicycle	0.0007	(0.004)	$[-0.008, 0.009]$	0.51
Share of households with motorcycle	0.0016	(0.003)	$[-0.004, 0.007]$	0.13
Share of households without assets	0.0030	(0.004)	$[-0.005, 0.011]$	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	0.0015	(0.003)	$[-0.004, 0.007]$	0.06
Share of households w/ kerosene lighting	0.0000	(0.007)	$[-0.014, 0.014]$	0.48
Share of households with mud floors	0.0045	(0.004)	$[-0.004, 0.012]$	0.74
Share of households with thatched roof	-0.0042	(0.005)	$[-0.015, 0.006]$	0.22
Share of households dilapidated	-0.0032	(0.003)	$[-0.010, 0.003]$	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	0.0088	(0.013)	$[-0.017, 0.035]$	0.75
1/0 Post office in village	0.0056	(0.004)	$[-0.002, 0.013]$	0.03
1/0 Ag credit societies in village	0.0053	(0.004)	$[-0.002, 0.013]$	0.02
1/0 Water from tubewell in village	0.0040	(0.012)	$[-0.019, 0.028]$	0.46
Share of village area irrigated	-0.0063	(0.006)	$[-0.018, 0.005]$	0.36
Share of village area planted	0.0034	(0.006)	$[-0.008, 0.015]$	0.59

Notes: This table includes the same regressions as in Table 2.5.3, but running weighted least squares. We weight villages by their absolute distance from the 300-person cutoff, such that a $w_v \equiv 1 - \frac{|P_v - 300|}{150}$. The RD bandwidth includes 29,573 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. All specifications also include state fixed effects. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2.4.4 Fixed effects and controls

As mentioned above, while RD designs do not require fixed effects or controls for identification, they can greatly improve precision of RD point estimates. Section B.2.1.4 demonstrates that the RD results for nighttime brightness are robust to different fixed effects specifications, and our results for census outcomes are also not sensitive to cross-sectional fixed effects. Table B.2.15 and Figure B.2.8 present RD results estimating Equation (2.1) without fixed effects, while Table B.2.16 and Figure B.2.9 include more granular fixed effects at the district level. These results are quite similar to our preferred specification (with state fixed effects), with slight differences in precision.

Section B.2.1.4 also shows that our RD results for nighttime brightness depend on the inclusion of 2001 brightness as a village-level control. This reveals that there is enough cross-sectional variation in brightness levels (even within districts) to obscure within-village changes in brightness due to RGGVY electrification. In contrast, Table B.2.17 and Figure B.2.10 show that our RD results for census outcomes are much less sensitive to the exclusion of 2001 controls. The decrease in male agricultural employment at the 300-person cutoff remains small but precisely estimated, and the confidence intervals on other labor outcomes in Panel B are very similar to those in Table 2.5.3.

Comparing Tables 2.5.3 and B.2.17 highlights two key features of our census outcome data. First, Panels C and D report on outcomes from the Houselisting Primary Census Abstract (HPCA), which is not available at the village level for 2001 (see Section B.1.4.2). This means that for Table 2.5.3 regressions in Panels C and D, 2001 controls are at the block level.[56] Because these regressions cannot control for within-block baseline heterogeneity, it is unsurprising that removing these 2001 block-level controls does not affect their precision.

Second, our regressions in Panel E of Table 3 already do not include 2001 controls for 1/0 indicator variables. Since these outcomes are not continuous, including baseline controls would not greatly increase the precision of our RD point estimates. Moreover, including 2001 controls for 1/0 indicator variables would effectively remove observations for villages that did not change status (i.e. from 0 to 1) between 2001 and 2011. This illustrates why we strongly prefer nighttime brightness as a measure of electricity access: nighttime brightness is a more continuous measure of electricity access (a 0–63 scale, compared to a binary indicator); nighttime brightness measures luminosity, a better proxy for electricity *consumption* than binary indicators of electricity *access* (the latter does not account for variation in reliability or usage); the majority of RGGVY villages were targeted for "more intensive electrification", while already meeting the government's official definition of being "electrified" (we would not expect binary indicators to reflect more intensive electrification); and the official definition of "electrified" changed in 2004, meaning that 2001 and 2011 VD variables might not be apples-to-apples comparisons (see Section B.3 for further detail). Table B.2.17 reports RD results for end-use-specific 1/0 electric power indicators, and we see the RGGVY led to statistically significant increases

[56] The only exception is share of households dilapidated, which was only available in 2001 at the district level.

in electricity access for the commercial sector. We do not see corresponding increases for electricity access in the domestic or agricultural sectors, which reflects the fact that most treated villages already had low levels of electrification prior to RGGVY — meaning that the 1/0 indicators were coded as 1 in the baseline, and RGGVY did not cause them to increase.

B.2.4.5 2001 covariate smoothness

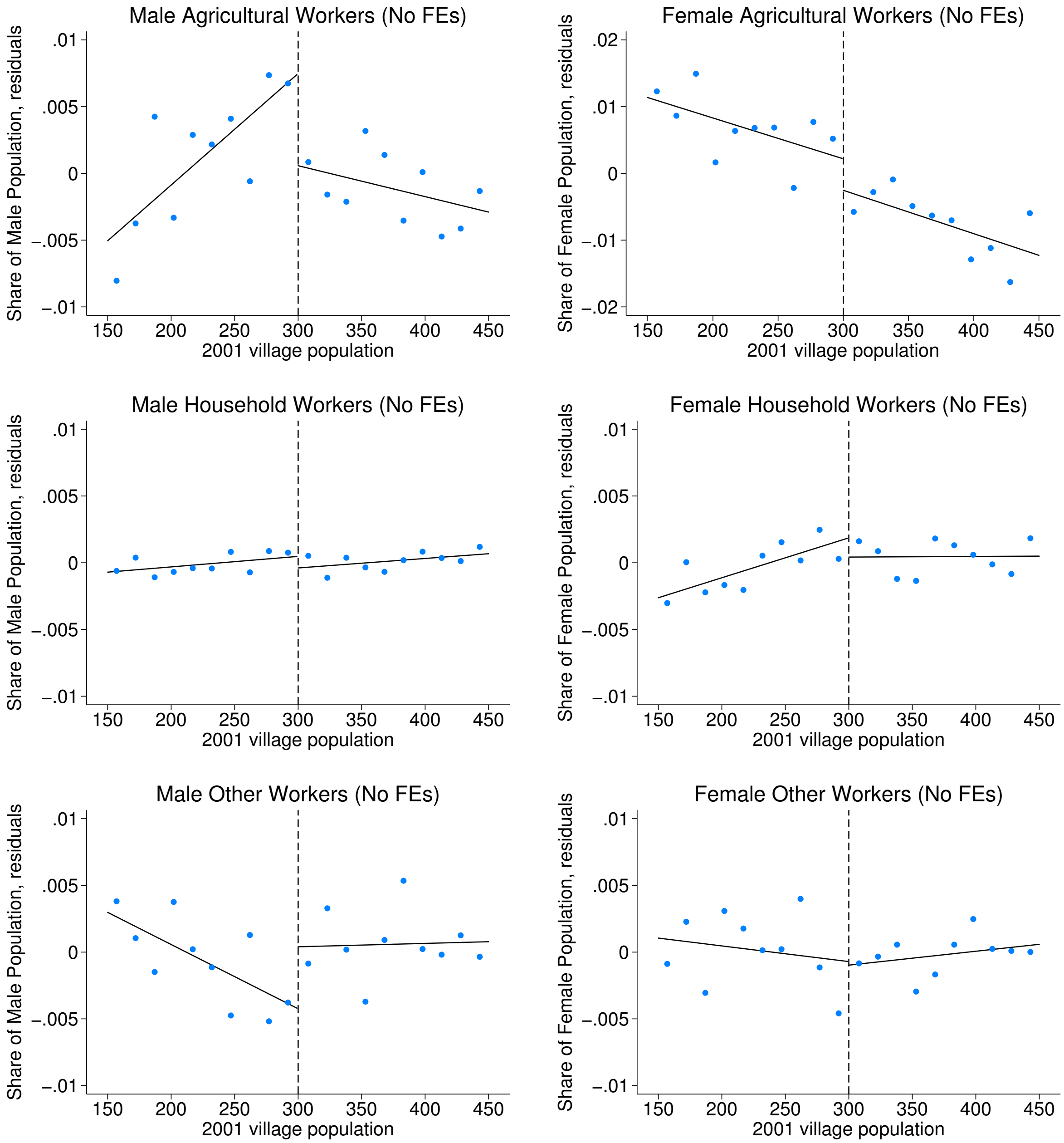
A key RD identifying assumption is smoothness of covariates across the RD cutoff. Section B.2.1.5 demonstrates that nighttime brightness is not discontinuous at the 300-person threshold prior to the 2005 announcement of RGGVY. Table B.2.18 shows that 2001 village-level covariates are also smooth across the 300-person cutoff in 2001 population. This includes all 2001 covariates corresponding to 2011 census outcomes reported in Table 2.5.3, which are available in 2001 at the village level. Figure B.2.11 presents a subset of these results graphically.

Table B.2.15: RD Sensitivity – Census Outcomes, No Fixed Effects

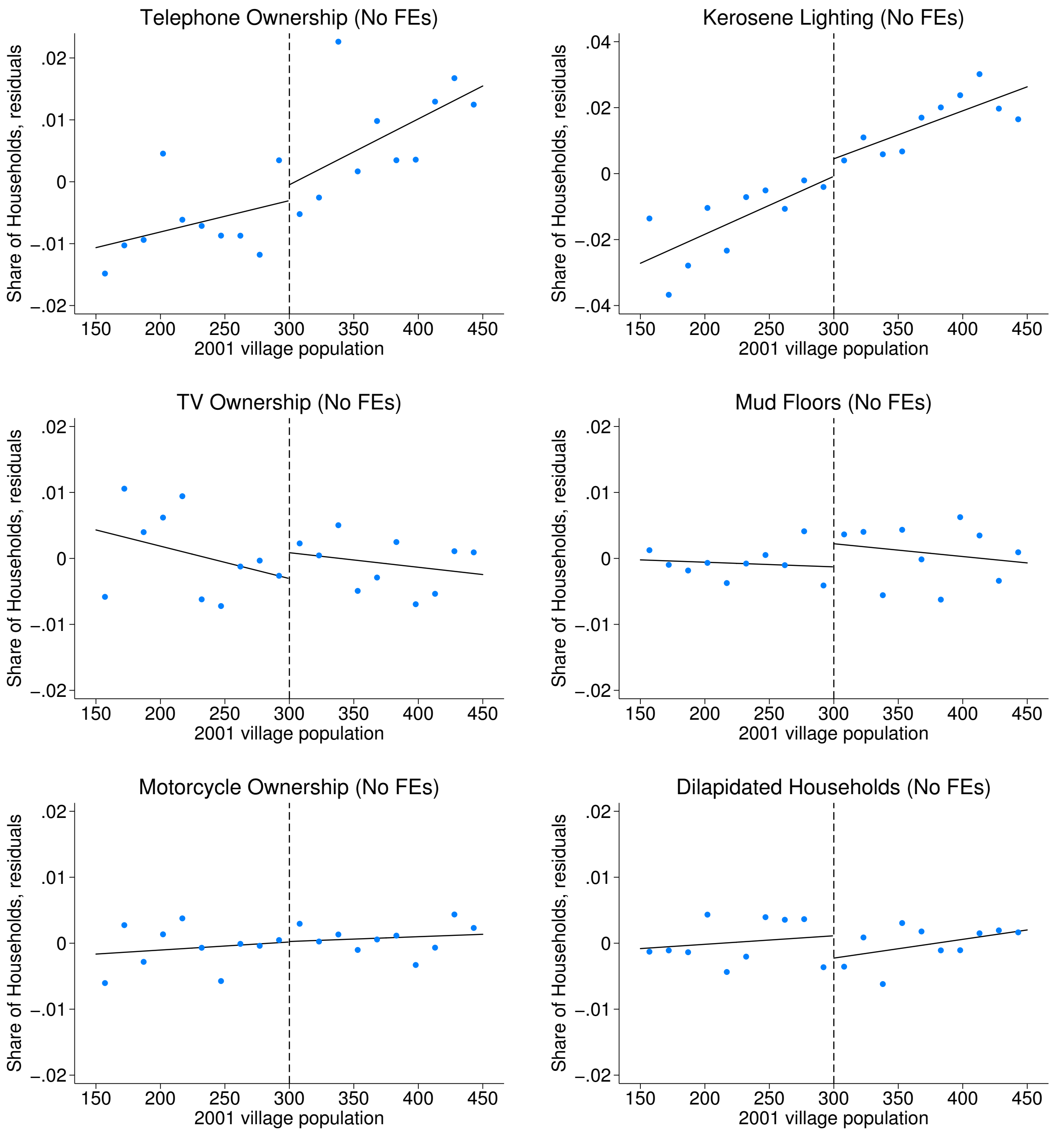
2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	-1.4365	(2.576)	$[-6.486, 3.613]$	271.09
0–6 cohort / total population	0.0011	(0.001)	$[-0.001, 0.003]$	0.14
Average household size	-0.0014	(0.013)	$[-0.027, 0.024]$	5.13
Literacy rate	-0.0025	(0.002)	$[-0.006, 0.002]$	0.57
B. Labor outcomes				
Male agricultural workers / male pop	-0.0069^{**}	(0.003)	$[-0.013, -0.001]$	0.42
Female agri. workers / female pop	-0.0047	(0.004)	$[-0.014, 0.004]$	0.29
Male household workers / male pop	-0.0009	(0.001)	$[-0.002, 0.000]$	0.01
Female household workers / female pop	-0.0014	(0.001)	$[-0.004, 0.001]$	0.01
Male other workers / male pop	0.0046^{**}	(0.002)	$[0.001, 0.008]$	0.10
Female other workers / female pop	-0.0003	(0.002)	$[-0.004, 0.004]$	0.05
C. Asset ownership				
Share of households with telephone	0.0026	(0.006)	$[-0.009, 0.015]$	0.54
Share of households with TV	0.0039	(0.004)	$[-0.004, 0.012]$	0.26
Share of households with bicycle	-0.0007	(0.004)	$[-0.009, 0.008]$	0.50
Share of households with motorcycle	0.0001	(0.003)	$[-0.005, 0.005]$	0.13
Share of households without assets	0.0034	(0.004)	$[-0.005, 0.011]$	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	0.0009	(0.003)	$[-0.004, 0.006]$	0.07
Share of households w/ kerosene lighting	0.0053	(0.006)	$[-0.007, 0.018]$	0.48
Share of households with mud floors	0.0035	(0.004)	$[-0.004, 0.011]$	0.73
Share of households with thatched roof	-0.0032	(0.005)	$[-0.013, 0.007]$	0.23
Share of households dilapidated	-0.0034	(0.003)	$[-0.009, 0.002]$	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	-0.0035	(0.012)	$[-0.027, 0.020]$	0.75
1/0 Post office in village	0.0020	(0.004)	$[-0.005, 0.009]$	0.03
1/0 Ag credit societies in village	0.0010	(0.004)	$[-0.006, 0.008]$	0.02
1/0 Water from tubewell in village	0.0006	(0.011)	$[-0.022, 0.023]$	0.44
Share of village area irrigated	-0.0068	(0.006)	$[-0.019, 0.005]$	0.34
Share of village area planted	0.0059	(0.006)	$[-0.006, 0.018]$	0.58

Notes: This table includes the identical set of regressions in Table 2.5.3, without controlling for state fixed effects. The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.8: RD Sensitivity – Census Outcomes, No Fixed Effects



(Figured continued on next page.)



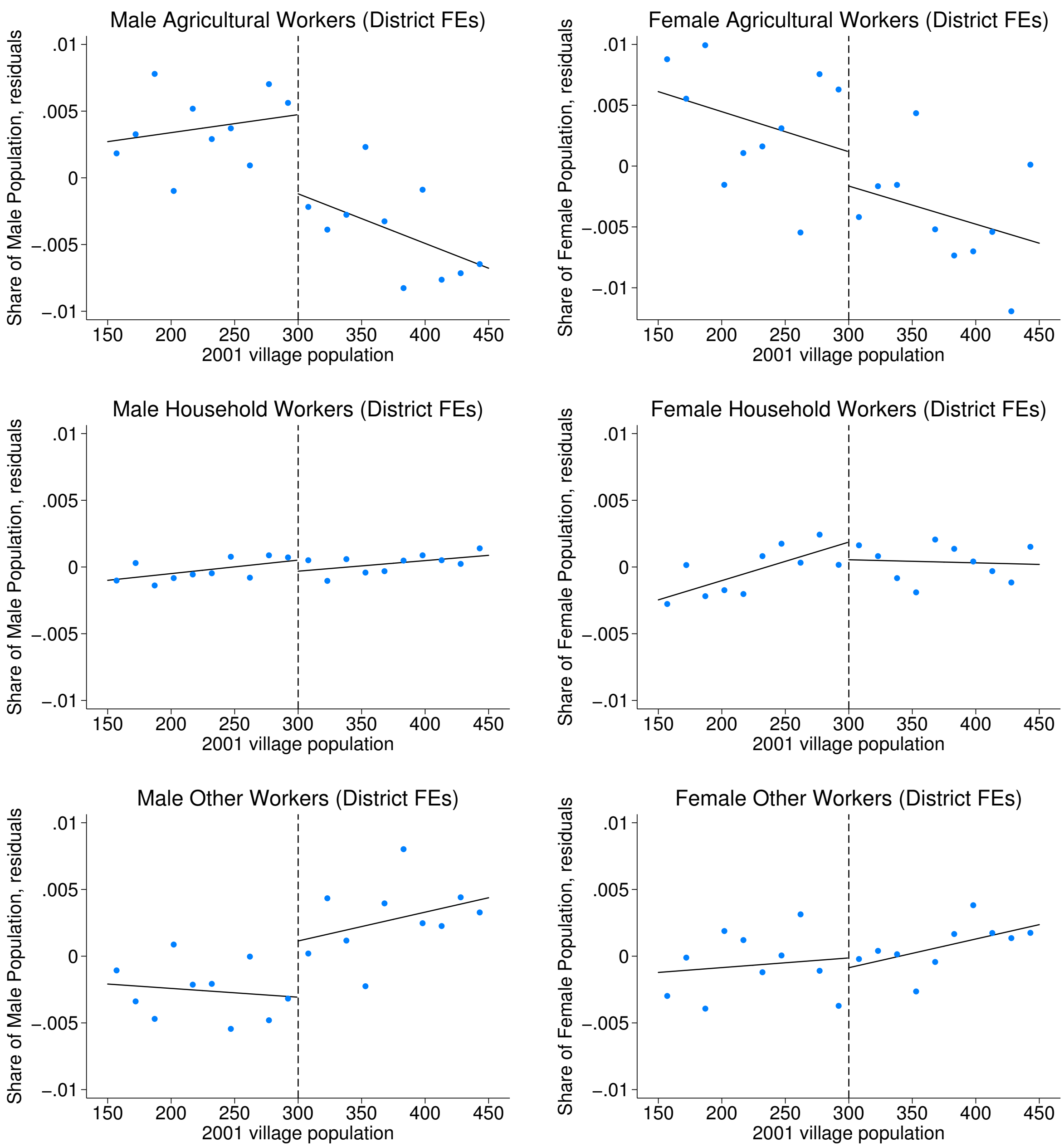
Notes: This two-page figure shows reduced form results estimating Equation (2.1) without fixed effects, as reported in Table B.2.15. Blue dots show average residuals form regressing the 2011 outcomes on the 2001 control. Each dot contains approximately 1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Table B.2.16: RD Sensitivity – Census Outcomes, District Fixed Effects

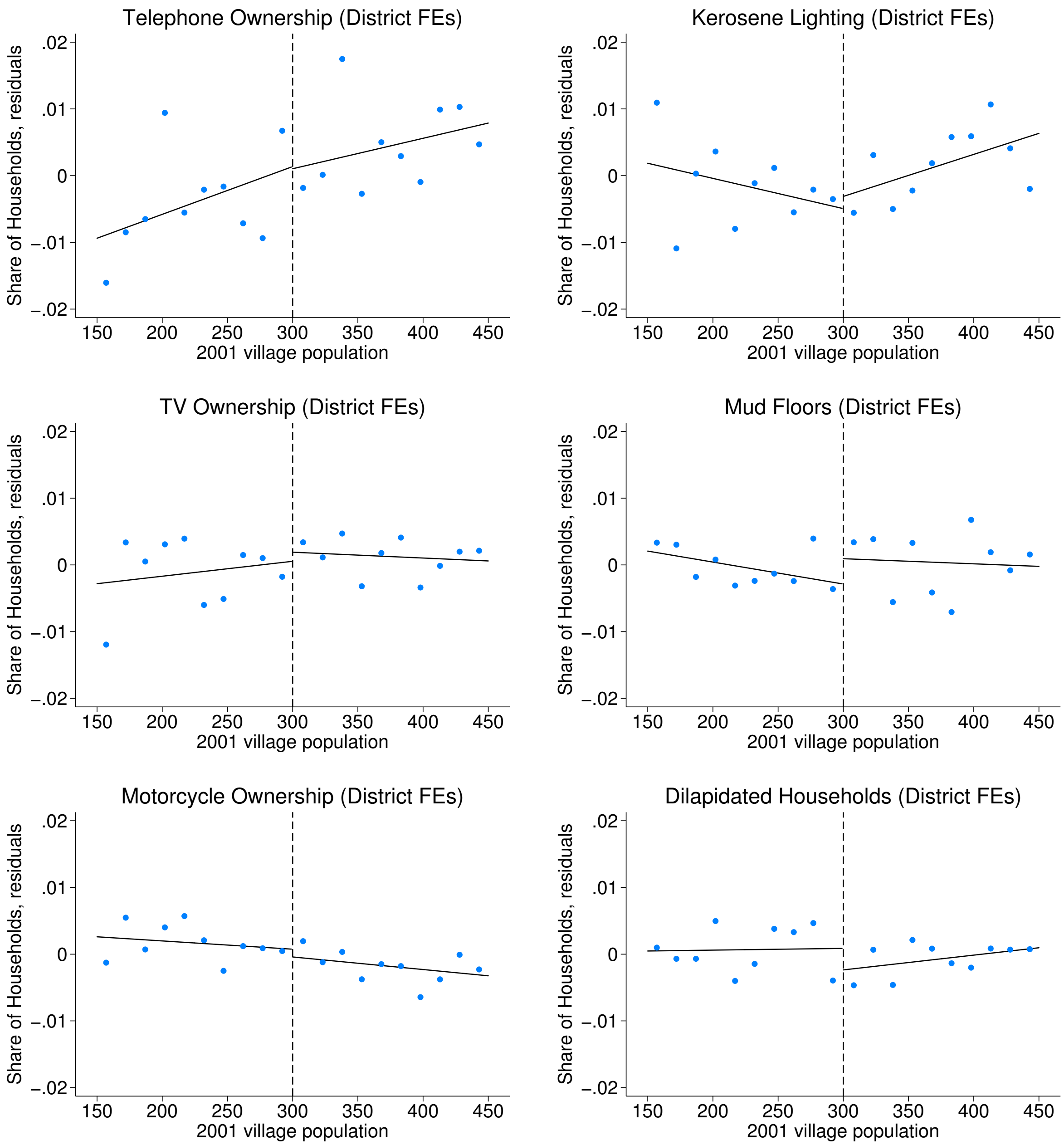
2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	-0.6014	(2.371)	$[-5.248, 4.045]$	271.09
0–6 cohort / total population	0.0010	(0.001)	$[-0.001, 0.003]$	0.14
Average household size	-0.0060	(0.012)	$[-0.030, 0.018]$	5.13
Literacy rate	-0.0031	(0.002)	$[-0.007, 0.001]$	0.57
B. Labor outcomes				
Male agricultural workers / male pop	-0.0060^{**}	(0.003)	$[-0.011, -0.001]$	0.42
Female agri. workers / female pop	-0.0029	(0.004)	$[-0.011, 0.005]$	0.29
Male household workers / male pop	-0.0008	(0.001)	$[-0.002, 0.000]$	0.01
Female household workers / female pop	-0.0013	(0.001)	$[-0.003, 0.001]$	0.01
Male other workers / male pop	0.0043^{**}	(0.002)	$[0.001, 0.008]$	0.10
Female other workers / female pop	-0.0007	(0.002)	$[-0.004, 0.003]$	0.05
C. Asset ownership				
Share of households with telephone	-0.0003	(0.005)	$[-0.011, 0.010]$	0.54
Share of households with TV	0.0014	(0.004)	$[-0.006, 0.009]$	0.26
Share of households with bicycle	-0.0016	(0.004)	$[-0.010, 0.006]$	0.50
Share of households with motorcycle	-0.0012	(0.002)	$[-0.006, 0.004]$	0.13
Share of households without assets	0.0041	(0.004)	$[-0.004, 0.012]$	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	-0.0005	(0.003)	$[-0.006, 0.005]$	0.07
Share of households w/ kerosene lighting	0.0018	(0.006)	$[-0.010, 0.014]$	0.48
Share of households with mud floors	0.0038	(0.004)	$[-0.003, 0.011]$	0.73
Share of households with thatched roof	-0.0018	(0.005)	$[-0.011, 0.008]$	0.23
Share of households dilapidated	-0.0033	(0.003)	$[-0.009, 0.002]$	0.07
E. Village-wide outcomes				
1/0 Mobile phone coverage in village	0.0022	(0.011)	$[-0.019, 0.024]$	0.75
1/0 Post office in village	0.0010	(0.004)	$[-0.006, 0.008]$	0.03
1/0 Ag credit societies in village	0.0013	(0.004)	$[-0.006, 0.008]$	0.02
1/0 Water from tubewell in village	-0.0013	(0.010)	$[-0.021, 0.018]$	0.44
Share of village area irrigated	-0.0059	(0.005)	$[-0.015, 0.004]$	0.34
Share of village area planted	0.0016	(0.005)	$[-0.007, 0.011]$	0.58

Notes: This table includes the identical set of regressions in Table 2.5.3, except controlling for district fixed effects (instead of state fixed effects). The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2001 level of the outcome variable, except for share of village area planted (where 2001 values are not available) and 1/0 indicator variables. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.9: RD Sensitivity – Census Outcomes, District Fixed Effects



(Figured continued on next page.)



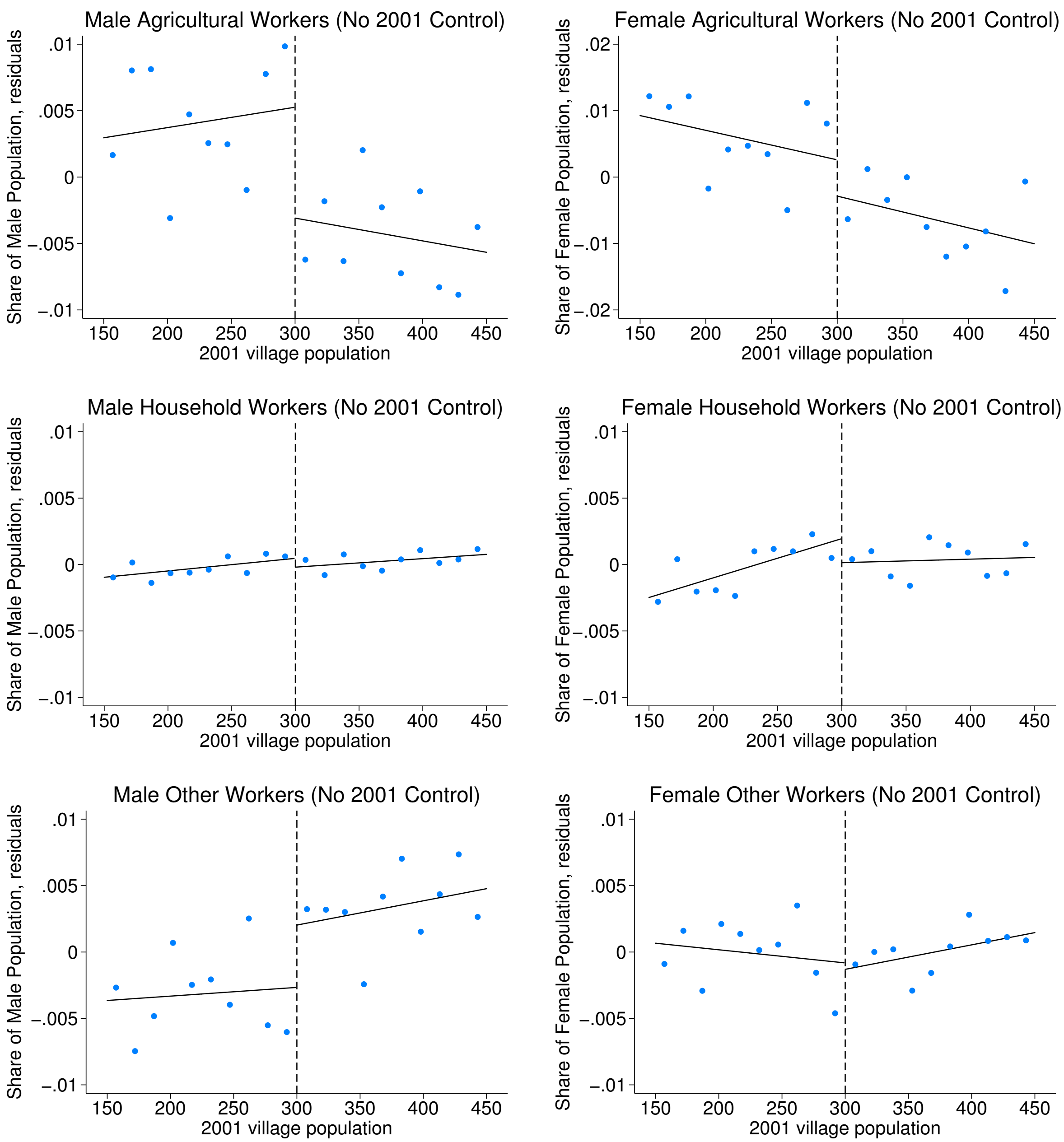
Notes: This two-page figure shows reduced form results estimating Equation (2.1) including district fixed effects, as reported in Table B.2.16. Blue dots show average residuals form regressing the 2011 outcomes on the 2001 control and district fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Table B.2.17: RD Sensitivity – Census Outcomes, No 2001 Controls

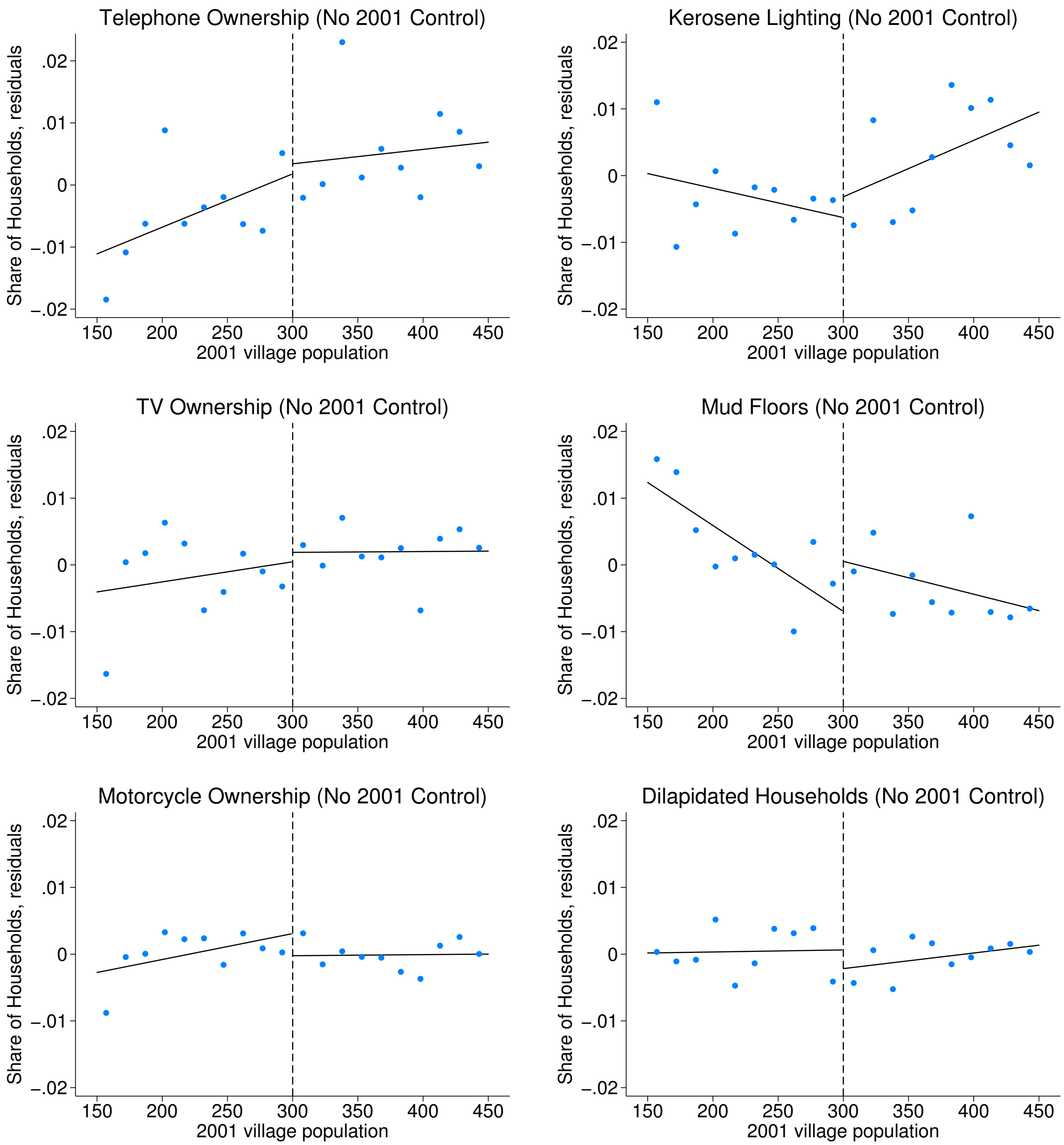
2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic outcomes				
Total population	−0.8647	(2.528)	[−5.820, 4.091]	271.09
0–6 cohort / total population	0.0007	(0.001)	[−0.001, 0.002]	0.14
Average household size	0.0022	(0.014)	[−0.026, 0.031]	5.13
Literacy rate	−0.0016	(0.003)	[−0.007, 0.004]	0.57
B. Labor outcomes				
Male agricultural workers / male pop	−0.0084***	(0.003)	[−0.014, −0.003]	0.42
Female agri. workers / female pop	−0.0055	(0.004)	[−0.013, 0.003]	0.29
Male household workers / male pop	−0.0007	(0.001)	[−0.002, 0.001]	0.01
Female household workers / female pop	−0.0018	(0.001)	[−0.004, 0.000]	0.01
Male other workers / male pop	0.0047**	(0.002)	[0.000, 0.009]	0.10
Female other workers / female pop	−0.0005	(0.002)	[−0.004, 0.003]	0.05
C. Asset ownership				
Share of households with telephone	0.0016	(0.006)	[−0.009, 0.013]	0.54
Share of households with TV	0.0014	(0.004)	[−0.007, 0.009]	0.26
Share of households with bicycle	−0.0034	(0.006)	[−0.015, 0.008]	0.50
Share of households with motorcycle	−0.0033	(0.003)	[−0.010, 0.003]	0.13
Share of households without assets	0.0032	(0.005)	[−0.006, 0.012]	0.22
D. Housing stock				
Share of households w/ elec/gas cooking	−0.0006	(0.003)	[−0.006, 0.005]	0.07
Share of households w/ kerosene lighting	0.0031	(0.006)	[−0.009, 0.015]	0.48
Share of households with mud floors	0.0075	(0.005)	[−0.003, 0.018]	0.73
Share of households with thatched roof	−0.0022	(0.006)	[−0.013, 0.009]	0.23
Share of households dilapidated	−0.0028	(0.003)	[−0.009, 0.003]	0.07
E. Village-wide outcomes				
1/0 Electricity (domestic use)	−0.0086	(0.006)	[−0.021, 0.004]	0.90
1/0 Electricity (agricultural use)	−0.0098	(0.010)	[−0.030, 0.010]	0.61
1/0 Electricity (commercial use)	0.0243**	(0.011)	[0.003, 0.046]	0.44
1/0 Electricity (all end uses)	0.0240**	(0.011)	[0.002, 0.046]	0.43
Share of village area irrigated	−0.0021	(0.006)	[−0.013, 0.009]	0.35

Notes: Panels A–D of this table included the same sets of regressions in Table 2.5.3. Panel E includes sector-specific 1/0 indicator variables for electricity availability at the village level, while omiting 4 regressions from Table 2.5.3 that already did not include 2001 controls. All RD regressions in this table control only for state fixed effects, not for the 2001 level of the outcome variable. The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.10: RD Sensitivity – Census Outcomes, No 2001 Controls



(Figured continued on next page.)



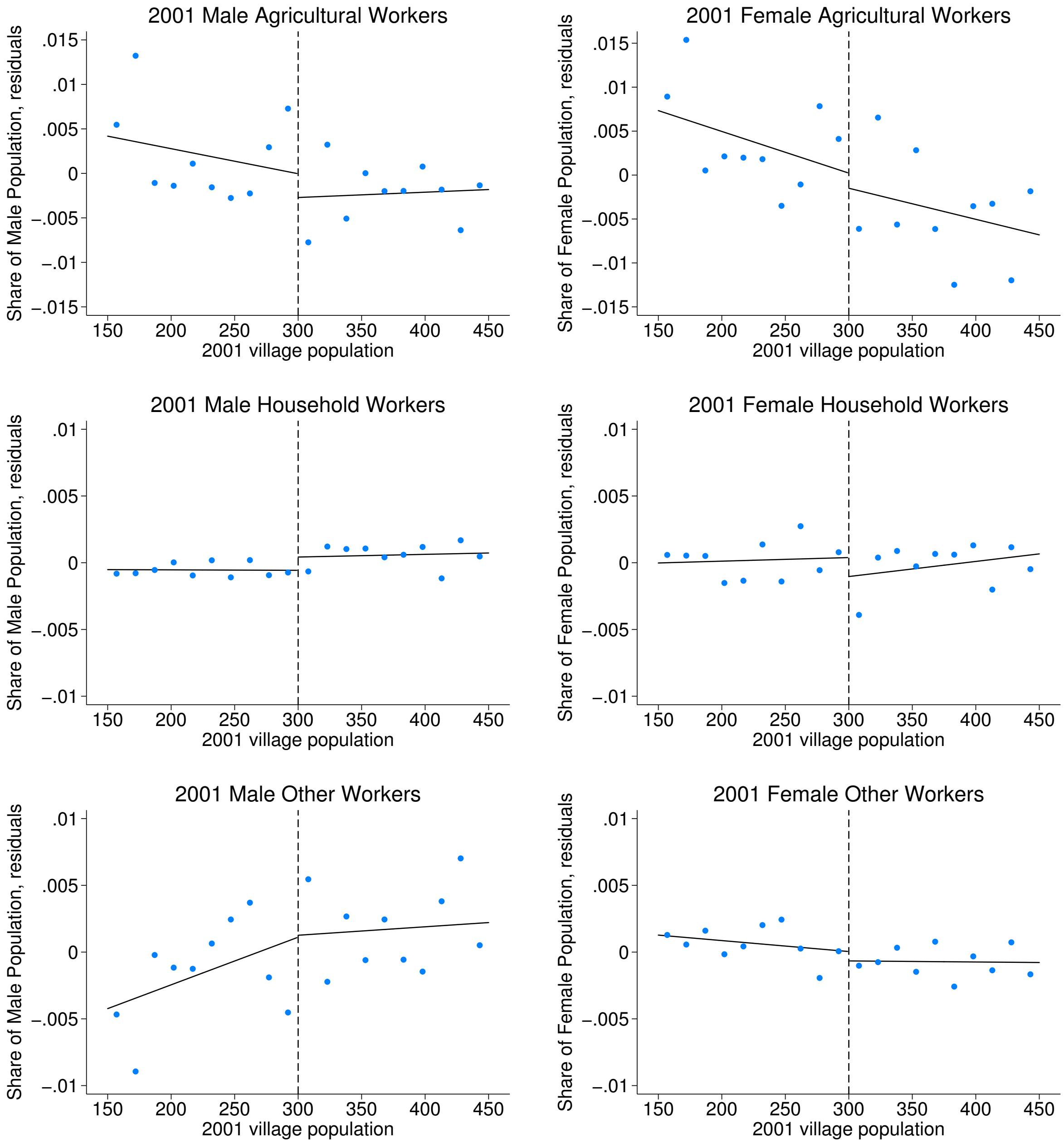
Notes: This two-page figure shows reduced form results estimating Equation (2.1) without 2001 controls, as reported in Table B.2.17. Blue dots show average residuals form regressing the 2011 outcomes on state fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

Table B.2.18: RD Sensitivity – Census Outcomes, 2001 Covariate Smoothness

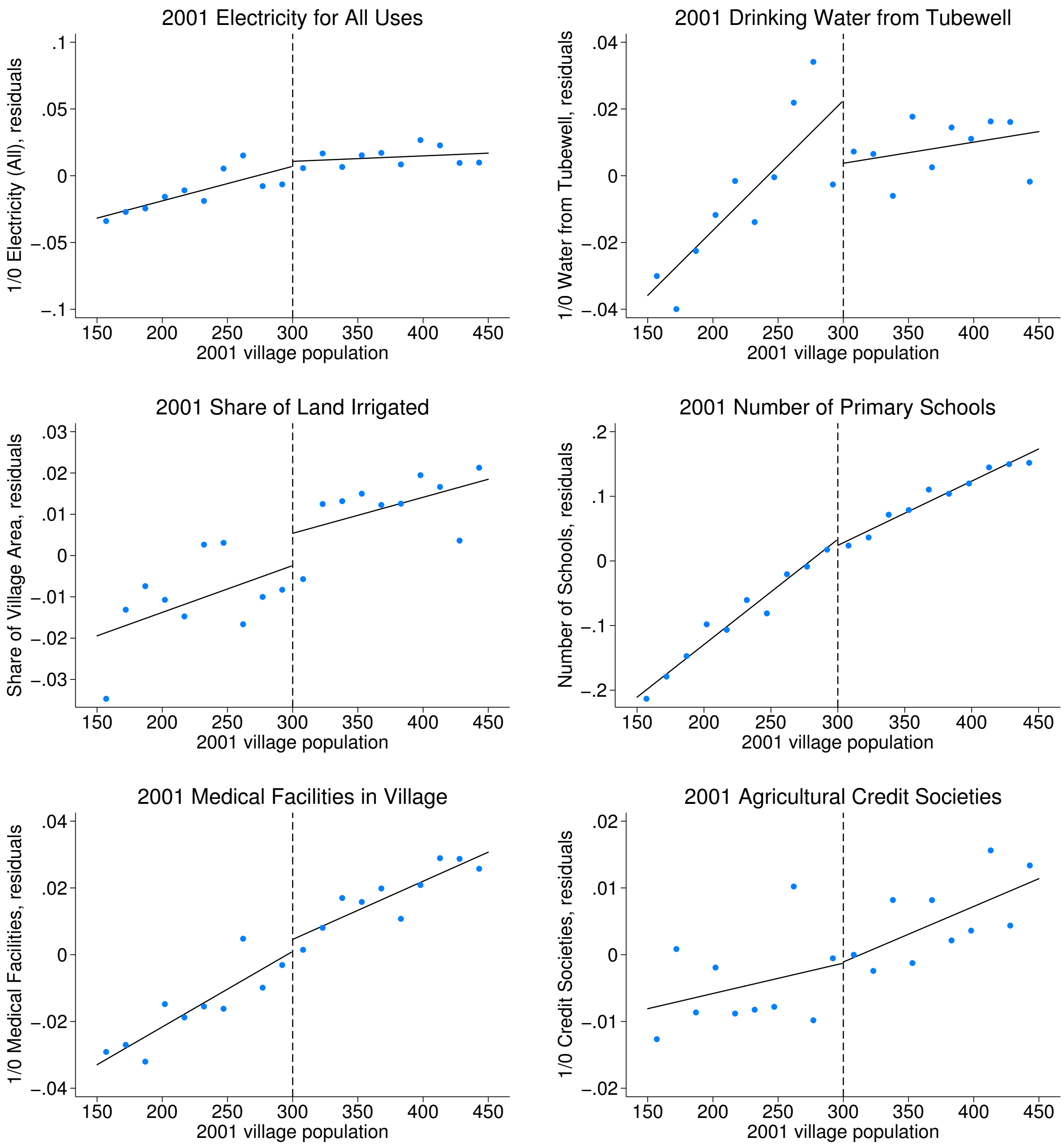
2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Demographic indicators				
0–6 cohort / total population	−0.0009	(0.001)	[−0.003, 0.001]	0.18
Share or population SC or ST	−0.0080	(0.007)	[−0.021, 0.005]	0.33
Literacy rate	0.0015	(0.003)	[−0.005, 0.008]	0.45
B. Labor indicators				
Male agricultural workers / male pop	−0.0027	(0.003)	[−0.008, 0.003]	0.43
Female agri. workers / female pop	−0.0017	(0.005)	[−0.011, 0.007]	0.31
Male household workers / male pop	0.0010	(0.001)	[−0.000, 0.002]	0.01
Female household workers / female pop	−0.0014	(0.001)	[−0.004, 0.001]	0.01
Male other workers / male pop	0.0001	(0.002)	[−0.004, 0.004]	0.09
Female other workers / female pop	−0.0007	(0.002)	[−0.004, 0.003]	0.03
Male main workers / male pop	−0.0017	(0.004)	[−0.009, 0.005]	0.43
Female main workers / female pop	−0.0047	(0.004)	[−0.013, 0.004]	0.17
Male marginal workers / male pop	0.0002	(0.003)	[−0.005, 0.006]	0.09
Female marginal workers / female pop	0.0008	(0.004)	[−0.007, 0.009]	0.18
C. Village-wide indicators				
1/0 Electricity (all end uses)	0.0038	(0.007)	[−0.009, 0.017]	0.29
1/0 Electricity (domestic use)	0.0018	(0.010)	[−0.018, 0.021]	0.59
1/0 Electricity (agricultural use)	−0.0014	(0.009)	[−0.020, 0.017]	0.41
1/0 Water from tubewell in village	−0.0188*	(0.010)	[−0.039, 0.001]	0.51
Share of village area irrigated	0.0079	(0.007)	[−0.006, 0.022]	0.33
1/0 Educational facilities	−0.0081	(0.010)	[−0.028, 0.012]	0.50
Number of primary schools	−0.0093	(0.012)	[−0.032, 0.014]	0.49
Number of secondary schools	0.0039	(0.003)	[−0.001, 0.009]	0.01
1/0 Medical facilities	0.0036	(0.007)	[−0.010, 0.018]	0.10
1/0 Banking facilities	0.0012	(0.002)	[−0.003, 0.006]	0.01
1/0 Agricultural credit societies	0.0002	(0.004)	[−0.008, 0.009]	0.02
1/0 Post office	0.0006	(0.004)	[−0.007, 0.008]	0.03
1/0 Bus service	0.0058	(0.014)	[−0.022, 0.033]	0.41

Notes: This table regresses 2001 village covariates on the RD variables in Equation 2.1 and state fixed effects. Each row represents a separate regression, with the second column reporting the RD point estimate ($\hat{\beta}_1$). This table includes all available 2001 village-level covariates that correspond to a 2011 outcome reported in Table 2.5.3. SC and ST refer to official scheduled caste and scheduled tribe designations. "Main" workers work at least six months of the year, while "marginal" workers work less than six months. The RD bandwidth includes 29,765 villages with 2001 populations between 150 and 450, across 225 districts. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.11: RD Sensitivity – Census Outcomes, 2001 Covariate Smoothness



(Figured continued on next page.)



Notes: This two-page figure shows reduced form results estimating Equation (2.1) without 2001 controls, as reported in Table B.2.17. Blue dots show average residuals form regressing the 2011 outcomes on state fixed effects. Each dot contains approximately 1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts.

B.2.5 Census Outcome Results: Intensive Margin of Labor

Our main results examine the effects of electrification on extensive-margin employment. Thus far, we have presented results on the fraction of men and women working in agriculture, in the home, and in other sectors (see Table 2.5.3). We find evidence that eligibility for RGGVY shifts a small number of men out of agriculture and into the more formal sector, and we find a small but not statistically significant impact of RGGVY eligibility on female employment. However, it is also possible that electrification changes employment on the intensive margin, by causing previously employed workers to either increase of decrease the amount they work.

We test for effects on this intensive margin using data from the PCA, which reports separately employment counts for workers are "main", working 6 or more months out of the year, or "marginal", working fewer than 6 months.[57] We conduct an analogous employment analysis using the number of main workers divided by the number of total (main plus marginal) workers for each gender and sector as the dependent variable. In doing this, we drop villages with no workers in a particular gender-sector group, where this fraction is undefined. (This essentially amounts to dropping cases where this intensive margin does not exist.) Our results should therefore be interpreted as the effects of eligibility for electrification on the share of workers working at least 6 months per year, conditional on each village employing people of a given category.

Table B.2.19 and Figure B.2.12 display the results of this exercise. We find no evidence of a discontinuity in the intensive margin of labor at the 300-person RD threshold across any gender-sector category. Our point estimates across all categories are negative, with the exception of female agricultural workers. This demonstrates that RGGVY eligibility did not increase labor on the intensive margin. None of our results are statistically significant, but our confidence intervals are relatively tight, and we can reject small effects.

B.2.6 Census Outcome Results: Validity Tests

Section 2.5 in the main text supports the validity of the nighttime brightness RD results using a placebo test, a randomization test, and three falsification tests. Below, we conduct the analogous validity tests on the two census outcomes with the strongest non-zero results — the share of male agricultural workers, and the share of male "other" workers. While even these RD results are small in magnitude, they are statistically significant and have RD plots that display a level shift at the 300-person eligibility cutoff.[58] This implies that these validity tests can support our use of RD inference to test the effects of RGGVY program eligibility on these labor outcomes.

[57]For all employment results in Table 2.5.3 and Appendix B.2.4, we sum main and marginal workers into a single pooled employment metric for each subcategory.

[58]Male agricultural labor is statistically significant at the 5 percent level in every set of sensitivity results we present. Male other labor is mostly significant at the 5 percent level, and quite robust to different RD specifications.

Figure B.2.13 reports the results of placebo and randomization tests for 2011 male labor outcomes. We implement these tests separately on each outcome, exactly as described in Section 2.5 in the main text. The decrease in the share of men working in agriculture falls below the 3rd percentile of the placebo distribution and below the 1st percentile of the randomization distribution. This indicates that that this RD result is very unlikely to reflect spurious volatility in the relationship between male agricultural labor and village population data. The increase in the share of male "other" labor also passes the placebo and randomization tests (above the 99th and 97th percentiles, respectively).

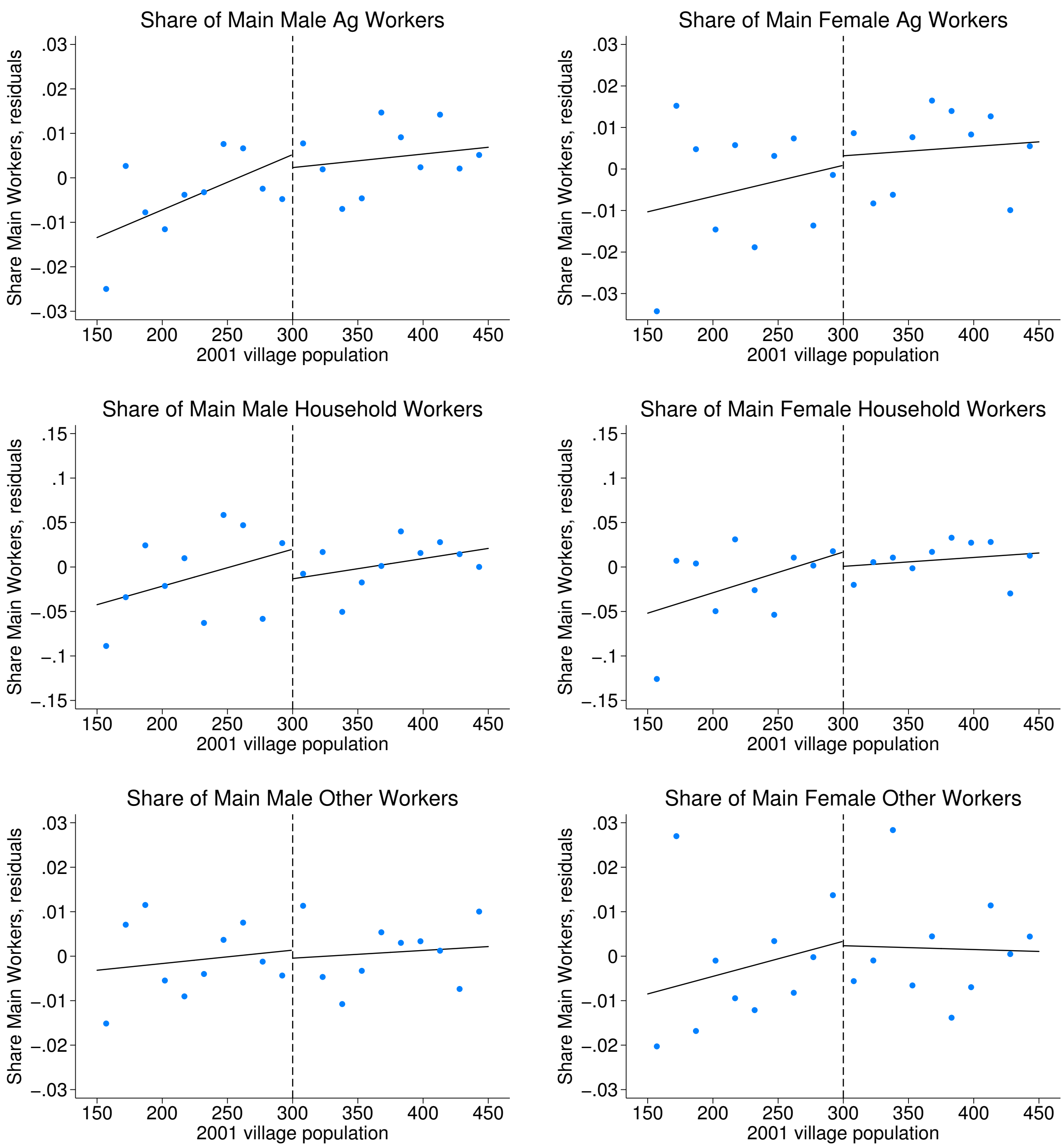
Table B.2.20, Figure B.2.14, and Figure B.2.15 conduct three falsification tests, by estimating Equation (2.1) on subsets of villages for which the 300-person cutoff in village population was not the relevant criterion determining RGGVY eligibility (because these villages contain multiple habitations and/or faced a 100-person eligibility cutoff). As expected, none of the three alternative samples shows statistically significant discontinuities at the 300-person cutoff. This provides further evidence that the RGGVY program has driven the small-but-statistically-significant shift of male labor out of the agricultural sector.

Table B.2.19: RD Results – Share of "Main" Workers by Sector

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome	Total Villages
Share main male ag workers	−0.0029	(0.007)	[−0.016, 0.011]	0.72	29,572
Share main female ag workers	0.0023	(0.010)	[−0.017, 0.022]	0.51	26,487
Share main male hh workers	−0.0333	(0.023)	[−0.079, 0.013]	0.65	5,324
Share main female hh workers	−0.0161	(0.035)	[−0.084, 0.052]	0.42	2,807
Share main male other workers	−0.0018	(0.008)	[−0.017, 0.013]	0.79	24,855
Share main female other workers	−0.0010	(0.013)	[−0.027, 0.025]	0.69	14,389

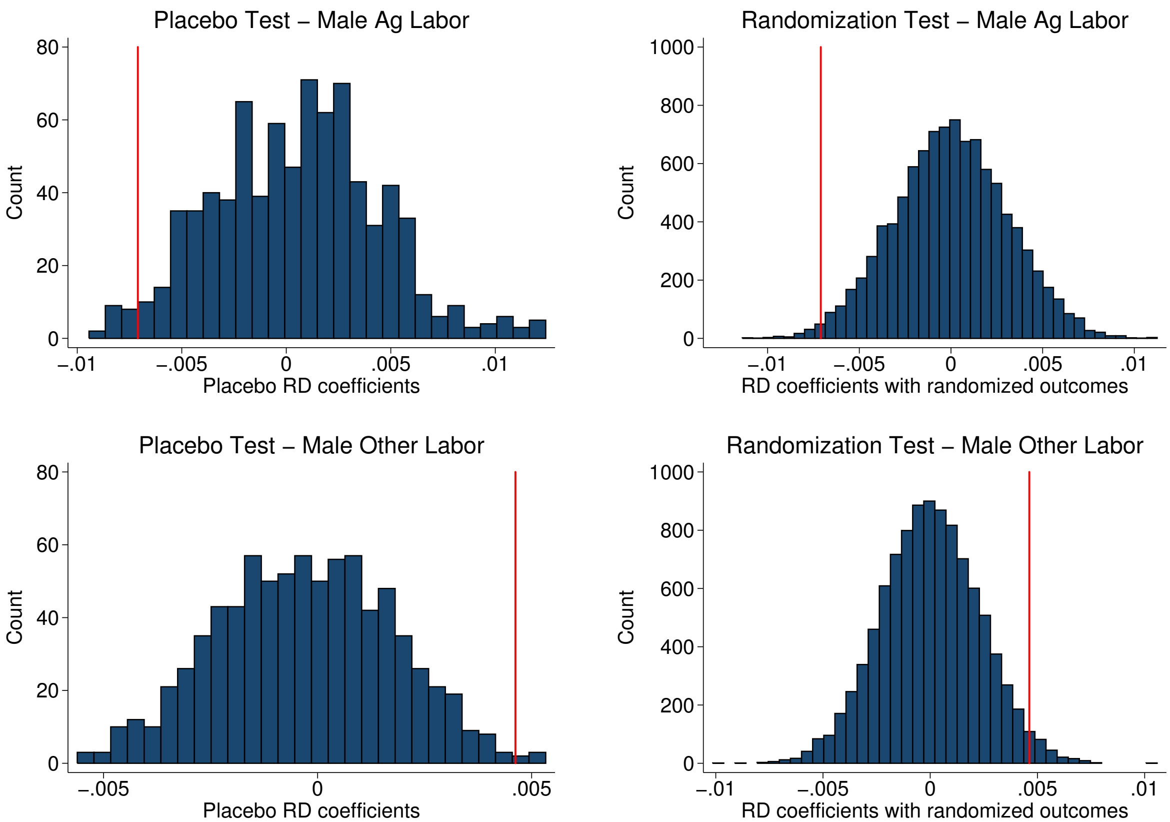
Notes: This table estimates Equation (2.1) on the share of "main" workers for each sector/gender. By definition, "main" workers work for at least 6 months of the year, while "marginal" workers work for less than 6 months of the year. Our main results in Table 2.5.3 pooled main and marginal workers, thereby testing for effects on the extensive margin of labor. In this table, we test the intensive margin using outcome variables that divide the number of main workers by the number of (main + marginal) workers in each category. Each row represents a separate regression, with the second column reporting the RD point estimate ($\hat{\beta}_1$). All specifications control for state fixed effects and the 2001 level of the outcome variable. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. The sixth column shows the number of villages in each regression, which varies since the 2011 outcome and the 2001 control divide by total workers and many villages report zero total workers for a given category. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.12: RD Results – Share of "Main" Workers by Sector



Notes: This figure shows RD results estimating Equation (2.1) on the share of workers in each category that work at least 6 months of the year, as reported in Table B.2.19. Blue dots show average residuals form regressing the 2011 outcomes on the 2001 control and state fixed effects. Each dot contains between 150–1,500 villages, averaged in 15-person population bins, including all 29,765 single-habitation villages between 150 and 450 people, in 10th-Plan districts, with nonzero total workers for both 2001 and 2011.

Figure B.2.13: Male Labor Shares – Placebo and Randomization Tests



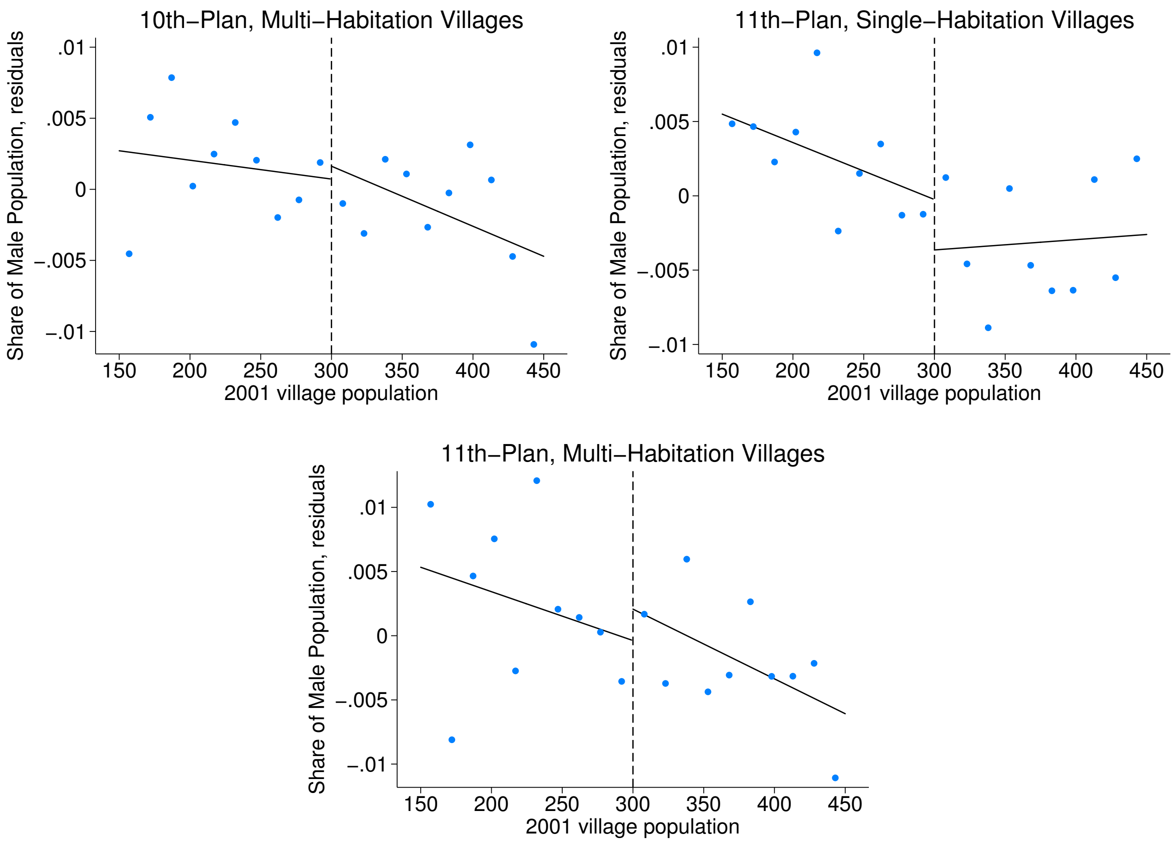
Notes: This figure presents the distributions of placebo RD coefficients and randomized RD coefficients. The left panels were generated by estimating Equation (2.1) using 801 placebo RD thresholds, representing all integer values in $[151, 275] \cup [325, 1000]$. We omit placebo thresholds within 25 people of the true 300-person threshold to ensure that placebo RDs do not detect the true effects of RGGVY eligibility, and we exclude thresholds below 151 due to our 150-person bandwidth. The right panel was generated by scrambling village brightness 10,000 times and re-estimating Equation (2.1) each time. The red lines represent our estimates of the RD coefficients from Table 2.5.3, using the correct 300-person threshold with unscrambled lights data. The RD point estimate for the share of male agricultural workers falls below the 3nd percentile of the placebo distribution and below the 1st percentile of the randomization distribution. The RD point estimate for the share of male other workers falls above the 99th percentile of the placebo distribution and above the 97th percentile of the randomization distribution.

Table B.2.20: RD Sensitivity – Falsification Tests

2011 Outcome Variable	10th Plan Single-Hab.	10th Plan Multi-Hab.	11th Plan Single-Hab.	11th Plan Multi-Hab.
	(1)	(2)	(3)	(4)
Male ag workers / male pop	-0.0071^{**}	0.0009	-0.0034	0.0024
	(0.0028)	(0.0039)	(0.0033)	(0.0042)
Male oth workers / male pop	0.0046^{**}	0.0032	0.0027	-0.0003
	(0.0019)	(0.0033)	(0.0027)	(0.0033)
RD bandwidth	150	150	150	150
Number of observations	29,765	16,481	24,104	16,164
Number of districts	225	202	261	219

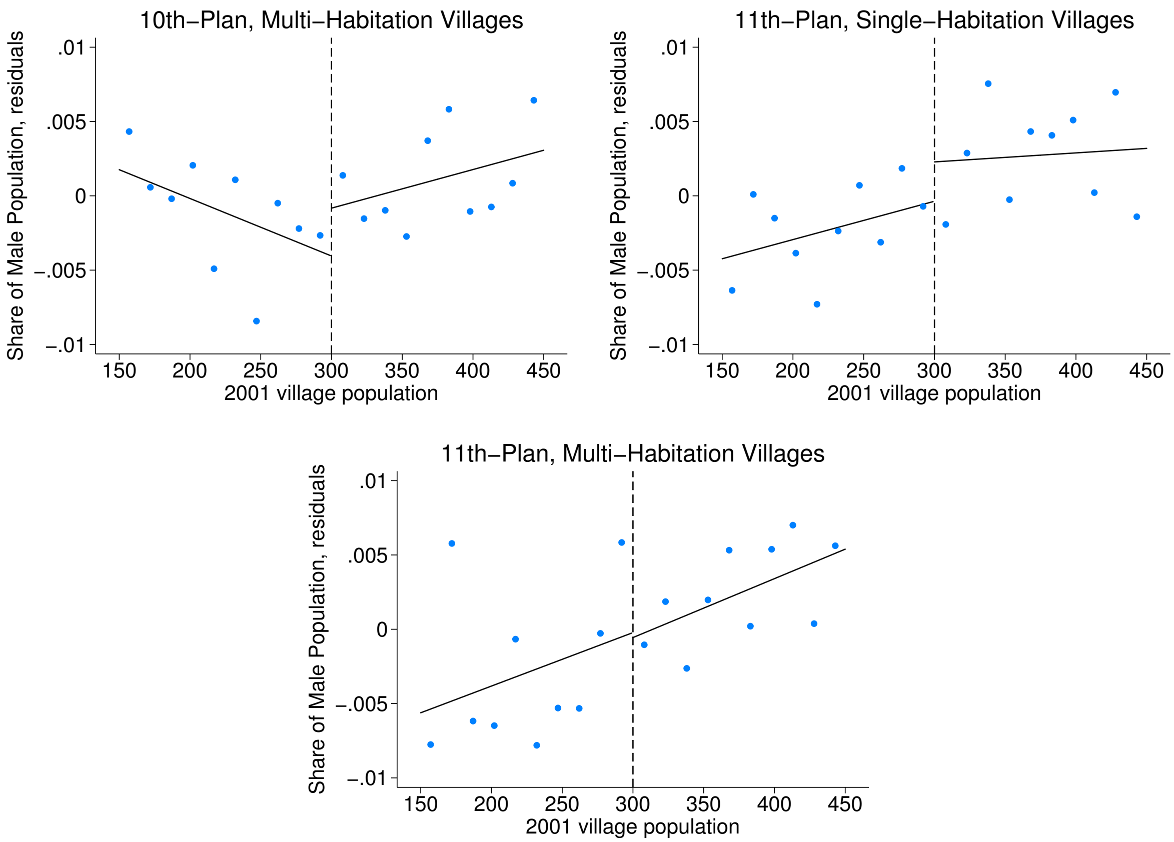
Notes: This table reports results from 8 separate RD regressions, estimating Equation (2.1) on four disjoint subsets of Indian villages. Column (1) reproduces results from Table 2.5.3, using our RD sample of single-habitation villages in 10th-Plan RGGVY districts. Columns (2) and (4) include villages with multiple habitations, for which the the running variable (village population) does not correspond to the habitation populations that determined village eligibility. Columns (3) and (4) includes villages that were eligibile for RGGVY under the 11th Plan, after the cutoff had moved from 300 to 100 people. Figures B.2.14 and B.2.15 present these falsification tests graphically. All specifications control for the 2001 level of the outcome variable and state fixed effects. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.14: Male Agricultural Labor – Falsification Tests



Notes: This figure presents three falsification tests for our RD on male agricultural labor, corresponding to the first row of Table B.2.20. The top-left and bottom panels include only villages with multiple habitations, for which the running variable of village population did not determine RGGVY eligibility. The top-rignt and bottom panels include only villages in districts that became eligible for RGGVY under the 11th Plan, for which the appropriate eligibility cutoff was lowered from 300 to 100 people. Blue dots show average residuals form regressing the 2011 outcomes on 2001 male agricultural employment and state fixed effects. Each dot contains approximately 800–1,500 villages, averaged in 15-person population bins. Lines are estimated separately on each side of the 300-person threshold, for villages within the 150–450 population RD bandwidth.

Figure B.2.15: Male Other Labor – Falsification Tests



Notes: This figure presents three falsification tests for our RD on male other labor, corresponding to the second row of Table B.2.20. The top-left and bottom panels include only villages with multiple habitations, for which the running variable of village population did not determine RGGVY eligibility. The top-right and bottom panels include only villages in districts that became eligible for RGGVY under the 11th Plan, for which the appropriate eligibility cutoff was lowered from 300 to 100 people. Blue dots show average residuals form regressing the 2011 outcomes on 2001 male other employment and state fixed effects. Each dot contains approximately 800–1,500 villages, averaged in 15-person population bins. Lines are estimated separately on each side of the 300-person threshold, for villages within the 150–450 population RD bandwidth.

B.2.7 Socio-Economic Caste Census Results: RD Robustness

Table 2.5.4 presents in the main text results from ten SECC outcomes relating to household wealth and adult employment. Figure 2.5.9 displays RD plots for four of these ten outcomes, and we report the remaining six RD plots below in Figure B.2.16. The two household figures show no visual evidence of discontinuities at the 300-person threshold, confirming our results from Table 2.5.4 that reject economically significant effect as a result of RGGVY electrification. The four employment figures are qualitatively similar to those in Figure 2.5.7, except for suggestive evidence of an increase in the share of adult women working on other jobs.

Figure B.2.17 conducts RD bandwidth sensitivities on two SECC outcomes reported in the main text: the share of total households with at least one poverty indicator, and the share of these households earning a monthly income greater than Rs 5,000. We see that for bandwidths above 80 people, our RD estimates are not sensitive to the choice of bandwidth.[59]

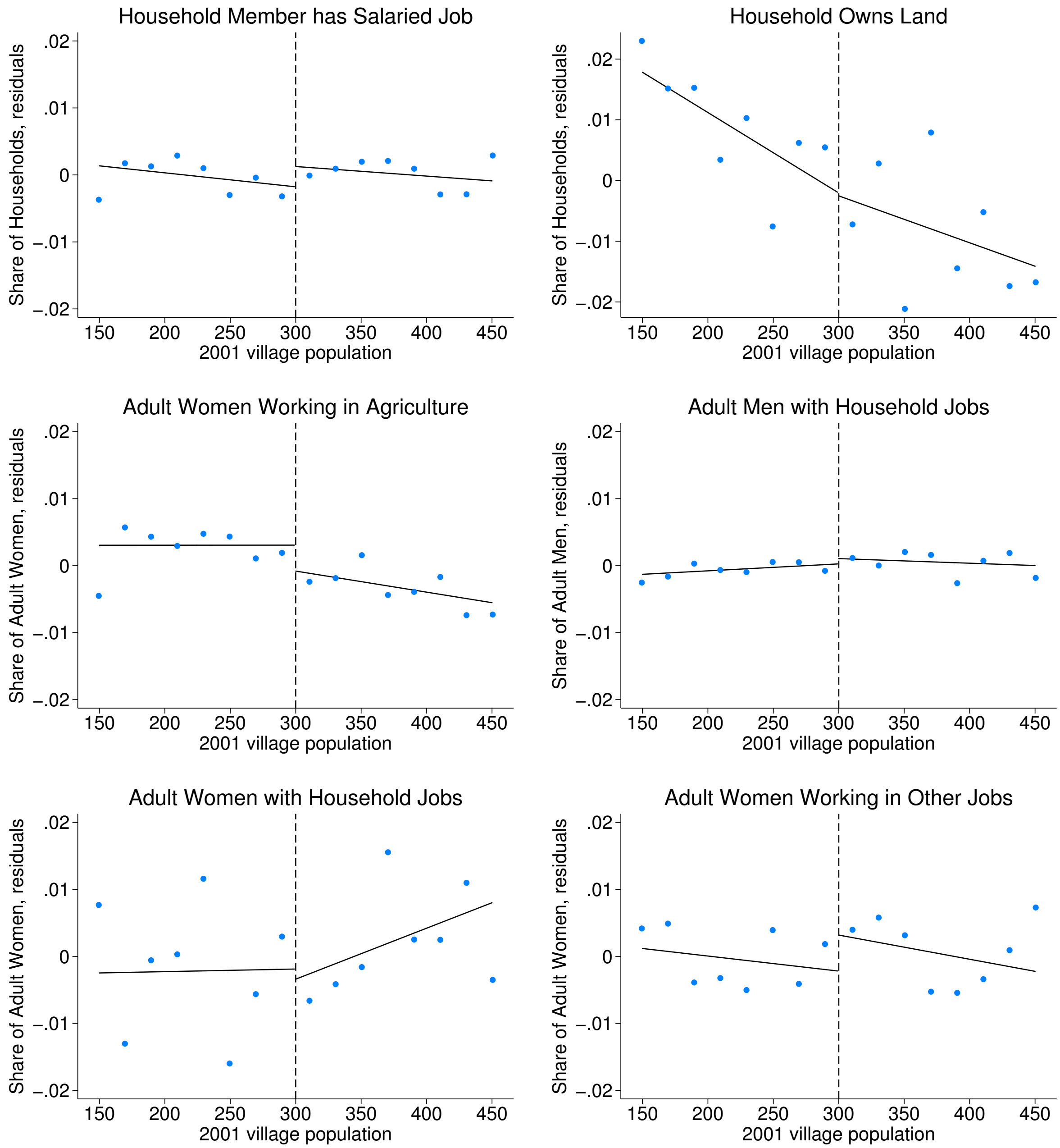
We might be concerned that when we average SECC outcomes to the village level for villages containing very few households with a poverty indicator in 2011, these averages are noisy and sensitive to outliers. Table B.2.21 tests for this possibility, by dropping villages with fewer than 10 percent of households in our SECC dataset in Panel A, and fewer than 20 percent of households in Panel B. In Panel C, we introduce 2001 controls to the SECC RD specifications, by selecting the 2001 Census variables that most closely align with each 2011 village level outcome. These sensitivities have very little effect on our results from Table 2.5.4.

Table B.2.22 presents results for additional village-level employment outcomes. In Panel A, we report results for the main source of household income earned by the household's main income earner. These income categories ("cultivation", "manual/casual labor", "non-farm enterprise") were recorded directly by SECC enumerators, and they do not map to the employment sectors used in other SECC outcomes. By contrast, the string-parsed employment categories for heads of household produce results very close to the adult male employment regressions in Panel B of Table 2.5.4. This is not surprising, as 83 precent household heads in our RD sample are adult men. Panel B of Table B.2.22 reports analogous results for youth (i.e. ages 0–16) employment, for the subset of households with at least 1 poverty indicator. We see that non-farm, non-household youth employment may increase slightly as a result of RGGVY eligibility, but we can reject effects larger than 1.6 percentage points.

We construct the employment categories "agricultural", "household", and "other" by string parsing occupations reported at the individual level and aggregating up to the household/village. While we try to recreate the three labor sectors reported in the Primary Census Abstract as closely as possible (in order to best facilitate apples-to-apples comparisons between Census and SECC outcomes), the "other" category contains an extremely wide range of occupations. In Panel C of Table B.2.22, we restrict this "other"

[59] The Imbens and Kalyanaraman (2012) optimal bandwidths range from 102 to 141 for these outcomes).

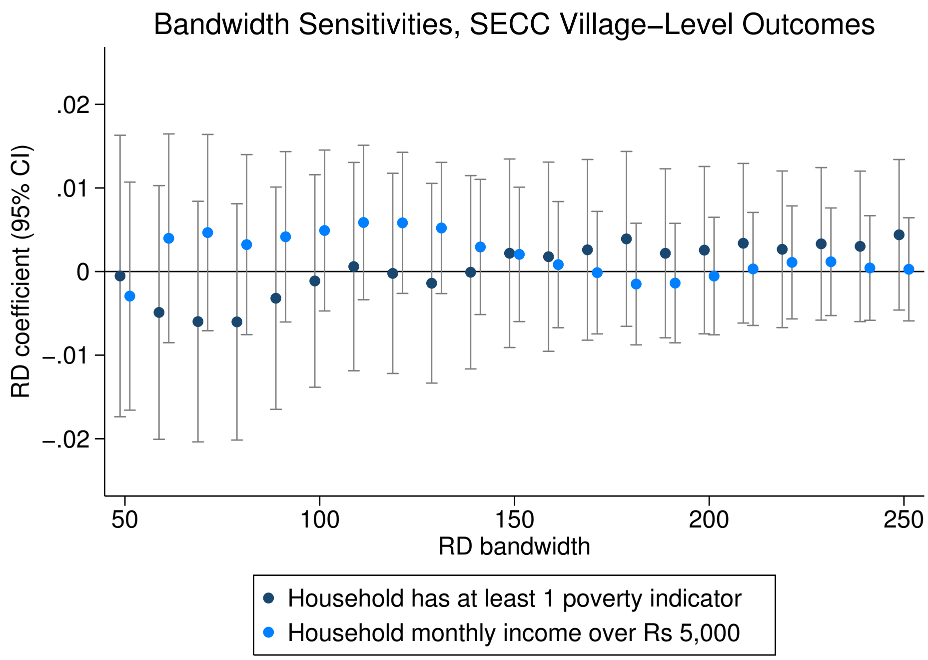
Figure B.2.16: RD Results – SECC Village-Level Outcomes



Notes: This figure presents RD results for SECC outcomes for our preferred specification, as a complement to Figure 2.5.9. They correspond to results reported in Table 2.5.4. Blue dots show average residuals from regressing the 2011 SECC village-level outcome (coded as the share of households in the village with at least one poverty indicator) on state fixed effects. Each dot contains approximately 1,600 villages, averaged in 20-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 10th-Plan single-habitation villages within our 150–450 population RD bandwidth, which match to the SECC dataset.

category to exclude occupations containing strings closely associated with manual labor: "labor" and "worker". This yields larger point estimates with larger t-stats, across all four gender/age combinations. Hence, this suggests that the (small) increases in "other" employment caused by RGGVY may represent a (small) shift towards relatively higher paying jobs outside of agriculture.

Figure B.2.17: RD Sensitivity – SECC, Bandwidths



Notes: This figure presents our bandwidth sensitivity analysis for two SECC outcomes (represented in the first two rows of Table 2.5.4). For each outcome, we estimate Equation (2.1) separately on bandwidths ranging from 50 (i.e., 250–350 people) to 250 (i.e., 50–550 people). Each dot represents the point estimate on the RD discontinuity at a given bandwidth around the 300-person cutoff, with 95 percent confidence intervals clustered at the district level. Our chosen bandwidth of 150 includes villages with populations between 150 and 450. The optimal RD bandwidth for these RD specifications ranges from 102 to 141 (calculated using the algorithm proposed by Imbens and Kalyanaraman (2012), using uniform, Epanechnikov, and triangular kernels).

Table B.2.21: RD Sensitivity – SECC Village-Level Outcomes

2011 Outcome	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Minimum 10 percent households with poverty indicator				
Monthly income greater than Rs 5,000	0.0017	(0.004)	[−0.006, 0.010]	0.08
One member holding salaried job	0.0030	(0.002)	[−0.001, 0.007]	0.02
Owning any land	0.0044	(0.009)	[−0.012, 0.021]	0.44
Male agricultural workers / adult men	−0.0108**	(0.005)	[−0.021, −0.000]	0.29
Female agricultural workers / adult women	−0.0021	(0.005)	[−0.011, 0.007]	0.08
Male household workers / adult men	0.0013	(0.001)	[−0.001, 0.004]	0.01
Female household workers / adult women	−0.0035	(0.007)	[−0.018, 0.010]	0.51
Male other workers / adult men	0.0056	(0.006)	[−0.007, 0.018]	0.42
Female other workers / adult women	0.0063	(0.005)	[−0.004, 0.016]	0.16
B. Minimum 20 percent households with poverty indicator				
Monthly income greater than Rs 5,000	0.0029	(0.004)	[−0.006, 0.011]	0.07
One member holding salaried job	0.0018	(0.002)	[−0.003, 0.006]	0.02
Owning any land	0.0049	(0.008)	[−0.011, 0.020]	0.43
Male agricultural workers / adult men	−0.0114**	(0.006)	[−0.022, −0.001]	0.29
Female agricultural workers / adult women	−0.0015	(0.005)	[−0.011, 0.008]	0.08
Male household workers / adult men	0.0011	(0.001)	[−0.002, 0.004]	0.01
Female household workers / adult women	−0.0044	(0.007)	[−0.018, 0.009]	0.51
Male other workers / adult men	0.0056	(0.007)	[−0.008, 0.019]	0.43
Female other workers / adult women	0.0064	(0.006)	[−0.004, 0.017]	0.17
C. Adding 2001 village controls				
At least one poverty indicator	0.0001	(0.006)	[−0.012, 0.013]	0.48
Monthly income greater than Rs 5,000	0.0053	(0.004)	[−0.003, 0.014]	0.08
One member holding salaried job	0.0030	(0.002)	[−0.002, 0.008]	0.02
Owning any land	−0.0005	(0.008)	[−0.017, 0.016]	0.44
Male agricultural workers / adult men	−0.0072	(0.005)	[−0.017, 0.003]	0.29
Female agricultural workers / adult women	−0.0036	(0.005)	[−0.013, 0.006]	0.08
Male household workers / adult men	0.0007	(0.001)	[−0.002, 0.004]	0.01
Female household workers / adult women	−0.0016	(0.008)	[−0.016, 0.013]	0.51
Male other workers / adult men	0.0044	(0.006)	[−0.008, 0.017]	0.42
Female other workers / adult women	0.0054	(0.005)	[−0.005, 0.016]	0.16

Notes: This table reports results of three sensitivity analyses on regressions reported in Table 2.5.4. Household-level outcomes are coded as the proportion of households with at least one poverty indicator, while adult employment outcomes are coded as the share of men (women) over 16 in households with a poverty indicator with an occupation in each sector. Each row represents a separate RD regression. Panel A includes villages where at least 10 percent of total households are included in our subset of the SECC data, resulting in 23,711 village observations. Panel B includes only villages where at least 20 percent of total households are included in our subset of the SECC data, restricting the analysis to only 21,072 village observations. Panel C includes all 25,942 SECC villages within our RD bandwidth, and also includes 2001 controls from the Census dataset. The share of households with a poverty indicator and monthly income regressions each include 13 village-wide controls from the 2001 Village Directory, while the salaried job and land regressions control for 2001 total other employment and 2001 village land area, respectively. The six employment regressions control for their corresponding village-wide employment shares from the 2001 PCA. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications include state fixed effects. Standard errors are clustered at the district level, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.22: RD Results – Additional SECC Village-Level Employment Outcomes

2011 Outcome	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
A. Share of households				
Main source of income: cultivation	0.0046	(0.007)	$[-0.010, 0.019]$	0.33
Main source of income: manual/casual labor	-0.0081	(0.007)	$[-0.022, 0.006]$	0.58
Main source of income: non-farm enterprise	-0.0006	(0.001)	$[-0.002, 0.001]$	0.01
Head of household occupation: agriculture	-0.0119^{*}	(0.007)	$[-0.025, 0.001]$	0.34
Head of household occupation: household work	0.0026	(0.003)	$[-0.003, 0.008]$	0.06
Head of household occupation: other work	0.0060	(0.007)	$[-0.008, 0.020]$	0.46
B. Youth employment				
Male youth ag workers / male youth	0.0006	(0.001)	$[-0.001, 0.003]$	0.01
Female youth ag workers / female youth	-0.0009	(0.001)	$[-0.003, 0.001]$	0.01
Male youth household workers / male youth	-0.0006	(0.001)	$[-0.003, 0.002]$	0.01
Female youth household workers / female youth	-0.0007	(0.001)	$[-0.004, 0.002]$	0.02
Male youth other workers / male youth	0.0070	(0.005)	$[-0.002, 0.016]$	0.07
Female youth other workers / female youth	0.0067	(0.004)	$[-0.001, 0.015]$	0.07
C. Excluding manual labor				
Adult male other workers / adult men	0.0107^{***}	(0.004)	$[0.003, 0.019]$	0.15
Adult female other workers / adult women	0.0104^{***}	(0.004)	$[0.003, 0.018]$	0.08
Male youth other workers / male youth	0.0086^{*}	(0.005)	$[-0.000, 0.017]$	0.06
Female youth other workers / female youth	0.0080^{*}	(0.004)	$[-0.000, 0.016]$	0.06

Notes: Each row represents a separate regression estimating Equation (2.1) on a different SECC village-level outcome. Outcomes are coded as the proportion of households (individuals) in each village, conditional on the (individual belonging to a) household having at least one poverty indicator in 2011. In Panel A, we report household income by source, based on categories coded in the SECC data; household head occupations are coded via string parsing each individual's reported occupation. Panel B presents youth employment results analogous to the adults labor results presented in Panel B of Table 2.5.4, defining all individuals less than 16 years old as "youth". Panel C reports adult and youth employment results where we narrow the definition of "other" to exclude occupations with "labor" or "worker" in their description. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications include state fixed effects, but they do not include any additional baseline control variables. The RD bandwidth includes 25,942 villages with 2001 populations between 150 and 450. These regressions contain fewer villages than regressions in Table 2.5.3 because only 87 percent of 10th-Plan, single-habitation, 150–450 villages match to the SECC dataset. Standard errors are clustered at the district level with 222 clusters, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.2.8 School Enrollment Results: RD Robustness

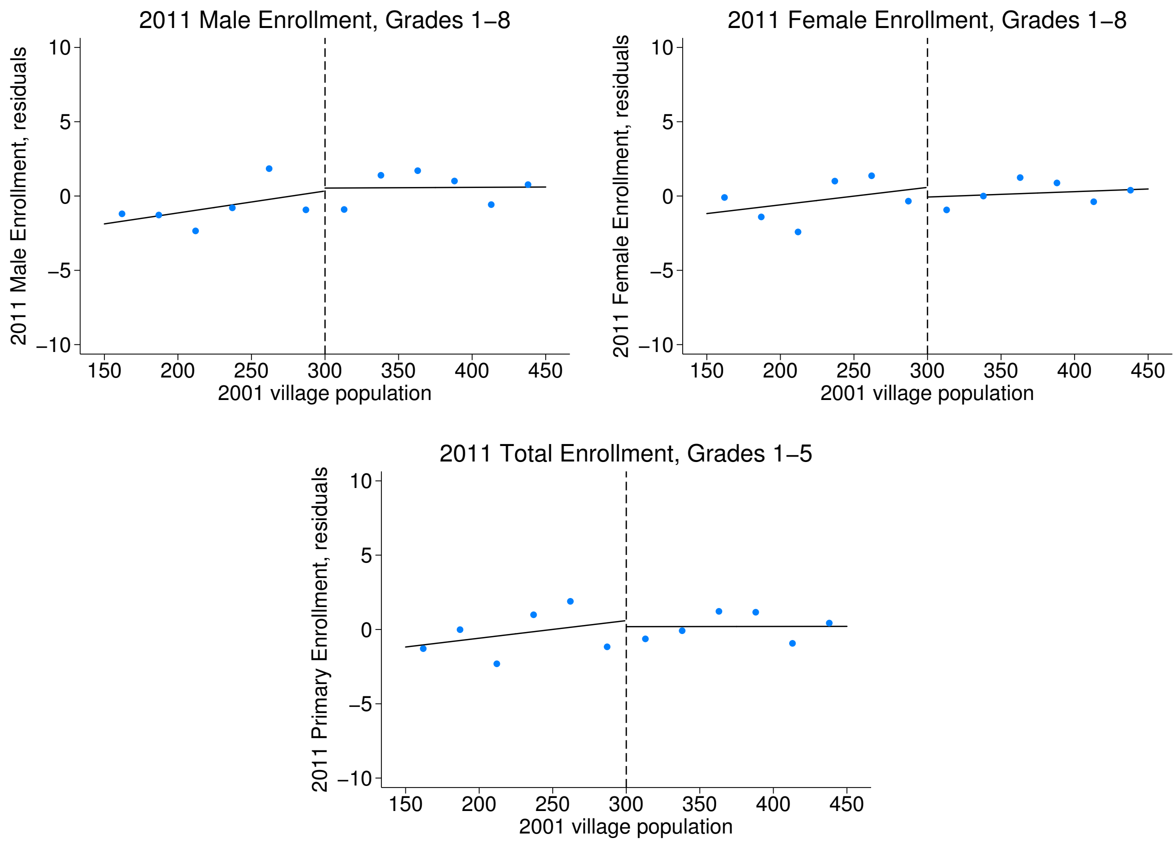
Table 2.5.5 reports results for five measures of village-level school enrollment, using our preferred RD specification. Figure 2.5.10 includes RD plots for two of these five outcomes, and we report the remaining three RD plots below in Figure B.2.18. We see no visual evidence of a discontinuity at the RD threshold, confirming our results in Table 2.5.5 that reject economically significant increases in school enrollment as a result of RGGVY electrification.

As an alternative RD specification, we re-estimate these enrollment regressions at the school level. Instead of aggregating enrollment counts across all schools in a village up to a single village-level observation, these regressions treat each school as a separate observation.[60] Table B.2.23 reports results for these school-level RD regressions, while Figure B.2.19 shows the analogous RD graphs. These estimates are very similar to the village-level results, with confidence intervals that reject 10 percent changes in enrollment as a result of RGGVY eligibility. The school-level RD plots in Figure B.2.19 likewise show no evidence of a discontinuity at the 300-person threshold.

Figure B.2.20 conducts RD bandwidth sensitivities using the total enrollment outcome, for both the village-level and school-level specifications. This demonstrates that these RD results are not sensitive to bandwidths above 100 people, which is below the smallest optimal bandwidth calculation for these outcomes (Imbens and Kalyanaraman (2012) optimal bandwidths range from 138 to 179). This provides further evidence in support of the assumptions underpinning our RD design. Table B.2.24 conducts additional specification sensitivities for our village-level enrollment RD. This shows that our RD point estimates are not sensitive to our choice of outcome year (i.e. total enrollment for the school year beginning in 2010, 2011, or 2012), control year (i.e. pre-RGGVY enrollment for the school year beginning in 2005 or 2006), or sample restriction (i.e. removing village-school matches most likely to be inaccurate). Across all sensitivities, we can reject even moderate increases to school enrollment around our 300-person RD threshold.

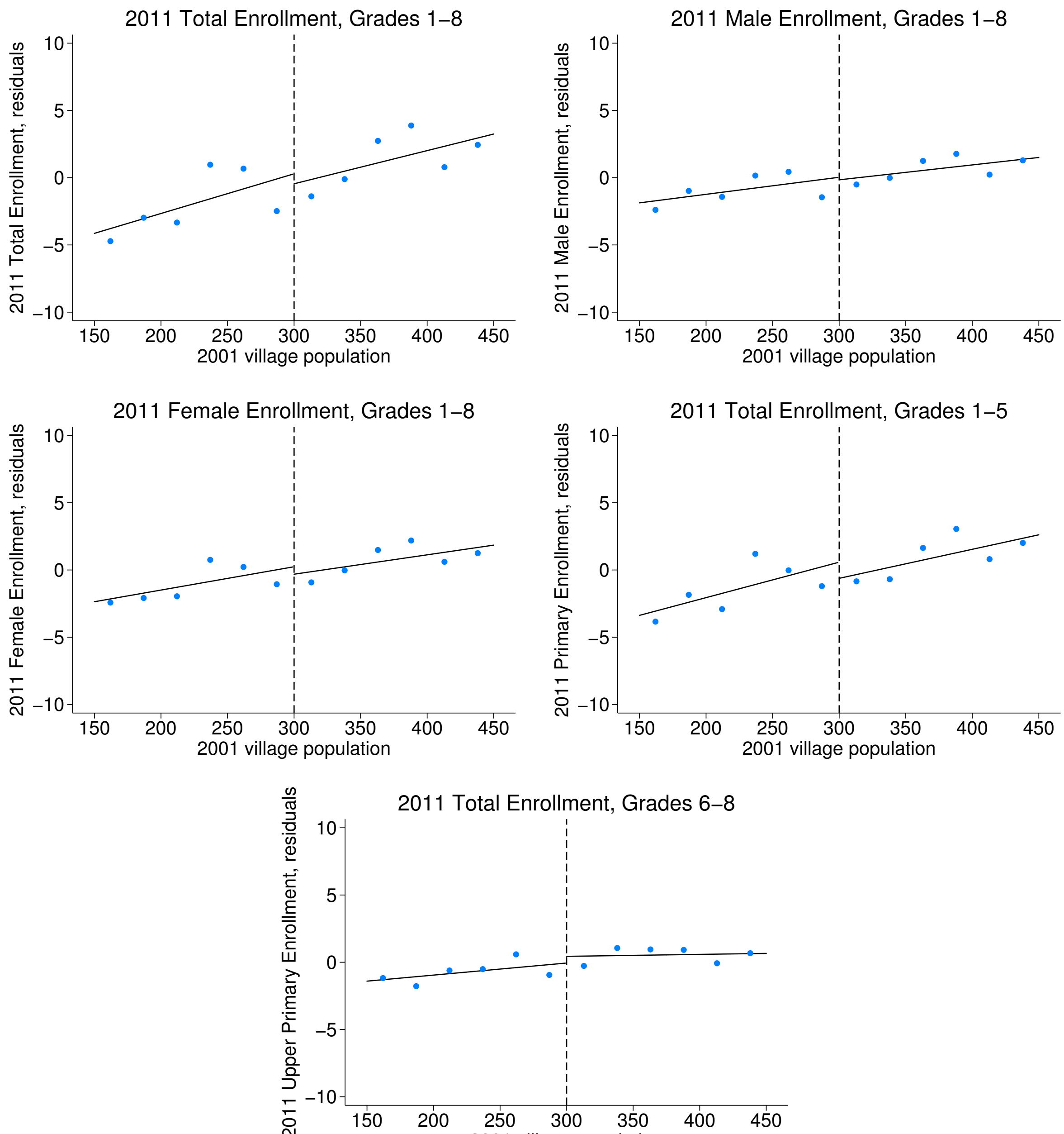
[60] 25 percent of the 15,215 RD bandwidth villages that matched to the DISE schools dataset (via the fuzzy matching algorithm detailed in Section B.1.7) matched to multiple schools reporting enrollment counts in 2005 and 2011.

Figure B.2.18: RD Results – School Enrollment



Notes: This figure presents RD results for school enrollment counts for our preferred specification, as a complement to Figure 2.5.10. They correspond to results reported in Table 2.5.5. Blue dots show average residuals form regressing the 2011 village-level enrollment on 2005 village-level enrollment and state fixed effects. Each dot contains approximately 1,000 villages, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 10th-Plan single-habitation villages within our 150–450 population RD bandwidth, with school-village matches and nonmissing enrollment counts for both 2005 and 2011.

Figure B.2.19: RD Sensitivity – School-Level Enrollment Regressions



Notes: This figure presents RD results for school-level enrollment, corresponding to results reported in Table B.2.23. Blue dots show average residuals form regressing the 2011 school enrollment on 2005 enrollment and state fixed effects. Each dot contains approximately 1,000 schools, averaged in 25-person population bins. Lines are estimated separately on each side of the 300-person threshold, for all 10th-Plan single-habitation villages within our 150–450 population RD bandwidth, with schools reporting nonmissing enrollment data for both 2011 and 2005.

Table B.2.23: RD Sensitivity – School Enrollment, School-Level Regressions

2011 Outcome Variable	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome
Total enrollment, grades 1–8	−0.742	(2.30)	[−5.25, 3.77]	61.19
Male enrollment, grades 1–8	−0.189	(1.13)	[−2.39, 2.02]	31.07
Female enrollment, grades 1–8	−0.556	(1.22)	[−2.95, 1.84]	30.12
Total enrollment, grades 1–5	−1.205	(1.98)	[−5.08, 2.67]	50.06
Total enrollment, grades 6–8	0.499	(0.75)	[−0.97, 1.97]	11.13

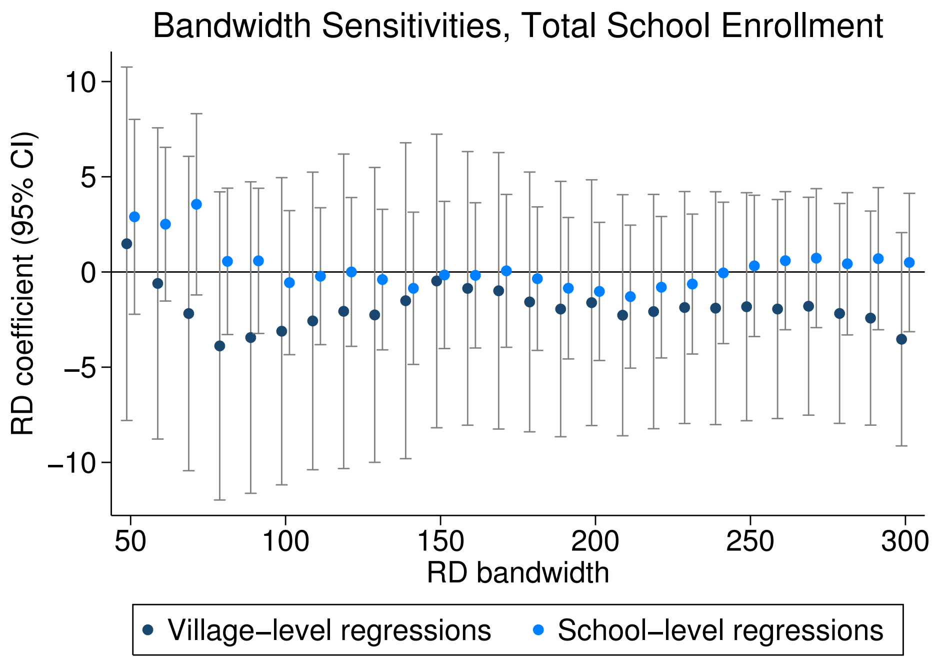
Notes: Each row represents a separate regression estimating Equation (2.1) at the school level, on a different enrollment count. The second column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for the 2005 level of the outcome variable and state fixed effects. The RD bandwidth includes 13,150 school-level observations, across 11,578 villages with 2001 populations between 150 and 450. These regressions contain fewer villages than regressions in Table 2.5.5 because some villages have nonmissing enrollment counts for for 2011 and 2005 only after summing across multiple schools (even though no single school in these villages has nonmissing data for both years). Standard errors are clustered at the district level, with 215 clusters, which we use to calculate 95 percent confidence intervals in the fourth column. The fifth column reports the mean of the dependent variable for each RD regression. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.24: RD Sensitivity – Total Grade 1–8 Enrollment, Village-Level Regressions

Outcome Year	Control Year	Subsample	RD Coeff	Std Error	95 Percent Confidence	Mean of Outcome	Village Obs
2011	2005		−0.472	(3.93)	[−8.18, 7.24]	74.05	12,251
2011	2006		2.196	(3.31)	[−4.30, 8.69]	74.05	12,651
2010	2005		−1.242	(3.66)	[−8.42, 5.94]	75.96	12,290
2012	2005		−2.952	(4.97)	[−12.69, 6.78]	69.94	12,205
2010–12	2005–06		0.528	(3.19)	[−5.73, 6.79]	73.30	12,801
2011	2005	VD school	0.825	(3.48)	[−6.00, 7.65]	66.21	8,837
2011	2005	PIN code	−1.449	(4.49)	[−10.25, 7.35]	73.81	9,560

Notes: This table conducts sensitivity analysis on the RD results for total school enrollment, at the village level. The first row reproduces our preferred specification from the top row of Table 2.5.5. The next three rows report results for the same regression, using outcomes and controls from adjacent years. Due to the unbalanced nature of DISE school panel dataset, nonmissing village observation counts are sensitive to the choice of outcome/control year. The fifth row averages village enrollment across 2010–2012 nonmissing values and controls across 2005–2006 nonmissing enrollment, which allows us to include 4 percent more schools than our preferred specification. The bottom two rows restrict our sample of village-schools matches to include only villages reported to have schools in the 2011 Village Directory; and only villages with Pincodes that match those reported in the DISE dataset. The fourth column shows the RD point estimate ($\hat{\beta}_1$) for each regression. All specifications control for total enrollment from 2005 (or 2006) and state fixed effects. The RD bandwidth includes single-habitation villages in 10th-Plan districts with 2001 populations between 150 and 450. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure B.2.20: RD Sensitivity – School Enrollment, Bandwidths



Notes: This figure presents our bandwidth sensitivity analysis for total school enrollment (grades 1–8), at the village level (represented in the first row of Table 2.5.5) and at the school level (represented in the first row of Table B.2.23). For each outcome, we estimate Equation (2.1) separately on bandwidths ranging from 50 (i.e., 250–350 people) to 300 (i.e., 0–600 people). Each dot represents the point estimate on the RD discontinuity at a given bandwidth around the 300-person cutoff, with 95 percent confidence intervals clustered at the district level. Our chosen bandwidth of 150 includes villages with populations between 150 and 450. The optimal RD bandwidth for these RD specifications ranges from 138 to 179 (calculated using the algorithm proposed by Imbens and Kalyanaraman (2012), using uniform, Epanechnikov, and triangular kernels).

B.2.9 Spatial Spillovers

Villages economies do not exist in isolation, and it is important to consider potential spillover effects of electrification on neighboring villages. Below, we test for spatial spillovers by modifying Equation (2.1) such that the dependent variable is the average of each 2011 outcome across villages within 10-, 20-, and 50-km radii of each village in our RD sample. These regressions still treat single-habitation 10th-Plan villages with 2001 populations between 150–450 as the unit of analysis; they simply test for effects on RGGVY eligible on outcomes in other surrounding villages.[61]

Table B.2.25 reports these results, where we find little evidence of spatial spillovers. Compared to our main specification (labeled "0 km"), spillover results attenuate at 10 km. Weakly significant effects within a 10km radius would not necessarily provide evidence of spillovers, given the measurement error inherent to assigning villages to shapefiles and calculating spatial averages in GIS. As we have no reason to suspect different degrees of measurement error across outcomes, we conclude that economically meaningful spatial spillovers are unlikely.

B.2.10 Heterogeneous RGGVY Implementation

One possible explanation for the small magnitudes of our RD results could be that RGGVY only impacted a subset of villages/districts/states in our sample. If this is were the case, and our RD estimates pooled villages with strong treatment effects and villages with no treatment effects, this would produce small average treatment effects. Below, we employ two strategies of subsampling our RD sample, by isolating districts and states most likely to demonstrate economically significant impacts from RGGVY.

First, we exploit the gradual rollout of RGGVY implementation under the 10th Plan. As shown in Table B.1.2, 10th-Plan districts received RGGVY funding as early as 2005 and as late as 2010. Even though the latest 10th-Plan funding predates our 2011 outcome data, it is quite possible that these 2011 data do not reflect the full impacts of electrification in districts where RGGVY implementation began in 2009 or 2010. This is especially likely to be true for medium-run economic outcomes. Table B.2.26 estimates Equation (2.1) on subsamples of RGGVY 10th-Plan districts that received RGGVY funds before 2007, 2008, 2009, and 2010. We see that while over half of 10th-Plan districts received funding in 2005 or 2006 (i.e. before 2007), this subsample yields RD point estimates very close to the full-sample averages. Variation in the timing of RGGVY rollout is unlikely to be obscuring large effects of electrification in districts with early RGGVY implementation.

Second, we consider heterogeneous power quality across states. Poor power quality could help to explain the small magnitudes of our economic results — newly electrified villages can only benefit from electricity infrastructure if power reliably flows through the grid. As a proxy for power quality at the state level, we use electricity demand surpluses/deficits as reported in the 2011–2012 Load Generation Balance Report (Central

[61] We exclude RD-sample villages when taking these spatial averages, in order to ensure that no village is simultaneously represented on the left-hand side and the right-hand side of these regressions.

Table B.2.25: Spatial Spillovers to Adjacent Villages

	Radius around within-bandwidth village			
2011 Outcome Variable	0 km	10 km	20 km	50 km
2011 nighttime brightness	0.1493**	0.0529	0.0475**	0.0178
	(0.0603)	(0.0360)	(0.0241)	(0.0189)
Male agri. workers / male pop	−0.0065**	−0.0004	−0.0004	0.0001
	(0.0033)	(0.0009)	(0.0007)	(0.0004)
Female agri. workers / female pop	−0.0051	−0.0012	−0.0004	−0.0002
	(0.0049)	(0.0015)	(0.0011)	(0.0006)
Male other workers / male pop	0.0056**	0.0009	0.0007	0.0003
	(0.0023)	(0.0007)	(0.0005)	(0.0003)
Female other workers / female pop	−0.0006	0.0010	0.0007	0.0002
	(0.0026)	(0.0009)	(0.0007)	(0.0004)

Notes: This table estimates Equation (2.1) on main RD samples, using the average 2011 outcomes of adjacent villages. For each single-habitation, 10th-Plan district within our 150–450 bandwidth, we calculate the average level of each outcome varaible for all adjacent villages *not* in the RD sample, within a 10-, 20-, and 50-kilometer radius of the village centroid. In each row, we present RD point estimates ($\hat{\beta}_1$) from four separate regressions — three regressions of the average outcome for adjacent villages within a given radius, and the main specification (i.e. 0 km) in the first column. Both nighttime brightness and labor regressions include 18,686 village observations, restricting the RD sample to the 12 states with available shapefiles that correlate with village areas. All specifications control for the 2001 level of the outcome variable and state fixed effects. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Electricity Authority (2011)). This defines demand surplus/deficit as the percentage of required electricity load that is available to the state. For the 12 states in our RD sample for nighttime brightness, this measure ranged from a 17.3 percent surplus in Chhattisgarh to a 19.4 percent deficit in Madhya Pradesh, with a 2011 national average shortfall of 10.3 percent of total load.

To test for heterogeneous effects in power quality, we estimate Equation (2.1) on the 7 states in our 12-state RD sample with 2011–2012 shortfalls that were better than the national average. These states are, in order of lowest-to-highest deficit (i.e. negative surplus): Chhattisgarh (−17.3 percent), Orissa (−15.4), Karnataka (−4.8), West Bengal (0.0), Gujarat (1.6), Haryana (6.0), and Rajasthan (7.0). Table B.2.27 shows that restricting the RD sample to these 7 states increases our RD point estimate for nighttime brightness by a factor of 1.7, without sacrificing precision. However, applying the same restriction to RD regressions on labor outcomes if anything attenuates our point estimates while barely affecting their confidence intervals. This demonstrates that even in states where RGGVY investments were likely coupled with above-average power availability, we can still reject economically significant changes in employment. Table B.2.28 presents analogous results for asset ownership, housing outcomes, and village-wide outcomes, while Tables B.2.29 and B.2.30 do so for the outcomes in Tables 2.5.4 and 2.5.5, respectively. The confidence intervals for the low-deficit sample remain broadly similar to those of both full 22-state RD sample and the 12-state lights sample, even at 32 percent smaller sample size.[62]

There are three exceptions. First, the share of households owning bicycles in Table B.2.28 is negative and statistically significant in the low-deficit sample. This result appears to be spurious, as it is not robust across other assets that require electricity. Second, the adult male other employment result in Table B.2.29 is larger and statistically significant in the low-deficit sample. While the 0.034 upper bound of the 95 percent confidence interval is twice as large as with the full sample of 22 states, 0.034 still represents a relatively small change in the share of adult male workers.[63] Third, the five school enrollment results in Table B.2.30 are negative and statistically significant for low-deficit states, suggesting that RGGVY led to *decreases* in student enrollment. We report six corresponding RD pictures in Figure B.2.21, and only the SECC adult male other employment result reveals visual evidence of a discontinuity at the 300-person threshold.

RGGVY implementation may have been heterogeneous in other ways that limited its effectiveness. Given that district-specific RGGVY projects were often carried out by

[62]For simplicity, B.2.27–B.2.30 use the 12-state RD lights sample for all outcome regressions. This excludes 11,079 villages in 10 states with low-quality or missing shapefiles, which are included in all other RD regressions on Census outcomes. 93 percent of these excluded villages are in Uttar Pradesh and Uttarakhand, two states with deficits well above the national average. Hence, the 12 vs. 22 state distinction is unimportant in this split-sample exercise, because the excluded states with below-average deficits are all very small.

[63]Our labor share results from Table 2.5.3 use the village's full male population as the denominator. For comparison, a 3.4 percentage point increase in adult men in households with at least one poverty indicator translates to less than a 1 percentage point increase in the full male population. This is because for the average village in our sample, 67 percent of the male population is 16 or older, and only 44 percent of the male population is included in our SECC dataset.

decentralized implementing agencies (i.e. state electricity boards, local distribution companies), the efficacy of project implementation might have varied widely across states or districts.[64] Even if all implementation efforts were identical, enforcement of the 300-person eligibility cutoff might have varied across implementers, which could reduce the power of our RD design. However, we lack a strong prior as to which states/districts are likely to have most strongly enforced the 300-person rule and most effectively implemented RGGVY projects. Interestingly, if we cherry-pick the 7 states with the largest RD point estimates for 2011 nighttime brightness, 6 of the 7 cherry-picked states also had below-average electricity deficits in 2011. This suggests that *if* poor implementation prevented barely eligible RGGVY villages in certain states from exhibiting increased nighttime brightness, these states were likely to have been states with above-average power shortfalls.

[64] Apart from implementer-specific effects, this could reflect socioeconomic or political differences across states. For example, newspaper stories have cited ethnic conflict as having caused RGGVY implementation delays in the northeastern states of Bihar, Assam, and Jharkhand.

Table B.2.26: Subsample – Districts Receiving Early RGGVY Funding

2011 Outcome Variable	Received RGGVY funding before 2007	2008	2009	2010
2011 nighttime brightness	0.1414	0.1172*	0.1172*	0.1493**
	(0.0977)	(0.0609)	(0.0609)	(0.0603)
Number of villages	10,833	17,960	17,960	18,686
Number of districts	90	126	126	130
Number of states	10	12	12	12
Male agri. workers / male pop	−0.0081**	−0.0063**	−0.0070**	−0.0073***
	(0.0035)	(0.0029)	(0.0028)	(0.0028)
Female agri. workers / female pop	−0.0056	−0.0036	−0.0042	−0.0050
	(0.0049)	(0.0041)	(0.0041)	(0.0041)
Male other workers / male pop	0.0047**	0.0041**	0.0046**	0.0046**
	(0.0024)	(0.0020)	(0.0020)	(0.0019)
Female other workers / female pop	0.0016	−0.0004	−0.0003	−0.0004
	(0.0022)	(0.0021)	(0.0020)	(0.0020)
Number of villages	20,958	28,489	28,973	29,703
Number of districts	167	211	217	223
Number of states	12	18	20	21

Notes: This table estimates Equation (2.1) on subsamples of villages, in districts that received RGGVY 10th-Plan funding *before* a given year. All 10th-Plan districts received funding before the end of 2010, and only 2 districts in Nagaland received funding after January 1, 2010 (and are not included in the rightmost column above). In each row, we present RD point estimates ($\hat{\beta}_1$) from four separate regressions on subsamples of single-habitation, 10th-Plan villages within our RD bandwith. The number of villages/districts/states differ across nighttime brightness and labor regressions, because our nighttime lights sample includes only the 12 states with available shapefiles that correlate with village areas. All specifications control for the 2001 level of the outcome variable and state fixed effects. Standard errors are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.27: Subsample – States with Low Power Deficits (Lights and Labor)

2011 Outcome Variable	Full RD Sample	Full Lights Sample	Low-Deficit States
2011 nighttime brightness		0.1493^{**}	0.2481^{***}
		(0.0603)	(0.0737)
		$[0.031, 0.268]$	$[0.104, 0.393]$
Number of villages		18,686	12,679
Number of districts		130	67
Number of states		12	7
ag workers / male pop	-0.0071^{**}	-0.0065^{**}	-0.0043
	(0.0028)	(0.0033)	(0.0039)
	$[-0.013, -0.002]$	$[-0.013, -0.000]$	$[-0.012, 0.003]$
ag workers / female pop	-0.0049	-0.0051	-0.0040
	(0.0040)	(0.0049)	(0.0061)
	$[-0.013, 0.003]$	$[-0.015, 0.005]$	$[-0.016, 0.008]$
oth workers / male pop	0.0046^{**}	0.0056^{**}	0.0054^{**}
	(0.0019)	(0.0023)	(0.0025)
	$[0.001, 0.008]$	$[0.001, 0.010]$	$[0.000, 0.010]$
Female oth workers / female pop	-0.0004	-0.0006	-0.0006
	(0.0020)	(0.0026)	(0.0036)
	$[-0.004, 0.004]$	$[-0.006, 0.005]$	$[-0.008, 0.006]$
Number of villages	29,765	18,686	12,679
Number of districts	225	130	67
Number of states	22	12	7

Notes: This table estimates Equation (2.1) on our full RD sample (in the first column), our 12-state lights RD sample (in the second column), and a subsample of 7 states with the lowest reported electricity demand shortfalls for 2011 (in the third column). We define this demand shortfall as the percent of total electricity demand not met by each state (Central Electricity Authority (2011)). These 7 states are (in order of increasing demand shortfall) Chhattisgarh, Orissa, Karnataka, West Bengal, Gujarat, Haryana, and Rajasthan. In each row, we present RD point estimates ($\hat{\beta}_1$) from three separate regressions, include all villages in the full RD sample, the nighttime lights RD sample or the 7-state subset of the nighttime lights RD sample (the latter two samples exclude states with shapefiles that are either unavailable or uncorrelated with village areas). All specifications control for the 2001 level of the outcome variable and state fixed effects. We report 95 percent confidence intervals in brackets. Standard errors (in parentheses) are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.2.28: Subsample – States with Low Power Deficits (Assets, etc.)

2011 Outcome Variable	Full RD Sample	Full Lights Sample	Low-Deficit States
C. Asset ownership			
Share of households with telephone	0.0025	−0.0007	−0.0010
	[−0.008, 0.013]	[−0.014, 0.013]	[−0.018, 0.016]
Share of households with TV	0.0026	0.0041	0.0016
	[−0.005, 0.010]	[−0.004, 0.013]	[−0.008, 0.011]
Share of households with bicycle	−0.0015	−0.0048	−0.0116**
	[−0.010, 0.007]	[−0.016, 0.006]	[−0.023, −0.000]
Share of households with motorcycle	−0.0008	−0.0008	−0.0020
	[−0.006, 0.004]	[−0.007, 0.005]	[−0.010, 0.006]
Share of households without assets	0.0039	0.0060	0.0098
	[−0.004, 0.012]	[−0.005, 0.017]	[−0.003, 0.022]
D. Housing stock			
Share of households w/ elec/gas cooking	0.0005	−0.0013	−0.0045
	[−0.005, 0.006]	[−0.006, 0.003]	[−0.010, 0.001]
Share of households w/ kerosene lighting	0.0029	−0.0047	0.0027
	[−0.009, 0.015]	[−0.020, 0.011]	[−0.009, 0.014]
Share of households with mud floors	0.0043	0.0046	0.0040
	[−0.003, 0.012]	[−0.004, 0.013]	[−0.007, 0.015]
Share of households with thatched roof	−0.0034	−0.0045	0.0037
	[−0.013, 0.007]	[−0.016, 0.007]	[−0.008, 0.015]
Share of households dilapidated	−0.0031	−0.0060*	−0.0060
	[−0.009, 0.002]	[−0.013, 0.001]	[−0.014, 0.003]
E. Village-wide outcomes			
1/0 Mobile phone coverage in village	−0.0008	−0.0006	−0.0193
	[−0.023, 0.021]	[−0.030, 0.029]	[−0.049, 0.011]
1/0 Post office in village	0.0017	0.0002	0.0001
	[−0.005, 0.009]	[−0.008, 0.008]	[−0.009, 0.009]
1/0 Ag credit societies in village	0.0013	−0.0032	−0.0035
	[−0.006, 0.008]	[−0.011, 0.005]	[−0.011, 0.004]
1/0 Water from tubewell in village	−0.0075	−0.0170	−0.0290
	[−0.036, 0.021]	[−0.050, 0.016]	[−0.065, 0.007]
Share of village area irrigated	−0.0057	−0.0033	−0.0051
	[−0.016, 0.004]	[−0.011, 0.004]	[−0.014, 0.003]
Share of village area planted	0.0015	0.0064	0.0036
	[−0.010, 0.013]	[−0.008, 0.021]	[−0.014, 0.022]

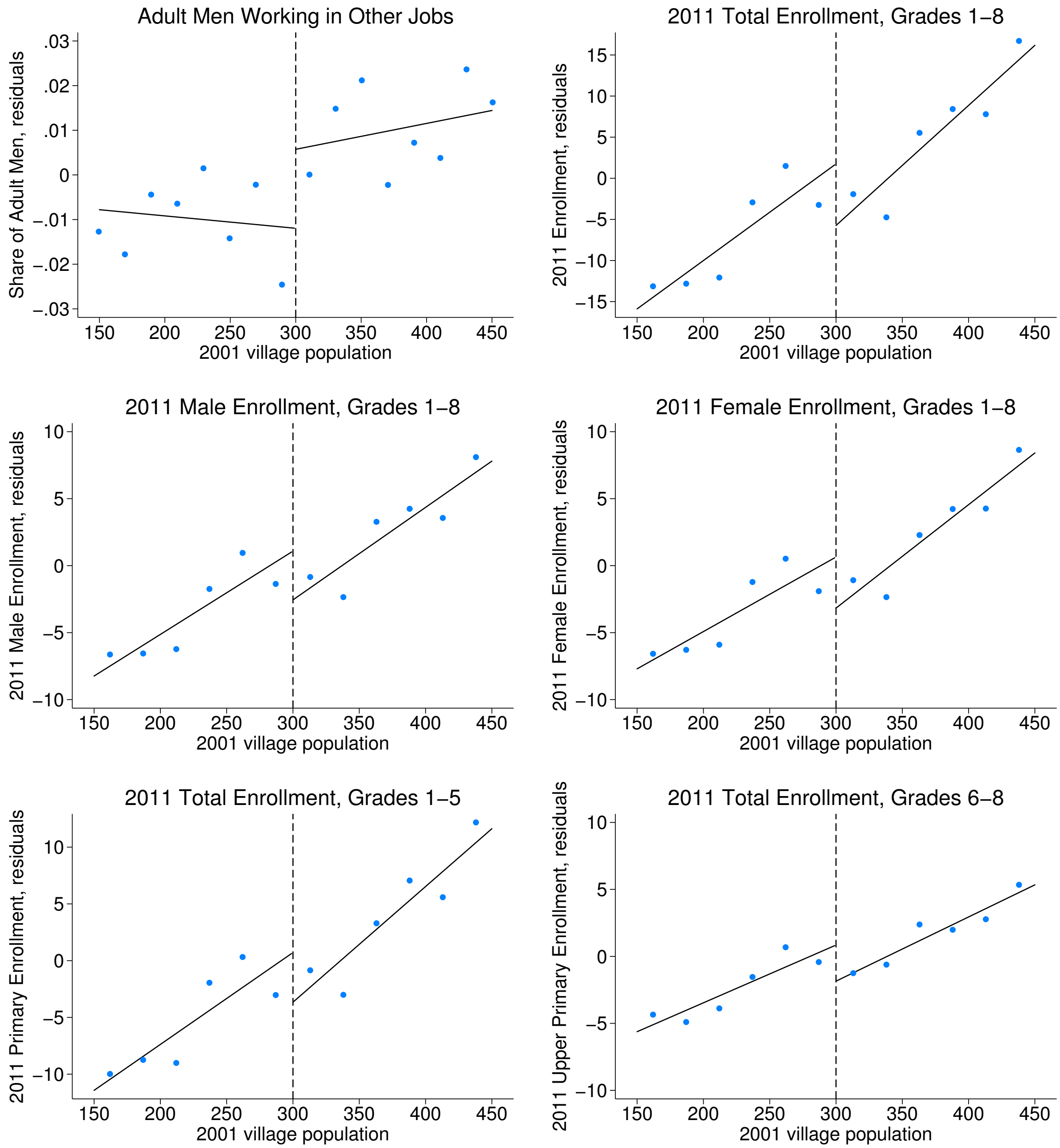
Notes: This table is exactly analogous to Table B.2.27, except that it reports full sample, lights sample, and low-deficit-states results for outcomes in Panels C, D, and E of Table 2.5.3. Standard errors are omitted for brevity, but 95 percent confidence intervals are still presented in brackets. Please refer to the notes below Table B.2.27.

Table B.2.29: Subsample – States with Low Power Deficits (SECC Outcomes)

2011 Outcome Variable	Full RD Sample	Full Lights Sample	Low-Deficit States
A. Share of households			
At least one poverty indicator	0.0006	−0.0034	−0.0021
	[−0.011, 0.012]	[−0.019, 0.012]	[−0.022, 0.018]
Monthly income greater than Rs 5,000	0.0043	0.0069	0.0059
	[−0.004, 0.013]	[−0.003, 0.017]	[−0.006, 0.018]
One member holding salaried job	0.0030	0.0046**	0.0029
	[−0.002, 0.008]	[0.000, 0.009]	[−0.002, 0.008]
Owning any land	−0.0005	−0.0007	0.0068
	[−0.017, 0.016]	[−0.015, 0.014]	[−0.011, 0.024]
B. Adult employment			
Male agricultural workers / adult men	−0.0091*	−0.0078	−0.0047
	[−0.019, 0.001]	[−0.021, 0.005]	[−0.022, 0.013]
Female agri. workers / adult women	−0.0039	−0.0002	−0.0001
	[−0.013, 0.006]	[−0.012, 0.012]	[−0.017, 0.017]
Male household workers / adult men	0.0008	0.0013	0.0014
	[−0.002, 0.004]	[−0.002, 0.005]	[−0.004, 0.007]
Female hhold. workers / adult women	−0.0015	0.0065	0.0111
	[−0.016, 0.013]	[−0.014, 0.027]	[−0.019, 0.041]
Male other workers / adult men	0.0052	0.0124*	0.0177**
	[−0.007, 0.017]	[−0.002, 0.027]	[0.001, 0.034]
Female other workers / adult women	0.0054	0.0017	0.0033
	[−0.005, 0.016]	[−0.011, 0.014]	[−0.010, 0.017]

Notes: This table is exactly analogous to Table B.2.27, except that it reports full RD sample, lights RD sample, and low-deficit-state results for outcomes in Table 2.5.4. Regressions in the first column contain 25,942 village observations (as in Table 2.5.4), while regressions in the second and third columns contain 16,240 and 11,027 village observations, respectively. All regressions control for state fixed effects, but do not include any additional controls. Standard errors are omitted for brevity, but 95 percent confidence intervals are still presented in brackets. Please refer to the notes below Table B.2.27.

Figure B.2.21: RD Sensitivity – Selected Regressions, Low-Deficit States



Notes: This figure presents RD results for the subset of states with above-average power quality (i.e., below-average deficits). They correspond to the adult male other workers regression in Table B.2.29, and all five village-level enrollment regressions in Table B.2.30. Only the upper-left RD plot for the share of adult men working in "other" job reveals visual evidence of a discontinuity, even though all six RD point estimates are statistically significant. Blue dots show average residuals form regressing the 2011 outcome on state fixed effects and the 2005 level of the outcome (for enrollment variables only). Each dot in the upper-left plot contains approximately 700 villages, averaged in 20-person population bins. For the five enrollment plots, each dot contains approximately 450, averaged in 25-person population bins.

Table B.2.30: Subsample – States with Low Power Deficits (DISE Outcomes)

2011 Outcome Variable	Full RD Sample	Full Lights Sample	Low-Deficit States
Total enrollment, grades 1–8	-0.4725	-2.7663	-7.4681^{**}
	$[-8.182, 7.237]$	$[-8.174, 2.642]$	$[-13.325, -1.611]$
Male enrollment, grades 1–8	0.1966	-1.5136	-3.6321^{**}
	$[-3.715, 4.108]$	$[-4.304, 1.277]$	$[-6.737, -0.527]$
Female enrollment, grades 1–8	-0.6504	-1.1691	-3.7973^{***}
	$[-4.612, 3.311]$	$[-3.910, 1.572]$	$[-6.678, -0.917]$
Total enrollment, grades 1–5	-0.4080	-0.8513	-4.3510^{**}
	$[-6.191, 5.375]$	$[-4.848, 3.146]$	$[-8.133, -0.569]$
Total enrollment, grades 6–8	0.0513	-1.4067	-2.6964^{*}
	$[-2.892, 2.994]$	$[-4.083, 1.270]$	$[-5.815, 0.423]$

Notes: This table is exactly analogous to Table B.2.27, except that it reports full RD sample, lights RD sample, and low-deficit-state results for outcomes in Table 2.5.5. We report the corresponding RD figures in B.2.21. Regressions in the first column contain 12,251 village observations (as in Table 2.5.5), while regressions in the second and third columns contain 8,569 and 5,482 village observations, respectively. All regressions control for state fixed effects and the 2005 level of the outcome variable. Standard errors are omitted for brevity, but 95 percent confidence intervals are still presented in brackets. Please refer to the notes below Table B.2.27.

B.2.11 Difference-in-differences Results

We estimate a two-period difference-in-differences (DD) model in Section 2.6.3, as an alternative to our RD strategy. While invoking much stronger identifying assumptions (i.e. parallel trends, no time-varying unobservables, as-good-as-random selection into 10th-Plan treatment), this model allows us to include larger villages that are far from our 300-person RD cutoff.[65] Despite estimating treatment effects on a larger sample with a different counterfactuals (comparing 10th- vs. 11th-Plan villages, as opposed to barely eligible 10th-Plan vs. barely ineligible 10th-Plan villages), our DD estimates for nighttime brightness and male agricultural employment are quite comparable to our (preferred) RD estimates (see Figure 2.6.11).

Figure B.2.22 reports the analogous DD results for female agricultural, male other, and female other employment. We see that the DD point estimates are quite close to our RD estimates, except for female other employment.[66] This DD figure also shows relatively constant treatment effects for labor outcomes across the population support, which is consistent with the DD results for male agricultural employment in Figure 2.6.11. By contrast, DD treatment effects for nighttime brightness are positive and increase monotonically in village population.

Together, these results suggest that the small magnitude of our RD labor results is not simply an artifact of restricting our sample to villages smaller than 450 people. Even though the effects of electrification on nighttime brightness are increasing in population, labor effects are constant. This corroborates our RD evidence that electrification does not transform labor markets.

Table B.2.31 compares our preferred RD point estimates to DD estimates estimated using a pooled version of Equation (2.2):

$$Y_{vst} = \gamma_0 + \gamma_1 \mathbf{1}[\text{10th} \times \text{Post}]_{vt} + \delta_t + \eta_v + \varepsilon_{vt} \tag{B.2}$$

Taken at face value, the pooled DD point estimates (in the second column) are fairly consistent with the RD point estimates, except for female other employment. Importantly, for all four labor outcomes, 95 percent confidence intervals reject effects larger than 2.2 percent for both RD and DD models. The third column of Table B.2.31 tests the parallel trends assumption by estimating a district-level model on pre-RGGVY employment:

$$Y_{dt} = \zeta_0 + \zeta_1 \mathbf{1}[\text{10th} \times 2001]_{dt} + \delta_t + \eta_d + \varepsilon_{dt}\,, \quad \text{for } t \in \{1991, 2001\} \tag{B.3}$$

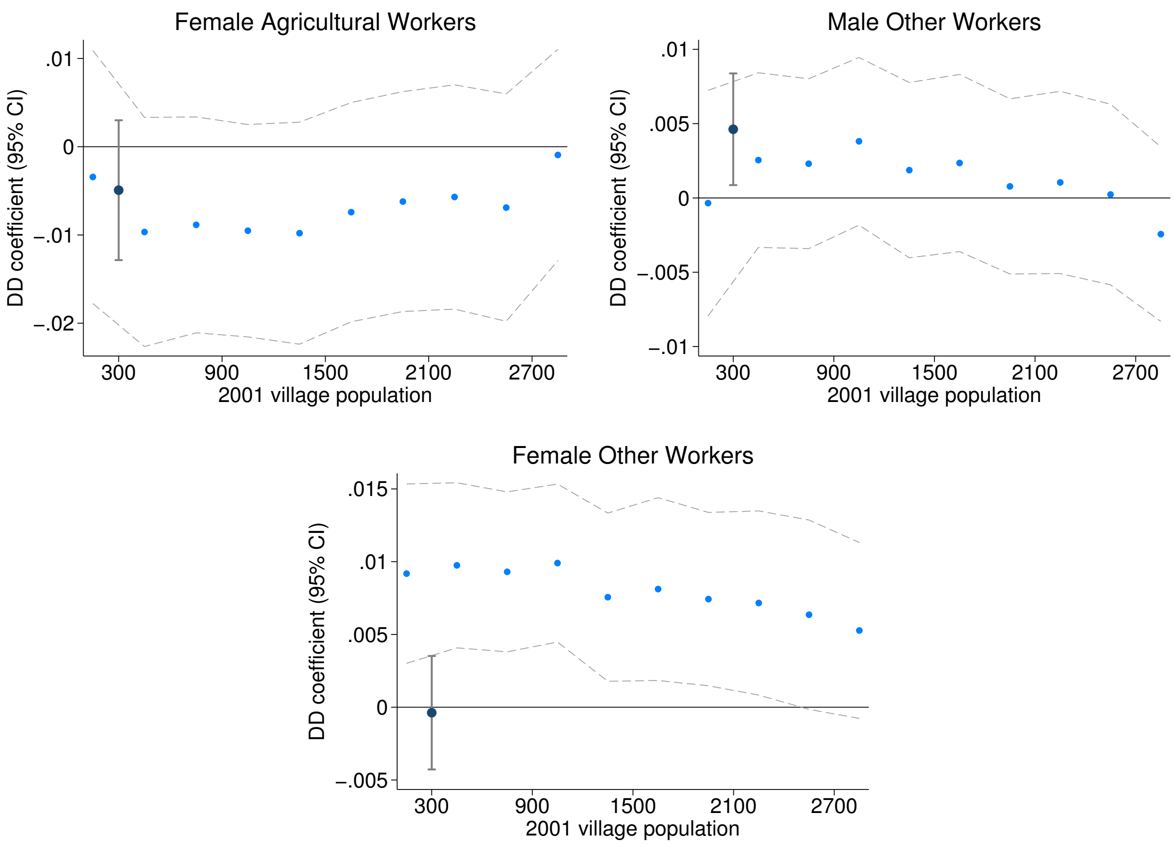
The parallel trends assumption necessary to identify Equations (2.2) and (B.2) requires that $\zeta_1 = 0$. However, Table B.2.31 reveals that we can reject parallel trends for three

[65] Our difference-in-differences analysis also allows us to incorporate multi-habitation villages and villages we could not match to the habitation dataset, since these restrictions are not necessary for identification in the DD context.

[66] Our 300-person DD bins estimate separate treatments for villages with 2001 populations between 0–300 and 300–600. Because our 150–450 RD window spans two bins, the RD estimate lies between these two DD binned estimates. We construct population bins this way in order to allow for heterogeneous DD effects on either side of the 300-person cutoff.

out of four labor outcomes, as $\hat{\zeta}_1$ is statistically different from zero. Hence, we interpret our DD results with caution.

Figure B.2.22: Difference-in-Differences Results



Notes: This figure compares the reduced form effects from our preferred RD specification (Equation (2.1)) to the results from our DD specification (Equation (2.2)), using 300-person population bins. Navy blue dots show the RD coefficients, and whiskers the RD 95 percent confidence interval. Light blue dots and dashed lines show the DD point estimates and 95 percent confidence intervals. From left to right, the three panels show effects for female agricultural, male other, and female other employment. Only the RD results for male employment is statistically significant at the 10 percent level. Table B.2.31 reports pooled DD results. DD regressions include 994,802 village-year observations.

Table B.2.31: RD vs. Difference-in-Differences Results

Outcome Variable	RD Point Estimate	DD Point Estimate	1991–2001 District Pre-Trend
Nighttime brightness	0.1493^{**}	0.4540^{*}	
	(0.0603)	(0.2659)	
	$[0.031, 0.268]$	$[-0.067, 0.975]$	
Number of villages	18,686	314,889	
Number of districts	130	307	
Male ag workers / male pop	-0.0071^{**}	-0.0136^{***}	-0.0112^{**}
	(0.0028)	(0.0042)	(0.0049)
	$[-0.013, -0.002]$	$[-0.022, -0.005]$	$[-0.021, -0.002]$
Female ag workers / female pop	-0.0049	-0.0069	0.0024
	(0.0040)	(0.0059)	(0.0101)
	$[-0.013, 0.003]$	$[-0.019, 0.005]$	$[-0.017, 0.022]$
Male oth workers / male pop	0.0046^{**}	0.0013	-0.0076^{**}
	(0.0019)	(0.0028)	(0.0033)
	$[0.001, 0.008]$	$[-0.004, 0.007]$	$[-0.014, -0.001]$
Female oth workers / female pop	-0.0004	0.0086^{***}	-0.0130^{***}
	(0.0020)	(0.0028)	(0.0039)
	$[-0.004, 0.004]$	$[0.003, 0.014]$	$[-0.021, -0.005]$
Number of villages	29,765	497,401	
Number of districts	225	499	499

Notes: This table compares our main RD regression results (from estimating Equation (2.1)) to results from a difference-in-differences model. The first column reproduces the RD results from Tables 2.5.2 and 2.5.3. The second column reports $\hat{\gamma}_1$ from separate regressions of Equation (B.2) on each outcome. The third column reports $\hat{\zeta}_1$ from separate district-level regressions of Equation (B.3) on each outcome. (weight districts by their 2001 rural populations). The number of villages and districts differs across nighttime brightness and labor regressions, because our nighttime lights sample includes only states with available shapefiles that correlate with village areas. We report 95 percent confidence intervals in brackets. All standard errors (in parentheses) are clustered at the district level. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B.3 Electrification in India: A (More) Detailed History

B.3.1 Before RGGVY

India has a long history of rural electrification programs.[67] Upon Independence in 1947, rural electricity access was virtually nonexistent; by 2012, 92 percent of India's villages were electrified, based on the government's official definition. This dramatic increase is attributable to a series of rural electrification schemes. Each of India's Five Year Plans has funded some sort of rural electrification program, beginning with the 1st Plan (1951–1956) and continuing through the current 12th Plan (2012–2017).

The 1st Plan (1951–1956) focused its electrification efforts on agricultural production and irrigation. A village was legally considered electrified if any electricity was used within its boundaries for any purpose. During these early years, the government's goal was to provide electricity in every 200th village. Ultimately, 4,231 villages were electrified. Under the 2nd Plan (1956–1961), the government's goals shifted towards providing electricity as a "social amenity." This accelerated electrification efforts, bringing 14,458 villages and 350 towns online by the end of the 2nd Plan. The 3rd Plan (1961–1966) motivated electrification as an anti-poverty tactic, with an additional 25,955 villages receiving access to electricity.

With rural electrification becoming increasingly expensive, the All India Rural Credit Review Committee recommended forming a financing agency focused on energy access. At the start of the 4th Plan in 1969, the Rural Electrification Corporation (REC) opened its doors. As a Public Sector Undertaking with a significant degree of fiscal autonomy, the REC funded rural electrification with the joint goals of reducing poverty and promoting productive activity. During that period, India's Green Revolution was increasing the economic returns to electricity in rural areas: electrified pumps enabled the irrigation systems necessary to support new high yield variety grains. The REC had a mandate to promote electrified pumpsets, and it targeted villages with populations of at least 5,000. In 1974, the beginning of the 5th Plan, the Minimum Needs Programme was begun in order to improve standards of living and provide for basic needs. This scheme targeted states with village electrification rates below than the national average and subsidized short distance connections between villages and the existing grid. Broader access to electricity, beyond for agriculture alone, was a key component of this legislation. As a result, over 200,000 villages gained access to electricity between 1969 and 1979.

Between 1980 and 1990, the 6th and 7th Plans funded a variety of schemes to promote access to electricity, including the Integrated Rural Energy Program and Kutir Jyoti Yojana.[68] These programs had strong distributional motivations, and were designed to

[67] The information from this section comes from a combination of Rural Electrification Corporation (2010) and Banerjee et al. (2014).

[68] Kutir Jyoti Yojana provided 100 percent subsidies for single point connections to BPL households.

decrease energy poverty amongst India's poorest households. The 1980s saw 237,371 newly electrified villages.

Under the 8th Plan (1992–1997), the government created both the Ministry of Power and the Ministry of New and Renewable Energy. However, funding challenges forced rural electrification efforts to slow dramatically, with only 11,666 villages electrified in this five year period. The 9th Plan (1997–2002) once again sought to promote electrification as an economic development program. In keeping with this goal, the Ministry of Power released a new definition of electrification in 1997. Villages were now only considered electrified if electricity was being used in the *inhabited* areas of the village. This exemplified the shift away from electrification solely for agriculture's sake. At the same time, new government programs, including Pradhan Mantri Gramodaya Yojana, began providing subsidies for electricity services. New provisions in the Minimum Needs Programme, as well as the launch of the Accelerated Rural Electrification Program, helped to provide individuals and states with the financing necessary to increase rural energy access. During the 9th Plan, 13,317 villages were electrified. By 2001, 86 percent of all villages were deemed electrified.

The 10th Plan (2002–2007) spanned several major changes to India's rural electrification efforts. The Electricity Act, 2003, codified the government's commitment to rural electrification, stating that "The Central Government shall ... formulate a national policy...for rural electrification and for bulk purchase of power and management of local distribution in rural areas." The Act also requires the government to "endeavour to supply electricity to all areas including villages and hamlets" (Ministry of Law and Justice (2003)). In 2004, the Ministry of Power created a new, stricter definition of electrification, which is still in use today. A village is now officially considered electrified only if basic infrastructure, including transformers and distribution lines exist in that village and in its constituent habitations; if public locations such as schools, government offices, health centers, and others have electricity; and if at least 10 percent of the village's households are electrified.

In 2004's National Electricity Policy, the Ministry of Power laid the groundwork for the future of rural electrification in India. Invoking the 2003 Electricity Act, the National Electricity Policy states that "The key development objective of the power sector is supply of electricity to all areas including rural areas...governments would jointly endeavour to achieve this objective at the earliest" (Ministry of Power (2005a)). In particular, subsequent rural electrification programs are supposed to create a "Rural Electrification Distribution Backbone" (REDB) of at a minimum one 33/11 kV or 66/11 kV substation in each Census block, with higher-load regions supplied with additional substations. These substations are to be connected to the state transmission grid. In addition, each village should have supply feeders and at least one distribution transformer, such that every household may be connected on demand to the grid via that transformer. Every household should be connected on demand to the village's transformer.[69] In keeping with

[69] Rural Indian households typically pay for their own electricity connections, unless they are specifically subsidized through an electrification program.

the theme of electrification as a tool for development, the Policy requires that electricity infrastructure be able to support the load from agriculture, textiles and other industries, small and medium enterprises, cold-chain (refrigeration) services, and other public services such as health and education. The Policy also stipulates that priority be given to electrification in "economically backwards" regions. Finally, the Policy makes the REC (now a division of the Ministry of Power) the nodal agency in charge of implementing the country's rural electrification goals.

B.3.2 RGGVY

In 2005, the REC initiated Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY), the Prime Minister's Rural Electrification Plan. RGGVY was the flagship Indian rural electrification program, created with the National Common Minimum Programme goal of universal electricity access in mind.[70] Upon launch, RGGVY enveloped the remaining ongoing electrification schemes, including the Accelerated Electrification of One Lakh Villages and One Crore Households and the Minimum Needs Programme.[71] The RGGVY scheme was detailed in a Ministry of Power Office Memorandum from March of 2005 (Ministry of Power (2005b)), and its rules followed directly from the National Electricity Policy.

Under RGGVY, states were required to "make adequate arrangements for supply of electricity," and to serve rural and urban customers equally. As stipulated by the National Electricity Policy, RGGVY was mandated to create the Rural Electricity Distribution Backbone, to electrify unelectrified habitations and villages, and to provide adequate distribution infrastructure in these newly electrified areas. This infrastructure was supposed to be able to support household load, as well as load from irrigation pumpsets, various industrial activity, cold chains, health care, education, and information technology. A small Decentralized Distributed Generation (DDG) provision was put in place for villages where grid connection would be infeasible or prohibitively expensive. RGGVY was specifically intended to "facilitate overall rural development, employment generation and poverty alleviation," and the policy specifically supported the poor by providing 100 percent subsidies for grid connections for below-poverty line (BPL) households.[72] As part of India's national anti-corruption efforts, details of proposed and completed electrification under RGGVY are available online.[73]

The Rural Electrification Corporation, serving as a nodal agency to the Ministry of Power, has been the main implementing agency for RGGVY, providing 90 percent of the capital needs as direct grants to states and loaning them the remaining 10 percent. The

[70] The program was designed to cover the entire country, but one state (Goa) and all of the union territories have been left out, since they had already achieved 100 percent village electrification by 2005.

[71] These other schemes had, by this point, been discontinued for financial reasons. In the South Asian numbering system, 1 lakh = 100,000 and 1 crore = 10,000,000.

[72] Note that this does not include free power - RGGVY only provides free connections to BPL households. Above-poverty line households pay their own connection charges.

[73] The new program website with these details is `http://www.ddugjy.in/`.

REC was responsible for providing detailed program guidelines, including requirements and standards for materials, equipment, and construction. Within each state that was eligible for RGGVY, the state government power utilities designated implementing agencies. These implementing agencies could be state power distribution companies, state electricity boards, state government power departments, central power sector undertakings (appointed by the state government), or co-operative societies.[74] These implementing agencies prepared Detailed Project Reports (DPRs) for each district under their jurisdiction, by carrying out surveys in every village. A DPR listed each village's electrification status, population, number of households (above/below the poverty line, and with/without electricity), and number of public places (with/without electricity). These reports proposed village-by-village RGGVY implementation plans for eligible villages, which included details on new electricity infrastructure and household connections to be installed (Ministry of Power (2014b)). These DPRs were then submitted to the REC after approval from the state government. After the REC conducted a comprehensive review of each DPR, it passed them on to the Ministry of Power for final approval. Once the Ministry of Power approved a DPR, it sanctioned that district's RGGVY project, and the REC disbursed funds to the implementing agency in charge of the project.[75]

Under the 10th Plan, RGGVY limited program eligibility to villages with constituent habitations of 300 people or above. It justified this population cutoff on the grounds of keeping program costs low. In all, the REC reports that RGGVY electrified 64,091 villages under the 10th Plan. This number included both "unelectrified" villages that did not meet the 2004 definition of electrification, as well as "de-electrified" that had previously been deemed electrified but no longer meet the official definition of electrification.[76]

In 2008, the Ministry of Power ordered the continuation of RGGVY in the 11th plan (2007–2012).[77] This second wave of RGGVY continued to target electrification for all, with the goal electrifying 115,000 un-electrified villages and providing free connections to 23.4 million BPL households. The REC continued as the nodal agency, with the same the 90/10 percent subsidy/loan capital split. Under the 11th Plan, RGGVY provided the same infrastructure as under the 10th Plan. However, states were now required to guarantee a minimum 6–8 hours of power supply for RGGVY villages before the REC would approved a given DPR proposal. The 11th Plan also included guidelines for electrifying villages where grid extensions would be cost-prohibitive with microgrids, under the small Decentralized Distributed Generation (DDG) carve-out.[78]

[74]The choice of implementing agency was left to the states, and depended on the administrative structure and relative capacity amongst different state agencies. Importantly, these were not local agencies: the REC specifically prohibited Gram Panchayats (local governments) from implementing RGGVY projects.

[75]All implementing agencies were required to bring their own teams to villages for RGGVY electrification; no hiring of local labor was permitted.

[76]These "de-electrified" village either had access to electricity and then lost this access due to infrastructure breakdown, or moved out of official electrified status when the definition became more stringent in 2004.

[77]A new Office Memorandum contains the details (Ministry of Power (2008)).

[78]For more details on this DDG carve-out, see: `http://powermin.nic.in/sites/default/files/uploads/Guidelines_for_Village_Electrification_DDG_under_RGGVY_0.pdf`

A three-tier quality monitoring mechanism was also put in place with the 11th plan. First, the state or sub-state level implementing agencies would conduct inspections to ensure that workmanship on RGGVY projects was up to standards. These agencies were mandated to randomly inspect 50 percent of RGGVY villages. Next, the REC was instructed to inspect materials before shipment to RGGVY sites, as well as a random subsample 10 percent of villages. Finally, the Ministry of Power hired National Quality Monitors to inspect 1 percent of villages. Amid complaints from state and local governments, the 300-person constraint was relaxed in the 11th plan. All villages with habitations of 100 people or more were supposed to be covered in RGGVY during this period. Between 2007 and 2012, the REC reports that an additional 46,206 villages were electrified.

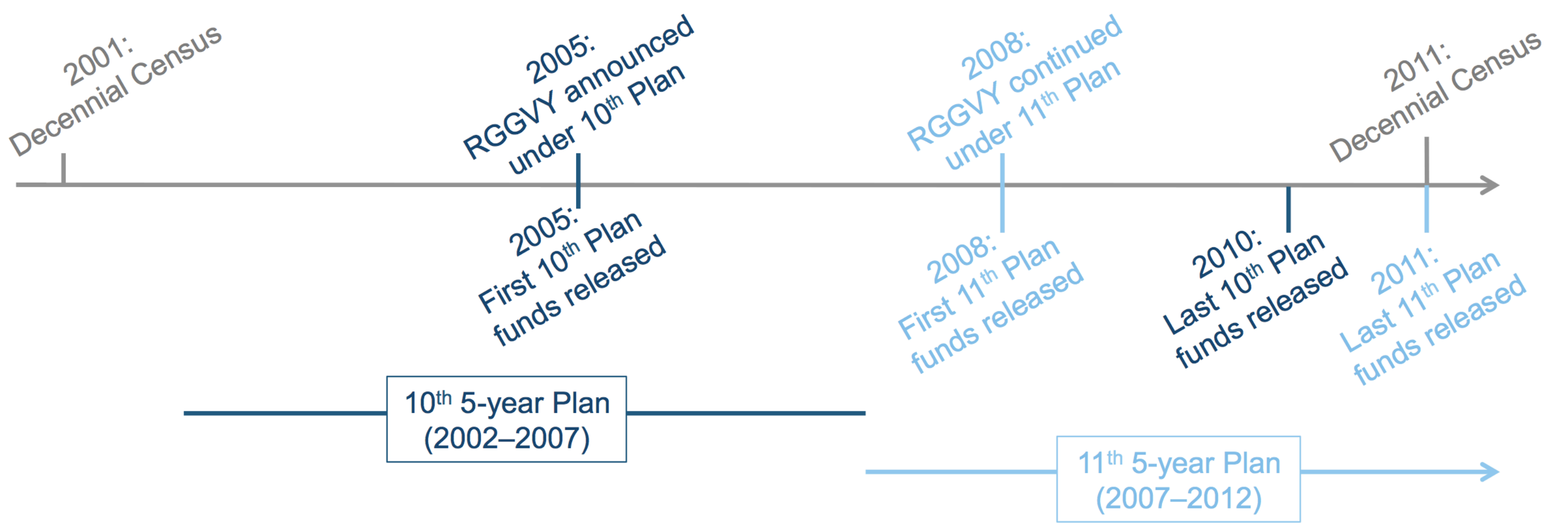
India is currently in its 12th Five-Year Plan (2012-2017). In September 2013, the Ministry of Power again extended RGGVY, this time to be covered under both the 12th and 13th Plans (forthcoming in 2017–2022).[79] This time, the scheme was sanctioned to complete unfinished projects from the 10th and 11th Plans and to cover all remaining habitations and villages with populations above 100. The third wave of RGGVY also continues to subsidize BPL connections. As of 2012, 92 percent of villages in India were officially classified as electrified.

Rural electrification work continues in India. In 2014, RGGVY was subsumed into a new program, Deendayal Upadhyaya Gram Jyoti Yojana (DDUGJY), the Deendayal Upadhyaya Village Light Plan.[80] This program is slated to carry out RGGVY's works as under the 2013 continuation document. Under DDUGJY, however, all villages and habitations are now eligible for rural electrification, regardless of population (Ministry of Power (2015)). DDUGJY also provides for the creation of new feeder lines to separate agricultural and non-agricultural consumers. It aims to strengthen sub-transmission and distribution infrastructure in the rural areas, with a particular focus on metering.

[79] Once again, an Office Memorandum marked the occasion (Ministry of Power (2013)).

[80] The Ministry of Power announced this decision in an Office Memorandum in December (Ministry of Power (2014a)).

Figure B.3.1: RGGVY Implementation Timeline



Notes: This figure shows the timing of the Indian decennial census, the 10th and 11th Five-Year Plans, and RGGVY under these Plans. Our RD analysis uses data from the 2001 and 2011 Census years, described in detail in Section 2.4.