Wiring America: The Short- and Long-Run Effects of Electricity Grid Expansion*

Gaurav Doshi†

This version: December 20, 2021
Click here for the most recent version

Abstract

According to the Net-Zero America study, the US needs to triple its electricity grid in order to decarbonize by 2050 (Larson et al. 2021). This paper examines the impact of large scale grid expansion on price-cost markups, emissions from fossil fuel generators, and wind investment. I focus on the rollout of a grid expansion project that linked windy areas in west Texas to population centers in the east. I find moderate declines in markups and emissions with total annual benefits of roughly $100 million in the short-run. Counties that received investment in transmission infrastructure saw significantly higher wind capacity (+202%) in the long-run, preventing $271 million worth of carbon emissions in 2019. Investments in grid level storage and grid-upgrades can be useful in enhancing the benefits of transmission expansion.

JEL Classifications: L11, Q40, Q41, Q53.
Keywords: Electricity Markets, Emissions, Transmission Expansion, Wind Energy

*I am grateful to my dissertation committee of Sarah Johnston, Sheldon Du, and Dan Phaneuf for their guidance and encouragement. Shreyans Saraogi provided excellent research assistance. This draft has benefited from comments and suggestions from Bill Provencher, Laura Schechter, Corbett Grainger, Andrew Stevens, Paul Dower, Akshaya Jha, Joseph Eggers, and seminar participants at AERE Summer Conference, EMEE Summer Workshop, MEA Annual Meeting, OSWEET, and Environmental and Natural Resources Seminar at UW Madison. I am also greatful to Ted Hailu and ERCOT Client Services for their assistance in accessing and using the data at ERCOT. All errors are my own.

†Department of Ag and Applied Economics, University of Wisconsin Madison, Madison, WI. E-mail: gdoshi2@wisc.edu

1 Introduction

The US aims to achieve a carbon-free power sector by 2035 and economy wide net zero emissions by 2050 (The White House 2021). Most wind and solar power plants in the US produce electricity in geographic locations far from where electricity is consumed. Transmission lines enable moving this electricity over long distances to the demand centers. Thus the transition to a carbon-free power sector and a decarbonized future will require substantial investment in transmission lines. [1]

Inadequate transmission capacity not only impedes integration of electricity from renewable sources but also enhances the market power exerted by fossil fuel generators (Borenstein et al. 2000; Joskow and Tirole 2005). The resulting welfare loss due to market power and the forgone benefits from lower emissions can be in the order of hundreds of millions of dollars annually (Woerman 2019; Fell et al. 2021). I add to the empirical evidence on this issue by analyzing the short-run impact of grid expansion on price-cost markups, local and global emissions from fossil fuel generators. The main innovation of my approach is to study the market and non-market impacts under a common empirical framework such that the potential benefits are comparable.

Another implication of grid expansion is the higher investment in renewable resources in the long-run. However, any analysis to quantify this is plagued by selection issues due to non-random siting of electricity transmission. Exploiting the rich spatial and temporal data from the rollout of a large scale transmission expansion project called Competitive Renewable Energy Zones (CREZ) in Texas, I provide first causal estimates in the economics literature on the magnitude of long term investment in wind energy in response to transmission expansion.

For the short-run analysis, I build a model of optimal bidding to understand how transmission line expansion affects the incentives of a marginal fossil fuel generator to set markups. This model is most closely related to Ryan (2021) who derives the optimal bidding condition for a fossil fuel generator and applies it in the context of the Indian electricity market. I extend it by deriving conditions under which integration of

1. This issue has been covered widely in both energy and popular news outlets, pointing out the imminent necessity to build transmission lines in order to dramatically cut carbon emissions and achieve ambitions energy goals (New York Times 2016; Temple 2017; Meyer 2021). Figure E1 shows the locations of all solar and wind projects ($\geq$10 MW) and the county level population density.

renewable energy would affect realized markups. I write this model in the context of a uniform auction wherein the generator participates by bidding the price and quantity of electricity. I specifically focus on the case of the marginal generator since its optimal bid determines the wholesale price. While the theoretical model is tailored to the empirical context in Texas, the findings are applicable to other markets in the US where transmission expansion is needed to integrate renewable resources to the grid.

The model yields several predictions on how large scale transmission expansion affects the marginal generator's markups. More generally, the ability of a generator to set markups depends on two factors. First, the extent to which wind generation displaces production from fossil fuel generators at the margin. Second, the degree to which integration of wind affects the slope of the electricity dispatch curve at the margin. The relative magnitudes and directions of these two effects determine whether the marginal generator sets higher or lower markups.

I derive a two-step estimator to examine how transmission expansion affects markups by solving the generator's optimization problem. In the first step, I estimate the effect of transmission expansion on hourly wind generation, followed by estimating the effect of wind generation on hourly markups set by marginal generator(s). I use a fixed effects model which flexibly controls for demand and seasonality in wind generation and markups to estimate the empirical analogues of this estimator. Therefore, identification comes from the variation in markups caused by changes in wind generation across similar hours for a given month and year of the sample. I find that CREZ expansion led to moderate decline in markups- about 2.5 percent during the peak demand hours and about 7 percent during the off peak hours. A counterfactual analysis suggests that CREZ led to a $44 million annual reduction in rents collected by marginal generators from retailers in the short-run and eventually consumers in the long-run.

I use the empirical framework above to study the impact of CREZ expansion on hourly emissions across different regions of Texas. I find a decline in emissions in the order of $53 million annually with about 60 percent of the decline due to local pollutants (SO_2 and NOx) and the remaining share from lower carbon emissions. While the value of damages prevented from carbon emissions is similar across different regions in Texas, the decline in local pollutants comes mainly from the west. However, I find an increase in emissions due to ramping up of coal generators at the margin as a result of wind intermittency during the early hours of the day.

Next, I estimate the extent to which investment in transmission spurs investment in wind generation in the long-run. The identification challenge is that locations with superior wind quality were selected to site CREZ lines. I implement Coarsened Exact Matching to address the selection issue. I match the counties on a wide range of pre-treatment observable dimensions that affected both selection into CREZ and investment in wind. These observables include geographic suitability, county demographics, factors affecting project costs, county specific wind regulation, and pre-transmission expansion wind capacity. Therefore, conditional on matching counties on these characteristics, selection into CREZ is as good as random.

Regressions using matched sample suggest that counties that received transmission infrastructure saw 73 MW (+202%) higher wind capacity, 40 more turbines (+245%), and about 33 MW (+121%) bigger wind projects over 2012 to 2019. A back of the envelope calculation shows that the wind capacity added due to CREZ prevented approximately $271 million in damages from carbon emissions in Texas in 2019.

The short-run impacts on markups and emissions highlight the role of grid-scale energy storage and careful siting decisions. These could mitigate the rise in emissions due to wind intermittency during early hours of the day. This would be beneficial in preventing damages from emissions especially in more populated regions. Grid level storage can be instrumental in reducing wind curtailment during periods of high wind and fossil fuel markups at peak hours when wind generation is low.

These results speak to the impact of large scale grid expansion on renewable investment and corresponding benefits from pollution reduction in other parts of the US as well. For instance, grid expansion can address the dropping out of thousands of megawatts of renewable energy projects from development due to inadequate transmission capacity in the Midwest (Thill 2020). Investments in grid upgrades would be beneficial in ensuring that additional wind does not strain the grid during periods of congestion and weather shocks.

Related Literature. This study builds on the insights from several sets of papers. First, it adds to the extensive literature on the incidence and consequences of market power in wholesale electricity markets. Studies focused on electricity markets post deregulation have found market power contributing to high wholesale prices (Borenstein et al. 2002) and misallocation due to sub-optimal bidding behavior (Hortacsu and Puller 2008; Hernández 2018). Existence of market power in sequential electricity markets is found to result in price premium across markets by causing lack of arbitrage (Saravia 2003; Borenstein et al. 2008; Ito and Reguant 2016). Several studies have highlighted the role of financial arbitrage (Borenstein et al. 2008; Birge et al. 2018; Mercadal 2018), vertical structures, and forward contracting in mitigating market power (Bushnell et al. 2008).

Second, I contribute to the growing literature focusing on the value of transmission infrastructure in mitigating market power in wholesale electricity markets. Theoretical studies in this area employ Cournot models and simulations to show how expansion in transmission capacity leads to higher competition and mitigates the effects of market power (Borenstein et al. 2000; Joskow and Tirole 2000, 2005). Recent empirical literature has looked at the welfare effects of geographical integration in electricity markets (Davis and Hausman 2016) and the effects of transmission constraints in exacerbating the market power exercised by generating firms (Woerman 2019; Ryan 2021).

Third, this paper adds to the nascent literature looking at the link between transmission expansion, wind energy, and the wholesale electricity prices. This builds upon the empirical literature in economics exploring the impact of renewable generation in lowering emissions in the power sector (Cullen 2013; Kaffine et al. 2013; Novan 2015; Fell and Kaffine 2018; Fell and Johnson 2021). Fell et al. (2021) study how CREZ expansion enhanced the environmental value of wind measured by emissions avoided. Recent papers find that CREZ led to significant reduction in wholesale market prices (LaRiviere and Lu 2020), congestion risk and the cost of hedging (Doshi and Du 2021).

Outline. The remainder of this paper is organized as follows. Section 2 describes the institutional context along with the CREZ expansion project. I provide a description of the data and some descriptive statistics in Section 3. The theoretical model for the short-run, empirical strategy, and the results are presented in Section 4. The long-run analysis is presented in Section 6, and Section 7 provides a concluding discussion and policy implications of the results.

2 Institutional Details

2.1 The Texas electricity market

The Texas electricity market is one of the major deregulated electricity markets in the US. Electric Reliability Council of Texas (ERCOT) is mandated to maintain system reliability and manage the wholesale and retail electricity markets in Texas. One of the tasks of ERCOT is scheduling supply from generators in order to meet demand for electricity at all times. It does so by organizing a series of sequential auctions and real-time market operations. In this paper, I focus solely on the real time-market decisions by fossil fuel generators.

Even though the ERCOT interconnection spans a single state geographically, it overlooks over 46,500 miles of electricity transmission and 700 generators serving electricity demand from over 26 million consumers. As of 2020, Natural Gas represented about 51 percent of electricity generating capacity followed by 25 percent by wind and 13.4 percent by Coal (ERCOT 2021). In terms of emissions, in 2019 power sector contributed to about 212.4 million metric tonnes of CO_2 emissions in 2017, about 12.3 percent of the total CO_2 emissions from the power sector in the US (EIA 2019). Clearly, Texas is an important context to study the behavior of fossil fuel generators and their environmental impact.

The ERCOT interconnection is comprised of six zones within Texas - Panhandle, West, North, South, Houston, and Coastal.[2] Figure 1 shows the distribution of all the wind projects and fossil fuel generators ($\geq$ 10 MW) in Texas along with the five major demand centers - Houston, Austin, Dallas, Forth Worth, and San Antonio. Most of the wind farms in Texas are located in the wind rich Panhandle and West. North and the South Zones host most of the fossil fuel generating firms along with the major demand centers. Houston forms the fourth zone which is a major demand center in itself. These zones are connected by a network of over 46,500 miles of transmission lines that carried about 74,820 MW of electricity at a record peak demand on August 12, 2019 (ERCOT 2021).[3]

2. Figure E2a provides a sense of geographic distribution of counties in these zones.

3. To put this in perspective, this amount of electricity is equivalent to powering about 15 million Texas homes during periods of peak demand (ERCOT 2021).

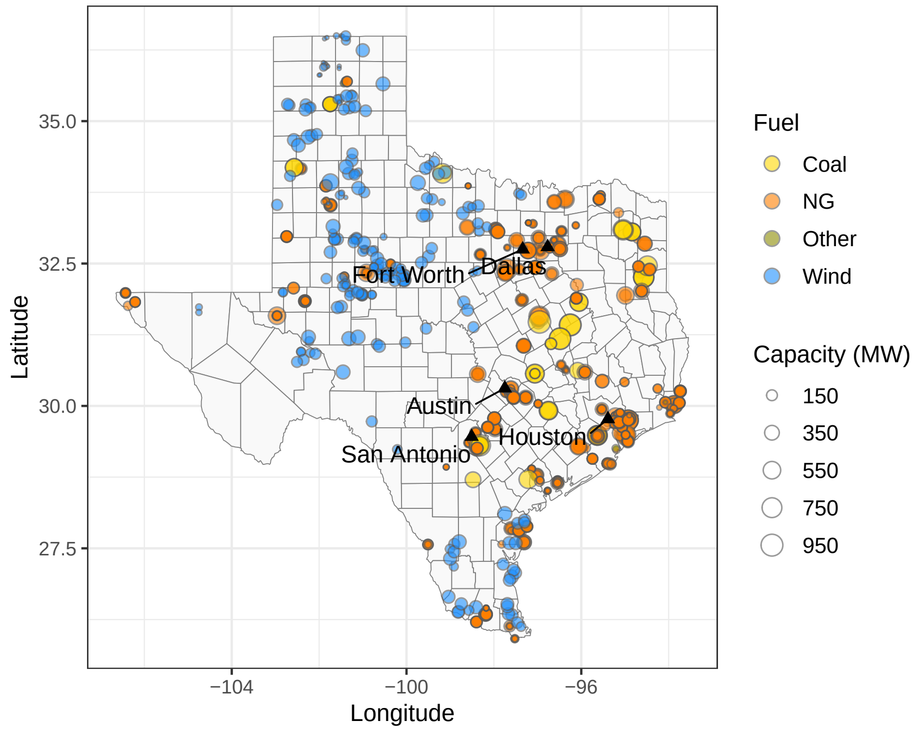


Figure 1: Wind farms and fossil fuel generators ($\geq$ 10 MW) in Texas

Note: 'NG' refers to Natural Gas generators. 'Other' fuel type includes petroleum, blast furnace gas, and other gas based generators. These generators are only 2 percent of total generators in Texas. Black triangles mark the locations of the five biggest population centers in Texas.

2.2 Competitive Renewable Energy Zones

An interesting aspect of the Texas electricity market is the increasing capacity of wind based power generation. Since ERCOT schedules lowest cost generation to dispatch first, wind based generators are always scheduled to dispatch first conditional on wind flow. Fossil fuel generators on the other hand are dispatched to meet the remaining demand as well as to meet any sudden surge in demand at Peak hours.[4]

Inadequate transmission capacity between the west and other parts of Texas could lead to transmission congestion thereby preventing the trade of electricity from wind rich west to demand centers in the east and the south. Presence of transmission con-

4. ERCOT defines Peak hours as hours ending in 07:00 to 22:00 from Monday through Friday. The remaining hours are classified as Off-Peak hours. Wind based generators and low marginal cost fossil fuel generators are usually the base-load units whereas Natural Gas units are typically used to meet peak demand because of their ability to ramp-up at low cost at short notice.

straints would cause ERCOT to schedule electricity from local generating units that are typically fossil fuel fired fired generators. This not only leads to CO_2 emissions that could've been offset by clean wind based electricity but also incentivize local fossil fuel generators to charge a markup over their marginal cost of production.

I examine this phenomenon under the backdrop of a recent transmission expansion project, Competitive Renewable Energy Zones (CREZ) in Texas. CREZ was a large scale transmission expansion project aimed at integrating electricity generation from wind farms located in West to the major demand centers in North, South, and Houston Zones. The project, commissioned in 2008 by the Public Utilities Commission of Texas was aimed to accommodate over 18.5 GW of electric power by building about 3,600 circuit miles of 345 kV electricity transmission lines. However, the transmission lines are open access meaning that the use is not limited to only wind generators (Billo 2017). Transmission lines were built over a period of 2011 through 2013 with a total cost of approximately $6.8 billion. All of the CREZ based transmission lines were brought in service by December 2013 (Lasher 2014).[5]

3 Data and Variables

I assemble multiple datasets with varying temporal resolution. For the short-run analysis of generator markups, I assemble a hourly generator level dataset from 2011 through 2014. For the long-run analysis on wind investment, I construct an annual dataset of wind projects from 2001 through 2019. Most of my data comes from publicly available sources like ERCOT, the Energy Information Administration (EIA), and the Environmental Protection Agency (EPA).

3.1 Markups

One of the main outcomes of interest for the short-run analysis is the generator markups. Markups are defined as $p - c$ where p is the Locational Marginal Price (LMP) and c is the marginal cost. LMP is defined as the price of supplying 1 MWh of electricity at a particular location. I use publicly data available from ERCOT to identify the price setting (marginal) generators and the corresponding LMP at each hour of the sample. The

5. Figure E2b in Appendix shows the location of CREZ transmission lines along with the county level population.

other component of markup is the marginal cost of generation. The generating technology assumes constant marginal cost of generation since fuel costs remains constant. The assumption of constant marginal costs is also common in the literature. The two major components of marginal cost are fuel costs and emissions permit costs based on emissions regulations for SO_2 and NOx. I calculate the marginal cost of each generator as the sum of these two components.

To compute fuel costs, I use weekly price data for coal and natural gas. For coal, I use Powder River Basin spot prices from EIA. For natural gas, I use Henry Hub Natural Gas prices from Quandl. I calculate fuel costs by multiplying fuel price and the heat rate (HR_i) of the generator.[6] I use hourly electricity generation data at the generator level from ERCOT and heat input data from EPA's Continuous Emissions Monitoring system (CEMS).

To compute the emissions permit costs, I use daily data on NOx and SO_2 allowances from S&P Global Market Intelligence. Using hourly emissions data from CEMS, I calculate the emissions rate (ER_i) for SO_2 and NOx by taking the ratio of emissions to net generation.[7] The generator's emission permit cost is thus the product of the permit price and emissions rate for each emission type. Thus, the marginal cost c_{it} of generator i in period t is:

$$c_{it} = \text{HR}_{it} \cdot p_t^{\text{fuel}} + \text{ER}_{it}^{\text{SO}_2} \cdot p_t^{\text{SO}_2} + \text{ER}_{it}^{\text{NO}_x} \cdot p_t^{\text{NO}_x} \tag{1}$$

Figure 2 shows the distribution of marginal costs ($/MWh) of coal and natural gas generators in the sample. The distribution of marginal cost for both the fuels is right skewed with the averages below $25/MWh for both the fuel types. The average marginal cost for coal generators is slightly higher than that of natural gas generators.

Table 1 reports descriptive statistics of key variables by fuel type. Each observation in the sample is a generator-hour combination. About 70 percent of the observations in the sample are natural gas generators and the coal generators are the remaining 30 percent. Even though the marginal cost of coal generators is on average $6.3/MWh higher

6. EIA defines heat rate as the amount of energy used by a power plant to produce 1 KiloWatthour (kWh) of electricity. It is calculated as a ratio of fuel input to net electricity generated and is expressed in British thermal units (Btu) per net kWh.

7. Due to Clean Air Act (CAA) electricity generators are subjected to emissions regulations for SO_2, NOx or both. Generators are required to purchase emission permits for each ton of emissions (SO_2 and NOx) they emit.

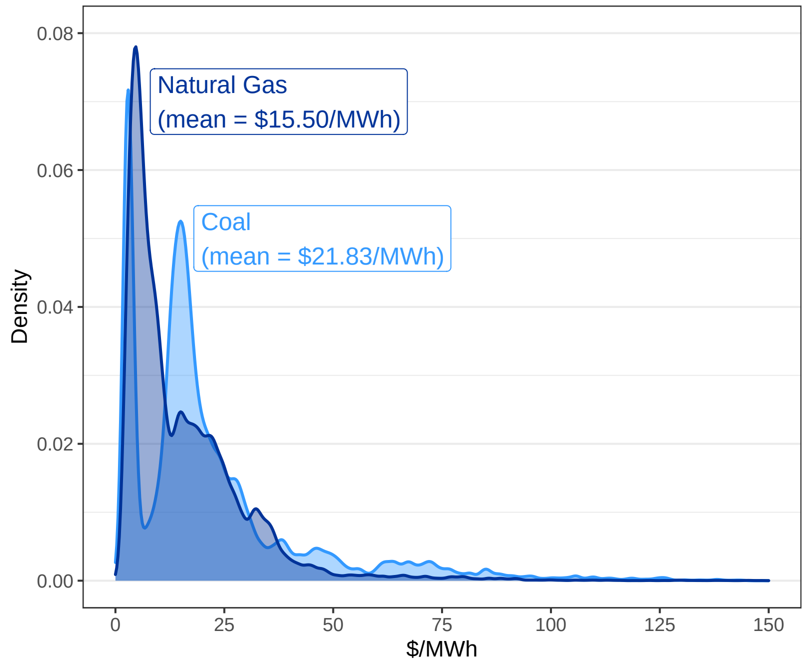


Figure 2: Distribution of marginal costs ($/MWh) of coal and natural gas generators

than the marginal cost of natural gas generators, the average markups set by marginal natural gas generators is about four times that of the coal generators. This is reflective of the pattern of generator dispatch in the Texas electricity market wherein coal generators tend to be at the margin during the night and early hours of the day whereas natural gas generators operate at the margin during the peak demand hours. Thus, marginal natural gas generators have greater incentives to set high markups during peak demand hours.

Generator markups exhibit quite a lot of hourly variation which is not apparent from Table 1. Figure 3b shows the average hourly markups of marginal generators from 2011 to 2014. From Figure 3b, we see that on an average markups were about $50/MWh during the peak hour of 16:00 in 2013 and over $30/MWh in 2011 and 2012. However, markups saw a dramatic drop in 2014 post CREZ expansion across peak hours of 14:00 to 17:00, perhaps most significant at 16:00. However, average markups show substantial hourly variation suggesting evidence of seasonality and generator idiosyncrasies.

Another key observation from Table 1 is that coal generators are much larger in capacity than natural gas generators. The average capacity of a coal generator in the sample is 602 MW wheres the average capacity of a natural gas generator is 190 MW. Further,

Table 1: Descriptive statistics of key variables by generator fuel type

Variable	Fuel	Mean	Std. Dev.	Min	Max
Marginal Cost	Coal	21.83	21.04	0.82	143.78
($/MWh)	Natural Gas	15.50	14.22	0.00	149.85
Realized Markups	Coal	4.18	31.97	−122.42	4597.90
($/MWh)	Natural Gas	16.58	60.40	−138.57	4899.21
Nameplate Capacity	Coal	602.37	200.99	174.60	1008.00
(MW)	Natural Gas	189.93	86.53	25.00	765.00
CO_2 damages/MWh	Coal	79.02	79.71	0.00	1375.50
(2020$)	Natural Gas	24.77	27.90	0.00	4233.44
SO_2 and NOx damages/MWh	Coal	102.40	138.37	0.00	4659.02
(2020$)	Natural Gas	0.76	2.87	0.00	725.97

Notes: This table presents descriptive statistics of key variables by generator fuel type. Sample is all generator-hour observations from August 2011 - December 2014. Total # generator-hour observations (N) is 619,864. Frequency of coal generators is 33.12% and Natural Gas generators are 66.88%. Damages computed using SCC of $44/ton for CO_2 emissions and county specific estimates from Holland et al. (2016) for SO_2 and NOx emissions.

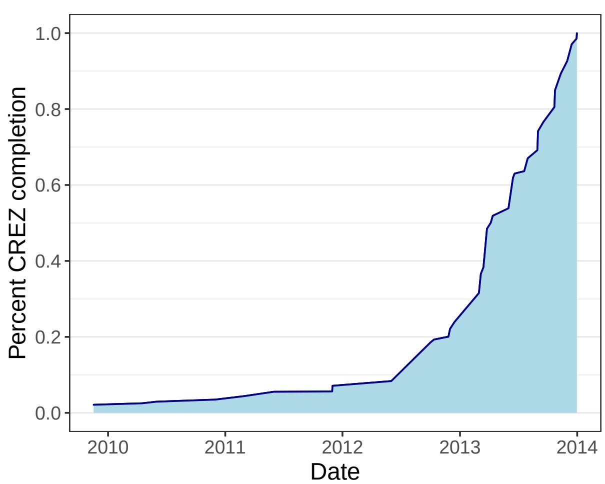
coal generators are much more polluting than natural gas. For the ease of comparison, I present damage estimates (2020$) for CO_2 emissions, and SO_2 and NOx emissions per MWh of power generated. Damages from carbon emissions from coal generators for each MWh of electricity generated is about $79 compared to $25 from natural gas generators. Even more striking is the difference in damages from local pollutants. For each MWh of power generated, damages from NOx and SO_2 from coal generators is on average $101 higher than natural gas generators.

3.2 Global and local emissions

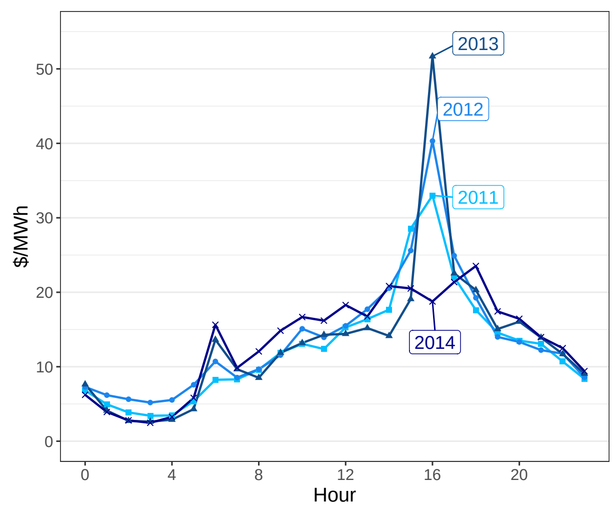
Another outcome of interest for the short-run analysis is the global (CO_2) and local (SO_2 and NOx) emissions. I use data on hourly CO_2, SO_2, and NOx emissions from fossil fuel generators from EPA's CEMS from 2011 to 2014. Since the impact of local pollutants vary across space due to differences in population densities, I use estimates of county-specific marginal damages due to additional ton of SO_2 and NOx from Holland et al. (2016).[8] I combine these county-specific damage estimates to SO_2 and NOx emissions from each generator to compute the $ value of the damages from these emissions.

3.3 CREZ Transmission Expansion

A key explanatory variable is the progress of CREZ transmission expansion. I use the publicly available Transmission Project and Information Tracking reports from ERCOT's website to construct a variable that tracks total miles of transmission lines built in a day under the CREZ expansion project. I express the CREZ progress variable as a cumulative ratio of total progress for the ease of interpretation. As shown in Figure 3a even though the CREZ started in 2010, over 80 percent of the project was completed in 2013.



(a) Daily progress of CREZ expansion.



(b) Average hourly generator markups ($/MWh) from 2011 - 2014.

Figure 3: Daily CREZ progress and generator markups over the years

Note: Figure 3a shows the cumulative share of CREZ lines (miles) completed each day from 2010 to 2014. Figure 3b shows the average hourly price-cost markups for the sample (N = 619,864).

8. The county specific damage estimates reported in Holland et al. (2016) uses the AP2 air pollution model to capture the geographic variation in the environmental costs imposed by local pollutants.

4 Short-run: Impact of CREZ Expansion on Markups

4.1 Transmission Constraints and Market Power

For this analysis, I focus on the real-time electricity market which sets the expectation for prices in the day-ahead and forward markets (Potomac Economics 2019). The main purpose of a real-time market is to match supply with demand while operating the transmission system within established limits. Real-time operations involve participation from various market participants like generators, retailers, transmission service providers, and distributors. The Electric Reliability Council of Texas (ERCOT) serves as the regulatory body that manages the efficient operation of the real-time market including scheduling the dispatch of generators to meet the demand at all times.

Transmission constraints play a central role in determining the cost effective dispatch of generating resources. To understand how, it is important to recognize the role played by transmission infrastructure in electricity markets. Presence of transmission essentially enables flow of electricity between two points. Typically, generating units are located at regions far away from the demand centers. Therefore, a transmission network that is able to carry electricity from supply to demand at all times is of prime importance.

One of the main features of transmission lines is that they operate under certain capacity limits that need to be maintained. Transmission constraints between two points A and B are said to be binding when transmission lines between them operate at their maximum capacity. This is another way of saying that the transmission lines are congested. There could be various reasons for transmission congestion or binding transmission constraints, like increase in demand due to weather conditions, outages, insufficient transmission infrastructure to name a few.

How does presence of transmission constraints translate to generating firms exercising market power? Generators submit monotonically increasing offer curves which is a function of price and quantity of electricity they are willing to supply. Generators anticipate demand and transmission constraints and hence submit a bid that is composed of the marginal cost of supplying electricity and a markup term.[9] As I show in the

9. In ERCOT, generators have access to demand forecasts and the information on transmission infrastructure. They use this publicly available information and any private information about the market to determine their offer curves.

theoretical model below, this markup term is dependent on the shape of the residual demand curve that the generator faces in the market which in turn is a function of the transmission capacity.

4.2 A Model of Optimal Fossil Fuel Markups

The theoretical model in this section aims to understand the effect of transmission expansion on the pricing decision of a profit-maximizing fossil fuel generator. I borrow elements of the merchant transmission investment model by Joskow and Tirole (2005), but extend it by including electricity generation from renewable sources. My model is based on Ryan (2021), however, I differ from it in two key ways. First, I introduce a wind generating sector which is isolated from the demand centers. Second, in my model transmission expansion affects fossil fuel generators mainly by integrating electricity from the wind generating sector. This mimics my empirical setting wherein CREZ expansion impacted fossil fuel generators by integrating electricity generated by wind farms from the West. In what follows, I present the optimal markup rule for a fossil fuel generator and provide intuition on how it is affected by the transmission expansion.

4.2.1 Model Setup

Consider two geographically distinct regions $\mathcal{W}$ and $\mathcal{S}$. Region $\mathcal{W}$ is a wind rich region comprising of wind farms and region $\mathcal{S}$ is comprised of several fossil fuel generators that serve a large demand center. The presence of electricity transmission capacity (K) enables trade of electricity generated from wind in region $\mathcal{W}$ to demand centers in region $\mathcal{S}$. [10]

In this model, I focus on the pricing decision of a profit maximizing fossil fuel generator i located in region $\mathcal{S}$. Generator i submits an offer curve that is a vector of supply quantities Q_i at bid prices b_i while incurring cost $C_i(Q_i)$. The optimization problem of i entails finding the offer curve that maximizes its profit function $\pi_i(p) = p \cdot Q_i(p) - C_i(Q_i(p))$, where p is the market clearing price that resolves in $\mathcal{S}$.

10. Figure E3 in Appendix illustrates this setup graphically.

However, the generator faces uncertainty over the offer schedules $\mathcal{S}_{-i} = (b_{-i}, Q_{-i})$ from other fossil fuel generators $(-i)$ in $\mathcal{S}$. Further, the generator has to consider any forward positions it has. I denote the forward price and quantity of generator as p^F and Q_i^F respectively. Therefore, the optimization problem is,

$$\max_{b_i, Q_i} \mathbb{E}_{\mathcal{S}_{-i}} \left[p \cdot Q_i(p) - C_i(Q_i(p)) + (p^F - p) Q_i^F \right] \tag{2}$$

The last term in Equation 2 is the payoff from the forward position that is resolved in the real time market. Market demand in $\mathcal{S}$ is denoted by D^S and is assumed to be perfectly inelastic.

Generator i faces a downward sloping residual demand curve $D_i^r(p, q_w; \mathrm{K})$ which is comprised of three elements: Market demand D^S; electricity generated from wind imported from $\mathcal{W}$, q_w; and the total electricity generated from competitor fossil fuel generators, $Q_f(q_w, p) = \sum_{j \neq i, j \in \mathcal{S}} Q_j(q_w, p)$. I express Q_f as a function of q_w because the dispatch of a fossil fuel generator depends on the amount of electricity generated by wind.[11] Recall that wind-based electricity generation incurs zero marginal cost and is always scheduled to dispatch first. $Q_f(q_w, p)$ is strictly increasing in p and strictly decreasing in q_w.[12] Mathematically, D_i^r can be written as,

$$D_i^r(p, q_w; \mathrm{K}) = D^S - q_w - Q_f(q_w, p) \tag{3}$$

The market clears when electricity generated by i equals residual demand, i.e. $Q_i(p) = D_i^r(p, q_w; \mathrm{K})$. The market clearing price p and the supply $Q_i(p, q_w)$ depend on the optimal bid price b_i that solves the generator i's problem,

$$\max_{b_i} \mathbb{E}_{\mathcal{S}_{-i}} \left[p(Q_i(p) - Q_i^F) + p^F Q_i^F - C_i(D_i^r(p, K)) \right]$$

11. I assume $D^S > q_w$, otherwise there wouldn't be any need to schedule electricity from fossil fuel generators as all of the market demand could be met by wind.

12. The interpretation of these assumptions is as follows:

1. $\frac{\partial Q_f}{\partial p} = \sum_{j \neq i, j \in \mathcal{S}} \frac{\partial Q_j}{\partial p} > 0$: generators have greater incentives to supply electricity at higher prices.
2. $\frac{\partial Q_f}{\partial q_w} = \sum_{j \neq i, j \in \mathcal{S}} \frac{\partial Q_j}{\partial q_w} < 0$: electricity generated from wind displaces a non-zero amount of electricity from fossil fuel generators.

Denote $Q_i(p, q_w) - Q_i^F$ as $Q_i^{net}(p, q_w)$. Taking first order condition with respect to b_i and rearranging,

$$\implies \mathbb{E}_{\mathcal{S}_{-i}} \left[\frac{\partial p}{\partial b_i} \left(Q_i^{net}(p, q_w) + \frac{\partial D_i^r(p, q_w)}{\partial p} \left[p - C_i'(D_i^r(p, q_w)) \right] \right) \right] \bigg|_{p=b_i} = 0 \tag{4}$$

Equation (4) is the optimal pricing rule for generator i wherein it sets price equal to marginal cost plus a markup term. $\frac{\partial p}{\partial b_i}$ is the slope of market clearing price in the bid price and is equal to one if the bid is marginal and zero otherwise. In this paper, I focus on the case when b_i is the marginal bid and therefore determines the market clearing price. Thus, I refer to i as the marginal generator as its optimal bid sets the price. For simplicity, I assume constant marginal cost i.e. $C_i'(D_i^r(p, K)) = c_i$ and full information on other generators' strategy. Equation (4) reduces to,

$$p - c_i = -\frac{Q_i^{net}(p, q_w)}{\partial D_i^r(p, q_w)/\partial p} \tag{5}$$

Equation 5 shows that the markups are dependent on the net-production of electricity and the slope of its residual demand curve which is a negative quantity. The numerator measures the extent to which generator's production decision affects the markups. With $Q_i^{net} > 0$, the generator is a net seller implying that it withholds output in the forward market to raise the market clearing price in the real-time market such that $p - c_i > 0$. Similarly, with $Q_i^{net} < 0$, the generator is a net buyer and pays price less than the marginal cost for the electricity generated.

The denominator which is the slope of residual demand curve determines the ability of the generator to set markups. A flatter residual demand curve implies that the generator has a lower potential to set markups whereas a steeper residual demand curve implies greater potential to set markups.

4.2.2 Predictions from the model

In order to characterize the effect of transmission line expansion (K) on markups, I perform comparative statics exercise by partially differentiating Equation (5) with respect to K,

$$\frac{\partial(p - c_i)}{\partial K} = \frac{\left[-\frac{\partial Q_i^{net}(p, q_w)}{\partial K} \cdot \frac{\partial D_i^r(p, q_w)}{\partial p} \right] + \left[Q_i^{net}(p, q_w) \cdot \frac{\partial^2 D_i^r(p, q_w)}{\partial p \partial K} \right]}{\left[\frac{\partial D_i^r(p, q_w)}{\partial p} \right]^2} \tag{6}$$

I express Equation (6) as a percentage change in markups by multiplying both sides by the inverse of Equation (5). Algebraic simplification allows me to split the resulting expression into two terms that measure the effect of transmission line expansion on markups. I call these terms Δ*Production* and Δ*Elasticity* based on how they affect net production and elasticity of the residual demand curve.

$$\frac{1}{p-c_i} \cdot \frac{\partial(p-c_i)}{\partial \mathrm{K}} = \underbrace{\left[\frac{1}{Q_i^{net}(p,q_w)} \cdot \frac{\partial Q_i^{net}(p,q_w)}{\partial \mathrm{K}}\right]}_{\Delta \text{Production}} - \underbrace{\left[\frac{1}{\partial D_i^r/\partial p} \cdot \frac{\partial^2 D_i^r(p,q_w)}{\partial p \partial \mathrm{K}}\right]}_{\Delta \text{Elasticity}} \tag{7}$$

ΔProduction. This term measures the percentage change in net-production by generator i in the real-time market due to change in transmission capacity K.

$$\frac{\partial Q^{net}(p,q_w)}{\partial \mathrm{K}} = \frac{\partial Q^{net}(p,q_w)}{\partial q_w} \cdot \frac{\partial q_w}{\partial \mathrm{K}} \tag{8}$$

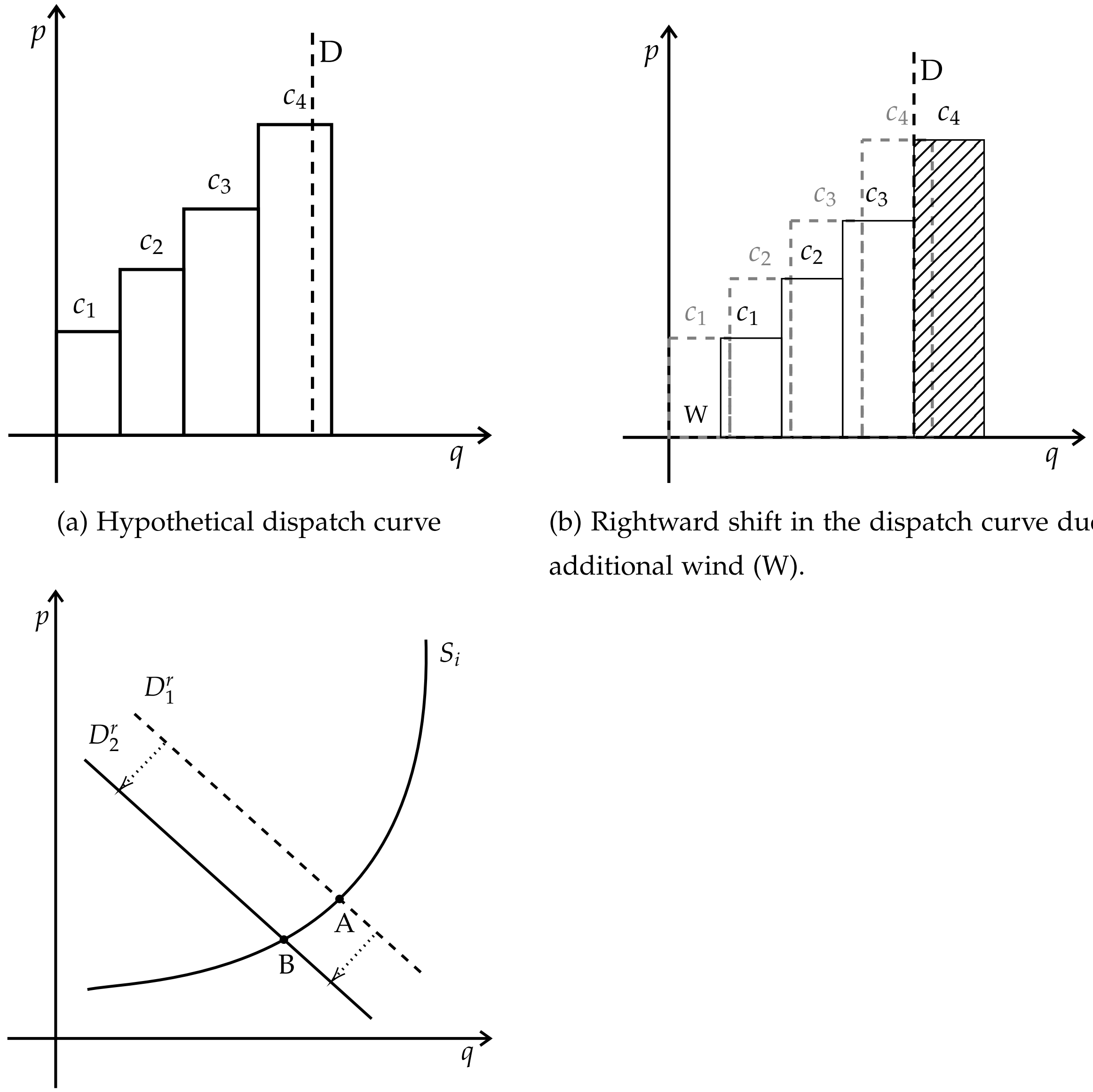
Proposition 1 *The addition of wind leads to an inward shift of marginal generator's residual demand curve.*

The first term in Equation (8) measures the extent to which production decision of generator i is affected by addition of wind energy. Consider a hypothetical electricity dispatch curve shown in Figure 4a. The supply side assumes four fossil fuel generators indexed by their offer/bid price $c_j (j = 4)$ of supplying electricity. The dispatch curve is a step function comprised of generators arranged in an increasing order of the offer price. The dotted vertical line (D) is the demand of electricity and is assumed to be fixed. Generators are dispatched in the increasing order of the offer price until the demand is met. The generator(s) with the highest offer price that is dispatched is the marginal generator and it determines the wholesale price of electricity.[13] In the scenario below, generator i submits the highest offer price c_4 and is thus the marginal generator.

Consider the scenario in Figure 4b wherein additional wind (W) displaces electricity generated from i shown as the hashed area. Mathematically, this can be written as:

$$\frac{\partial Q^{net}}{\partial q_w} < 0 \tag{9}$$

13. If there are multiple generators that have the highest offer price and are dispatched, all of them are referred to as marginal generators.



(a) Hypothetical dispatch curve

(b) Rightward shift in the dispatch curve due to additional wind (W).

(c) Shift in generator i's residual demand curve (D_i^r) due to the additional wind (W)

Figure 4: Hypothetical electricity dispatch curves and the effect of wind generation on marginal fossil fuel generator.

Note: c_i denotes generator i's offer/bid price of supplying electricity. Vertical dotted line denotes the inelastic demand for electricity (D). W is the wind integrated to the grid due to transmission expansion, and S_i denotes the supply curve of generator i.

Thus, electricity from wind shifts the dispatch curve to the right, displacing power generated by the marginal generator i. This is reflected as an inward shift of i's residual demand curve which in turn reduces i's potential to set higher markups. This is shown in Figure 4c with the generator moving from point A to point B of its offer curve post wind integration. Compared to point A, point B is associated with a flatter region of the offer curve thereby reducing i's ability to set higher markups.

Proposition 2 *Transmission expansion leads to integration of electricity from wind into the grid.*

The second term in Equation (8) measures the extent of integration of power generated by wind farms in $\mathcal{W}$ due to increase in the transmission capacity between regions $\mathcal{S}$ and $\mathcal{W}$. Assuming generating capacity of wind to remain fixed in the short-run, transmission expansion would enable higher imports of wind generation into $\mathcal{S}$, therefore $\partial q_w / \partial \mathrm{K} \geq 0$.

Δ**Elasticity.** This term measures the impact of transmission capacity (K) on the slope of marginal generator i's residual demand curve. To understand the direction of this term, I take the derivative of the slope of i's residual demand curve with respect to K. Since the demand for electricity (D^S) and wind generation (q_w) are invariant to changes in p, the slope of i's residual demand curve depends only on the production decisions of its competitors. Therefore,

$$\frac{\partial^2 D_i^r(p, q_w; \mathrm{K})}{\partial p \partial \mathrm{K}} = -\frac{\partial^2 Q_f(q_w, p)}{\partial p \partial q_w} \cdot \frac{\partial q_w}{\partial \mathrm{K}}$$

Let $\eta_f = \frac{\partial Q_f(q_w,p)}{\partial p}$ (> 0) denote the slope of the aggregate (marginal) fossil fuel generators supply curve. Rearranging,

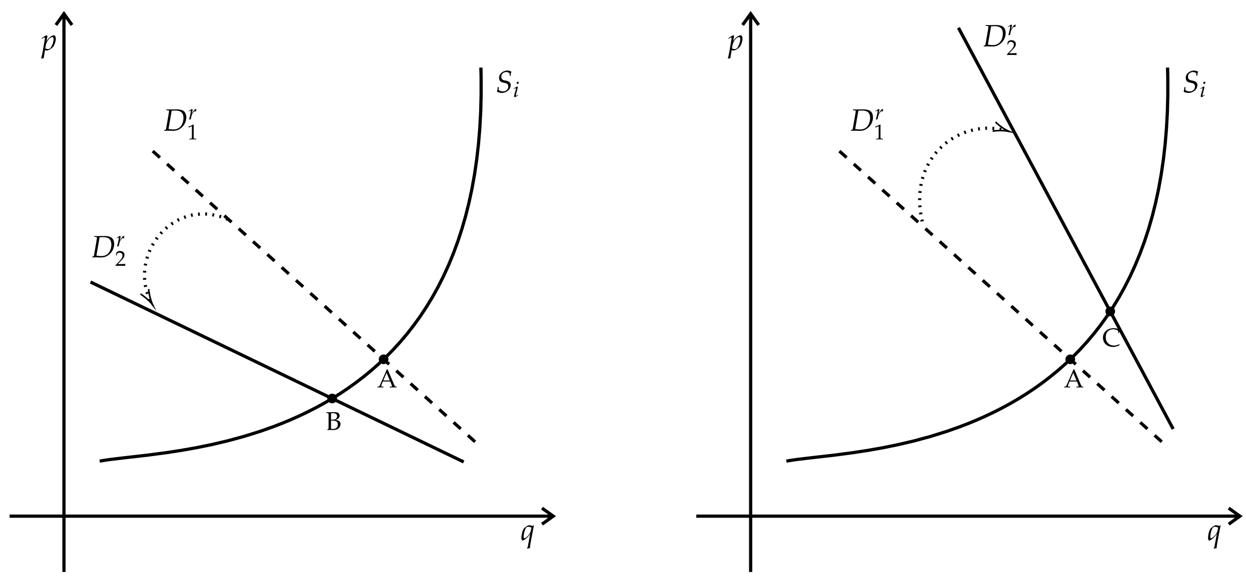
$$\frac{\partial^2 D_i^r(p, \mathrm{K})}{\partial p \partial \mathrm{K}} = -\frac{\partial \eta_f}{\partial q_w} \cdot \frac{\partial q_w}{\partial \mathrm{K}} \tag{10}$$

The slope of the generator supply curves at the margin determines the slope of the dispatch curve at the margin. Therefore, (10) shows that changes in the slope of the dispatch curve due to additional wind will affect generator i's residual demand curve. This leads to the following proposition:

Proposition 3 *The impact of transmission expansion on the elasticity of a marginal generator's residual demand curve is ambiguous.*

As shown in Figure 4b, integration of wind shifts the dispatch curve to the right. Heterogeneity in the cost of electricity generation from fossil fuel sources could lead to generator(s) with steeper or flatter supply curves operating at the margin. Figure 5 shows that this would lead to the rotation of residual demand curve of generator i.

A flatter dispatch curve due to additional wind would result in a flatter residual demand curve as shown in Figure 5a. In this case, generator i moves from point A to point B which is at the more elastic region of its offer curve, thus reducing its ability to set higher markups. However, we could expect a steeper dispatch curve especially during the periods of high demand. Thus in turn would result in a clockwise rotation of generator i's residual demand curve as shown in Figure 5b. In this case, the generator moves from point A to point C of its offer curve. Since point C is associated with the steeper region of the offer curve than point A, this enhances its ability to set higher markups.



(a) Anti-clockwise rotation of residual demand curve as a result of flatter dispatch curve at the margin.

(b) Clockwise rotation of residual demand curve as a result of steeper dispatch curve at the margin.

Figure 5: Rotation of fossil fuel generator's residual demand curve post integration of wind due to transmission expansion.

Note: D_1^r and D_2^r denote the residual demand curves of generator i pre and post transmission expansion respectively, and S_i denotes the supply curve of generator i.

Therefore, $\frac{\partial \eta_f}{\partial q_w}$ could be weakly positive or negative. This translates to Equation (10) being weakly negative or positive meaning a more inelastic or elastic residual demand curve respectively. Substituting the expressions for ΔProduction from Equation (8) and ΔElasticity from Equation (10) in Equation (7):

$$\frac{1}{p - c_i} \cdot \frac{\partial (p - c_i)}{\partial \mathrm{K}} = \left[\frac{1}{Q_i^{net}} \cdot \frac{\partial Q_i^{net}}{\partial q_w} + \frac{1}{\partial D_i^r / \partial p} \cdot \frac{\partial \eta_f}{\partial q_w} \right] \cdot \frac{\partial q_w}{\partial \mathrm{K}} \tag{11}$$

Equation (11) shows that the overall effect of transmission expansion on generator i's markups can be broken down into two pieces. First is the effect of wind generation on markups represented by the two terms in the square brackets in Equation (11). Second is the effect of expansion in transmission capacity on wind generation. Equation (11) can alternatively be expressed as:

$$\frac{\partial(p-c_i)}{\partial K} = \underbrace{\frac{\partial(p-c_i)}{\partial q_w}}_{\gtreqless 0} \cdot \underbrace{\frac{\partial q_w}{\partial K}}_{>0} \tag{12}$$

Thus, the effect of transmission expansion on markups is driven by the effect of wind generation on markups and the extent to which transmission expansion integrates the electricity generated from wind. From Equation (12), $\frac{\partial q_w}{\partial K}$ simply acts as a multiplier that scales up the effect of wind generation on fossil fuel markups.

4.3 Empirical Strategy

In this section, I estimate the impact of transmission expansion (K) on fossil fuel generator markups based on the relationship between transmission lines and markups described in Equation 12. I run the following regressions to estimate the empirical analogues of Equation 12:

$$y_{it} = \alpha_h \cdot w_t + f(D_t|\lambda) + \kappa_i + \delta_{hmy} + \epsilon_{it} \tag{13}$$

$$w_t = \beta_h \cdot crez_d + \gamma \cdot max_t + \eta_{hm} + \omega_t \tag{14}$$

where, y_{it} is the markup set by marginal generator i at hour t of the sample. Markup is defined as $(p-c)_{it}$ where p is the Locational Marginal Price (LMP) and c is the marginal cost of generator i at period t. Wind generation (GWh) in hour t is denoted by w_t, and $crez_d$ is the percentage completion of CREZ transmission project at day d of the sample. The parameters of interest are α_h which measures the impact of wind generation on markups, and β_h which measures the impact of transmission expansion on wind generation. Thus,

$$\alpha_h \approx \frac{\partial(p-c_i)}{\partial q_w}, \quad \beta_h \approx \frac{\partial q_w}{\partial K} \tag{15}$$

I use a wide variety of controls to account for potential confounding factors in Equation 13 and Equation 14. I use a quadratic polynomial of system wide electricity demand D_t in Equation 13 to account for variation in markups driven by spikes in electricity demand.[14] In Equation 14, I use the maximum predicted generation (max_t) of electricity from wind at hour t to control for the maximum energy production possible from wind at period t. [15] This is a useful control because it not only incorporates the generating capacity and technology of the wind generator but it also takes into account the real time meteorological conditions that could affect the amount of power generated through wind farms.

As shown in Figure 6 the actual electricity generated from wind (w_t) closely tracks the maximum predicted wind generation (max_t) for each hour from 2011 to 2014. The difference between the two curves arises due to inadequate transmission capacity needed to transport the power to demand centers. Therefore, this gap is the amount of wind generation curtailed by ERCOT so as to maintain grid stability. However, with the CREZ expansion in 2013 we see the gap between the maximum and actual wind generation decreasing with the lowest difference observed across all hours of 2014.

I use a battery of fixed effects to control for unobserved determinants of markups that could be correlated with wind generation. In Equation 13, I use generator fixed effects (κ_i) to control for any generator specific heterogeneity in markups. I use hour by month by year fixed effects (δ_{hmy}) to control for seasonality exhibited by the electricity market in Texas. This seasonality arises due to varying wind pattern at different hours of the day over the months in a year. For example, wind generation in Texas tends to be higher during the night than during the day. Similarly, wind flow is typically higher during the spring months than winter and summer months. Similarly, I use hour-by-month fixed effects (η_{hm}) in Equation 14 to control for seasonality in wind generation. ϵ_{it} and ω_t are the random error terms in Equation 13 and Equation 14 respectively.

The identifying variation for α_h in Equation 13 comes from the within generator variation in markups caused by changes in wind generation across hours h within a month m in a given year y. For example, α_{16} is the identified from deviations in markups from generator specific averages across all 16:00 hours within a month, in a given year.

14. Using zonal demand levels instead of system wide demand does not change the results.

15. Maximum predicted generation is technically referred to as the High System Limit (HSL) by ERCOT. HSL for a generation resource is defined as the maximum sustained energy production capability of that entity. HSL is established by the generator itself and is continuously updated in Real-Time.

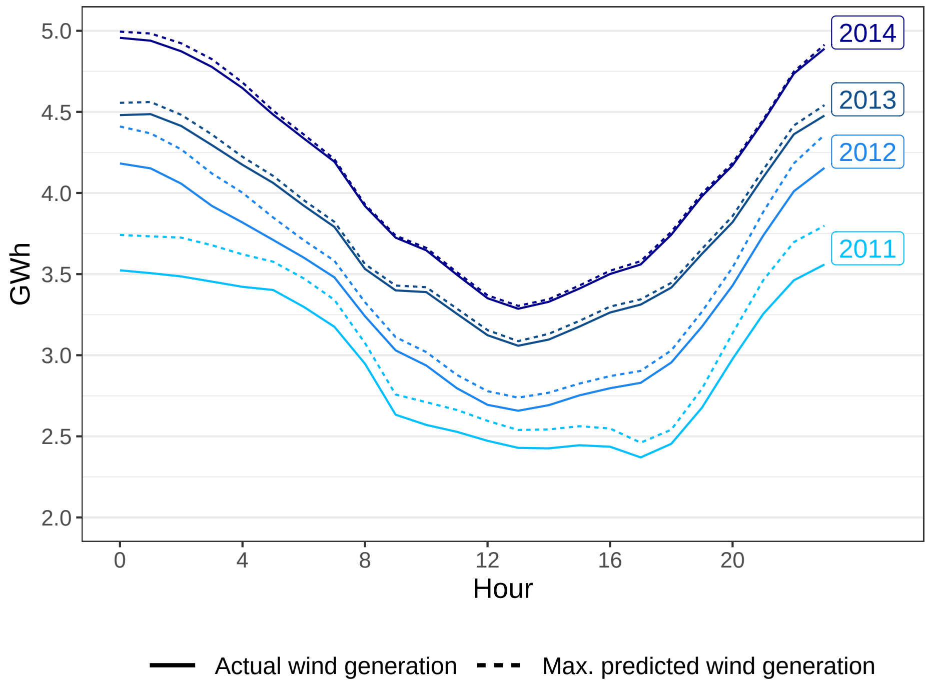


Figure 6: Hourly averages of actual wind generation (w_t) and maximum predicted wind generation (max_t) of wind generation from 2011 - 2014.

Note: max_t is the maximum energy production capability of the generator at period t. It is established by the generator itself and is continuously updated in Real-Time.

Similarly, the identifying variation for β_h in Equation 14 comes from variation in wind generation caused by daily transmission expansion across same hours in a given month.

Under the identifying assumption that the control variables and fixed effects account for confounding factors, α_h captures the unbiased effect of wind generation on generator markups and β_h is the unbiased effect of CREZ expansion on wind generation. Therefore, the estimator of the impact of CREZ on markups is, $\hat{\theta}_h = \hat{\alpha}_h \times \hat{\beta}_h$. Standard errors in Equation 13 are clustered at the generator level. I use Newey West auto-correlation corrected standard errors with a seven day lag structure for estimates in Equation 14.

4.4 Results

I first discuss the results of the effect of wind generation on generator markups in Figure 7, followed by the impact of CREZ expansion on integrating wind generation in Figure 8. Figure 7 shows the coefficient estimates $(\hat{\alpha}_h)$ of the magnitude of the decrease in fossil fuel markups ($/MWh) due to additional GWh of wind energy to the grid. We see that on average the drop in markups is strongest in magnitude at the peak hour of 16:00, about \$9/MWh. The coefficient estimates are smallest for the off-peak hours. Due to low electricity demand and high wind generation during the off-peak hours, fossil fuel generators typically operate on a smaller net demand curve as compared to the peak hours, thereby lowering their incentives to set high markups. In other words, the impact of additional wind in lowering fossil fuel markups is higher at the peak hours than at the off-peak hours.

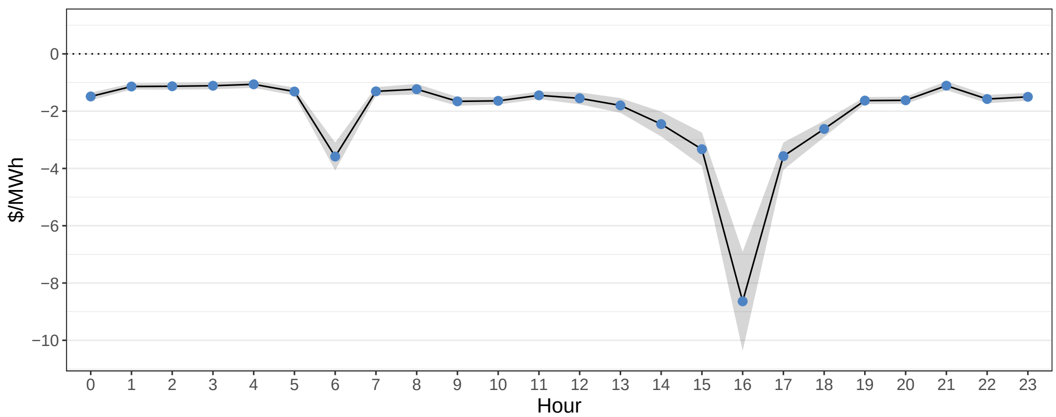


Figure 7: Coefficient estimates ($\hat{\alpha}_h$) of the effect of addition of a GWh of wind energy on generator markups (\$/MWh) and the 95 percent confidence intervals.

Figure 8 shows the impact of CREZ expansion on wind generation at each hour of the day. The effect is highest for the off-peak hours and lowest at the peak hours. [16] The coefficient estimates imply that keeping the stock of generating capacity fixed, CREZ integrated about 0.22 GWh of wind at 23:00 and 0:00, and about 0.10 GWh during the peak hours of 15:00 to 18:00. The hourly pattern of the coefficient estimates ($\hat{\beta}_h$) closely follows the hourly wind flow pattern in Texas where the wind flow is strongest in the evening compared to the day. This reflects that availability of transmission capacity is instrumental in integrating higher levels of wind generation in the Texas electricity market.

Recall that the overall impact of CREZ expansion on markups (θ) is given by the product of the effect of wind generation on markups (α_h) and the impact of CREZ on integrating wind generation (β_h). Figure 9a shows that the drop in markups due to CREZ is strongest at 16:00, about \$0.88/MWh. This is followed by 6:00 (about \$0.53/MWh)

16. The peak hours in ERCOT are defined as the hours ending in 7:00 to 22:00 CPT from Monday through Friday. I use this definition to discuss the results of my analysis as well.

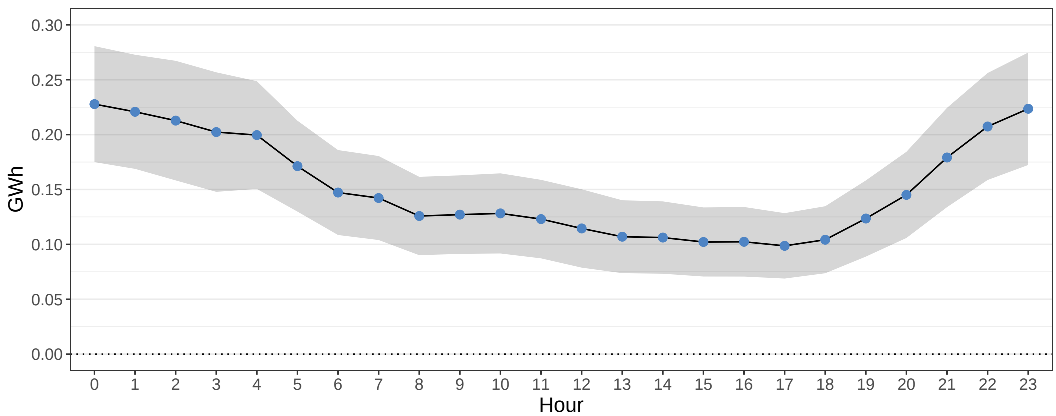


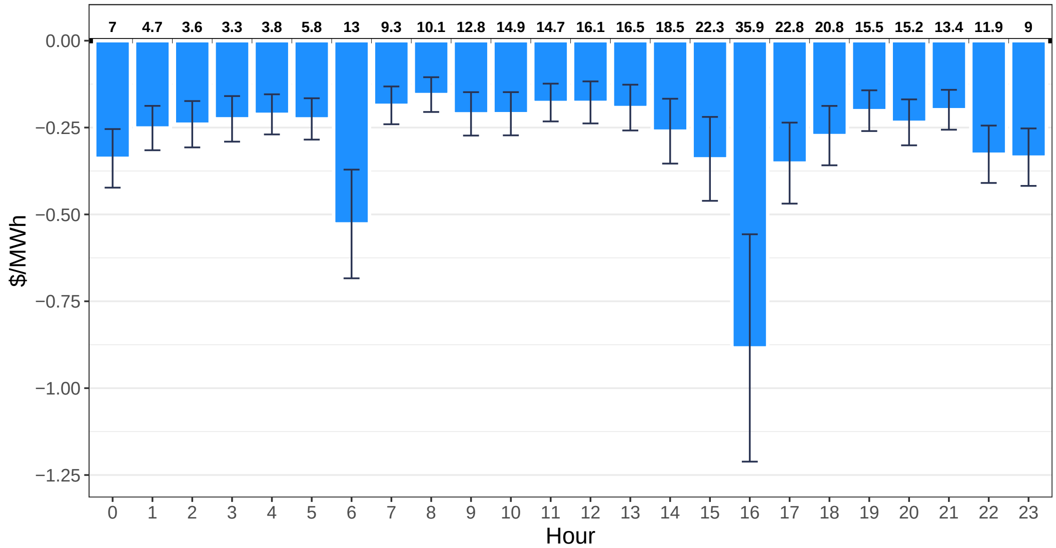
Figure 8: Coefficient estimates $(\hat{\beta}_h)$ for the effect of CREZ expansion $(crez_d = 1)$ on hourly wind generation (GWh) and the 95 percent confidence intervals.

which stands out from other hours in the earlier part of the day possibly because it marks the end of the off-peak hours of weekday in ERCOT. Compared to total wind power integrated at hour 0:00, CREZ only led to the integration of about half that amount at 16:00 ($\sim$ 0.10 GWh), but the subsequent drop in markups at 16:00 is about 2.5 times that at 0:00.

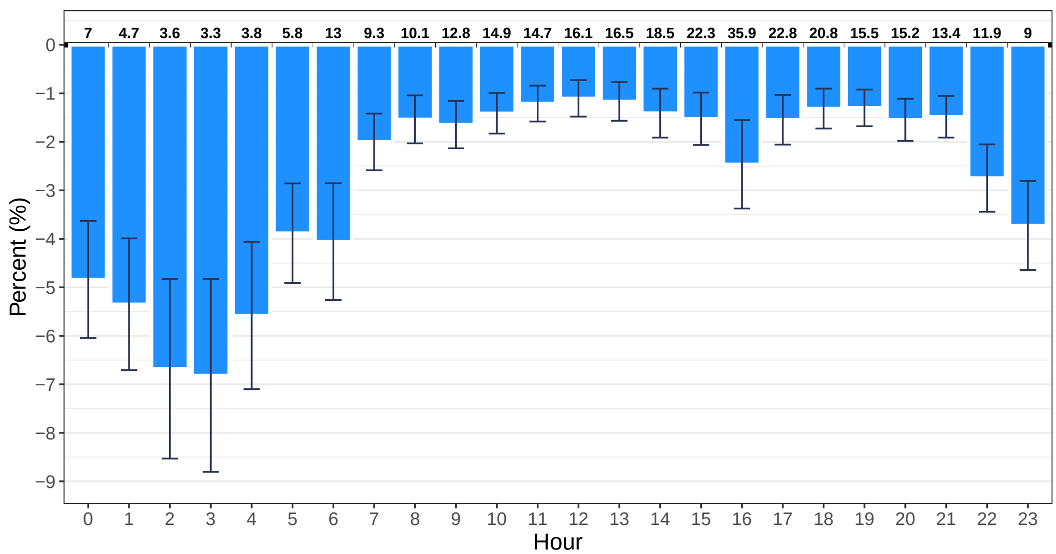
To provide a better sense of the magnitudes of $\hat{\theta}$ in Figure 9a, I present the semi-elasticity of markups in response to CREZ expansion in Figure 9b.[17] We see a clear distinction between the semi-elasticity of markups between off-peak v.s. peak hours. The magnitude of the semi-elasticity is highest for hours before 7:00 with the maximum decrease of 6.8 percent at hour 3:00. However, the percentage drop in markups for the peak hours (7:00 to 22:00) is less than 3 percent mainly because of the lower proportion of wind added to the grid during these hours.

A caveat of these results is that these are within generator changes in the markups as I use generator fixed effects in Equation 13. A simplified version of this effect is shown in Figure 4 in the theoretical model. In Figure 4c we see an inward shift in generator i's residual demand curve due to electricity generated from wind. This does not, however, capture the variation in markups due to the changes in the slope of the electricity dispatch curve as illustrated in Figure 5. I plan to incorporate this to the analysis in the near future.

17. I calculate semi-elasticity values by dividing the estimates in Figure 9a with the average markup at each hour.



(a) Effect of CREZ completion ($crez_d = 1$) on markups ($/MWh) and the associated 95 percent confidence intervals.



(b) Semi-elasticity of markups for CREZ completion ($crez_d = 1$) and the associated 95 percent confidence intervals.

Figure 9: Short-run impact of CREZ expansion on realized markups.

Note: Figure 9a shows the estimates and 95 percent confidence intervals of $\hat{\theta}_h = \hat{\alpha}_h \times \hat{\beta}_h$, where $\hat{\alpha}_h$ is the hourly impact of wind generation on markups from Figure 7 and $\hat{\beta}_h$ is the hourly impact of CREZ expansion on wind integration from Figure 8. Figure 9b is the semi-elasticity estimates corresponding to Figure 9a. Average markups for the sample are shown above the x axis.

4.5 Change in producer surplus from CREZ expansion

How do these changes in markups translate to gains or loss of producer surplus? In the short-run, marginal generators gain producer surplus in the form of excess rents from the purchasers of electricity like retailers by exercising market power. While the retail rates of electricity paid by end-use consumers remains fixed in the short-run, these excess rents are passed down from the retailers to the consumers in the long-run.

I conduct a counterfactual exercise to estimate the changes in annual rents collected by marginal fossil fuel generators due to lower markups as a result of transmission expansion. Using the parameter estimates from Equation 14, I first compute the counterfactual wind generation ($\widetilde{w}_t$) in the absence of CREZ expansion (i.e. $crez = 0$),

$$\widetilde{w}_t = \hat{\gamma} \cdot max_t + \hat{\eta}_{hm} + \hat{\omega}_t \tag{16}$$

I substitute $\widetilde{w}_t$ in the estimated Equation 13 to compute the counterfactual markups ($\widetilde{y}_{it}$) in the absence of CREZ expansion,

$$\widetilde{y}_{it} = \hat{\alpha}_h \cdot \widetilde{w}_t + f(\mathrm{D}_t|\hat{\lambda}) + \hat{\kappa}_i + \hat{\delta}_{hmy} + \hat{\epsilon}_{it} \tag{17}$$

The magnitude of increase in rents collected by generators, or more simply the change in surplus (ΔS) from the absence of CREZ expansion is:

$$\Delta S \approx \Delta(p - c) \times \widetilde{Q} \tag{18}$$

where, $\Delta(p - c)$ is the change in markups and $\widetilde{Q}$ is the power produced by the marginal generators in the absence of transmission expansion. I make two simplifying assumptions to compute ΔS. First, I assume that the gap between actual wind generation (w_t) and the counterfactual wind generation without CREZ ($\widetilde{w}_t$) is met by the power generated by marginal fossil fuel generators. In other words, this means that the wind generation integrated by grid expansion would have been met by the fossil fuel generators in the absence of the expansion. Second, I assume constant marginal costs, thus the change in markups is due to changes in the wholesale price of electricity (or LMP), for each generator. Therefore, $\Delta(p - c) = \widetilde{y}_{it} - y_{it} = p_{(\text{crez=0})} - p_{(\text{crez=1})}$.

I compute the total electricity produced by marginal generators without CREZ as:

$$\widetilde{Q}_t = Q_t + (w_t - \widetilde{w}_t) \tag{19}$$

where, Q_t is the observed hourly generation from the marginal generators and $w_t - \widetilde{w}_t$ measures the wind generation integrated by CREZ for each hour of the sample. For simplicity, I aggregate the generator data at the hourly level. This abstracts away from any distributional changes in the supply of electricity from generators in the absence of CREZ. Capturing these effects would require estimating the generator supply function which is beyond the scope of this paper.

The counterfactual analysis finds that marginal generators would have accrued about $110 million over the course of my sample in the absence of CREZ expansion. Since over 80 percent of CREZ expansion occurred throughout 2013, I use counterfactual profit values from the year 2014 to measure the annual surplus changes. I find that CREZ led to approximately $44 million annual reduction in transfers from retailers to marginal generators in the short-run and consumers of electricity in the long-run.

Note that the markup analysis only considers pricing behavior of the marginal fossil fuel generators. Thus the reduction in annual transfers is only for these marginal generators. Further this is the impact of wind integrated to the grid due to transmission expansion and not the system wide wind generation. the long-run, we expect higher wind capacity and thereby higher wind generation as a result of CREZ. Therefore, the annual benefits due to lower markups can be expected to increase as a result of higher wind capacity in the long-run.

5 Short-run: Impact of CREZ Expansion on Emissions

As shown in Section 4, the addition of wind to the grid would shift the electricity dispatch curve to the right. Further, the intermittent nature of wind generation is likely to affect which fossil fuel generator operates at the margin and therefore the emissions. In this section I examine how integration of wind due to CREZ expansion affected the emissions from the fossil fuel generator(s) at the margin at different hours of the day.

The closest empirical study in economics in this regard is Fell et al. (2021), where authors study how lower grid congestion as a result of CREZ enhanced the value of total wind generation measured by lower emissions. By contrast, I focus on how CREZ integrated more wind to the grid (keeping the generating capacity fixed) and the subsequent impact on emissions from marginal fossil fuel units in the short-run. The findings from this section are therefore complimentary to Fell et al. (2021).

The variation in the types of marginal generators over the course of a day makes such an analysis informative. For example, coal fired generators typically operate at the margin during the night whereas natural gas generators are the marginal units during the day since they are quicker to ramp up or down to meet any sudden changes in demand. The additional electricity from wind in the night could therefore displace high polluting coal generators from the margin with significant implications in terms of emissions. I run the following regression to estimate the impact of additional wing on marginal emissions:

$$E_{zt} = \rho_{zh} \cdot w_t + f(D_{zt,t-1}|\lambda) + \alpha_{zy} + \delta_{hmy} + \epsilon_{zt} \tag{20}$$

where, E_{zt} is the total emissions from fossil fuel generators at the margin in zone z and w_t is the wind generation at hour t of the sample. The parameter of interest is ρ_{zh} which measures the impact of an additional GWh of system wide wind generation on the marginal emissions in zone z at hour h.[18]

Marginal generators typically respond to changes in demand of electricity over the course of the day by ramping up or down. I use a cubic polynomial of contemporaneous and lagged demand of electricity $D_{zt,t-1}$ at the zone level to control for the variation in marginal emissions due to changes in the demand. Fixed effects δ_{hmy} control for average emission levels at hour h in month m in year y. Conditioning on these averages controls for seasonal patterns in wind generation that could also be correlated with variation in emissions. To account for longer run changes in the generation mix of wind capacity, I use zone by year fixed effects denoted by α_{zy}. Standard errors are clustered at the hour-month level to account for serial correlation.

18. I restrict my analysis to the four main zones: West, North, South, and Houston since these contain all the marginal generators affected by wind generation added as a result of CREZ.

5.1 Results

5.1.1 Impact on marginal carbon emissions

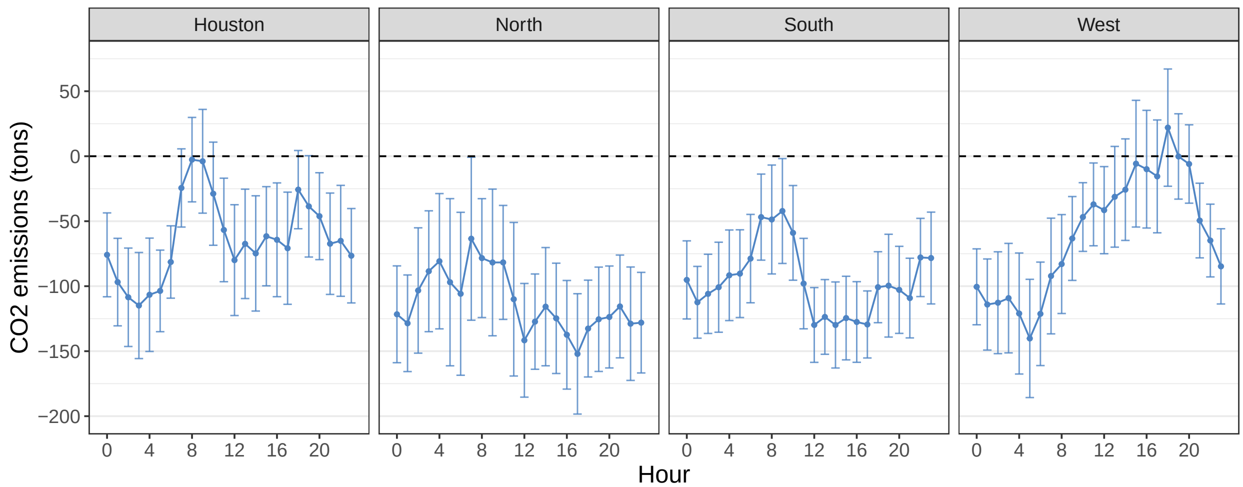
Figure 10a shows the average changes in the marginal carbon emissions (tons) for each hour of the day across the four zones in response to an additional GWh of wind energy. There is a clear decrease in emissions across all the zones throughout the day with significant spatial heterogeneity. The magnitude of decline in emissions is highest between the hours of 12:00 to 20:00 in North, South, and Houston. However, this pattern is different for the wind rich West, wherein the decline in emissions is highest at the night when the flow of wind is strongest.

The pattern in Figure 10a could be the result of the heterogeneity in the types of generators at the margin at different times of the day. To explore this, I estimate Equation 20 separately for the sample of marginal emissions from coal and natural gas. The coefficient estimates are shown in Figure 10b. Two key insights emerge. First, the hourly pattern of estimates for coal is very similar to the pattern in Figure 10a, suggesting that the carbon emissions are mainly driven by emissions from coal generators. Second, the drop in emissions from marginal natural gas generators is mostly stable and negative over the hours across all the four zones. This would shift the emissions estimates from coal downwards when aggregated by fuel type giving rise to the pattern in Figure 10a.

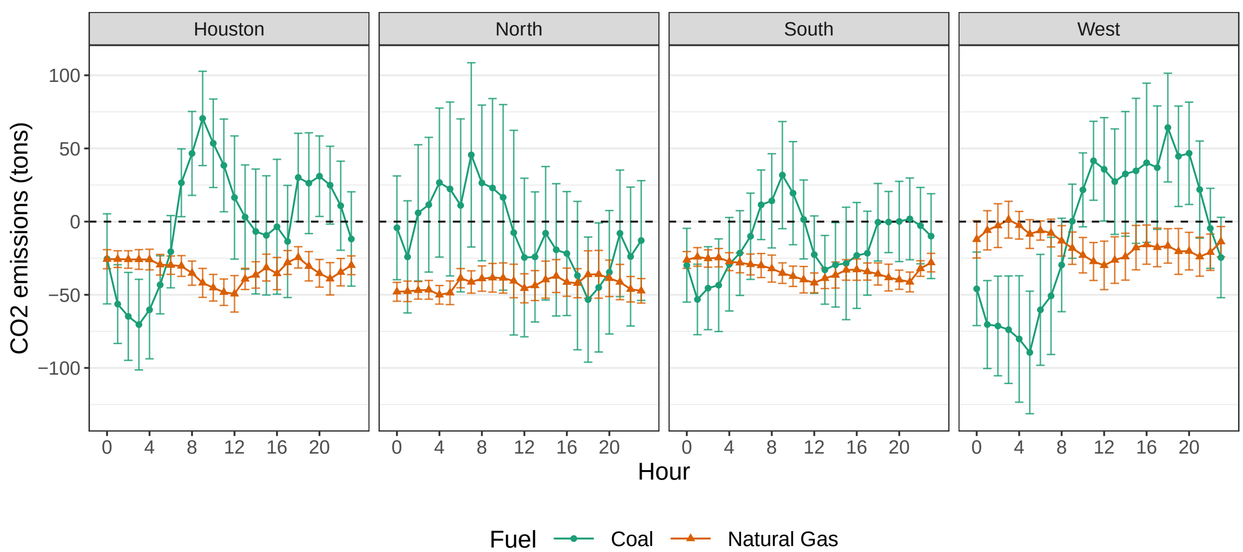
The coefficient estimates for coal generators suggests that electricity from wind has a significant effect in lowering emissions from coal generators at the margin during the night. However, we can notice the rise in emissions from coal units during the early hours of the day especially in Houston and West. This could be a consequence of intermittent wind generation during the early hours of the day leading to ramping up of coal fired power plants to meet the demand.

5.1.2 Impact on marginal local pollution (SO_2 and NOx)

To estimate the impact of hourly wind generation on damages from local pollutants, I use SO_2 and NOx emissions (tons) from marginal generators as the dependent variable in Equation 20. Figure 11a shows the coefficient estimates for SO_2 and NOx. The pattern of SO_2 emissions is very similar to that of carbon emissions from coal generators in



(a) Impact of additional wind generation (GWh) on marginal CO_2 emissions



(b) Impact of additional wind generation (GWh) on CO_2 emissions from coal and natural gas marginal generators

Figure 10: Short-run impact of wind generation on tons of CO_2 emissions.

Figure 10b.[19] This is because SO_2 is a byproduct from burning of coal in power plants due to the presence of sulphur impurities. SO_2 emissions from natural gas power plants are low because of low amounts of sulphur in pipeline quality natural gas. NOx on the other hand is released from burning of any fossil fuel due to the mixing of fuel and air (EPA 1998).

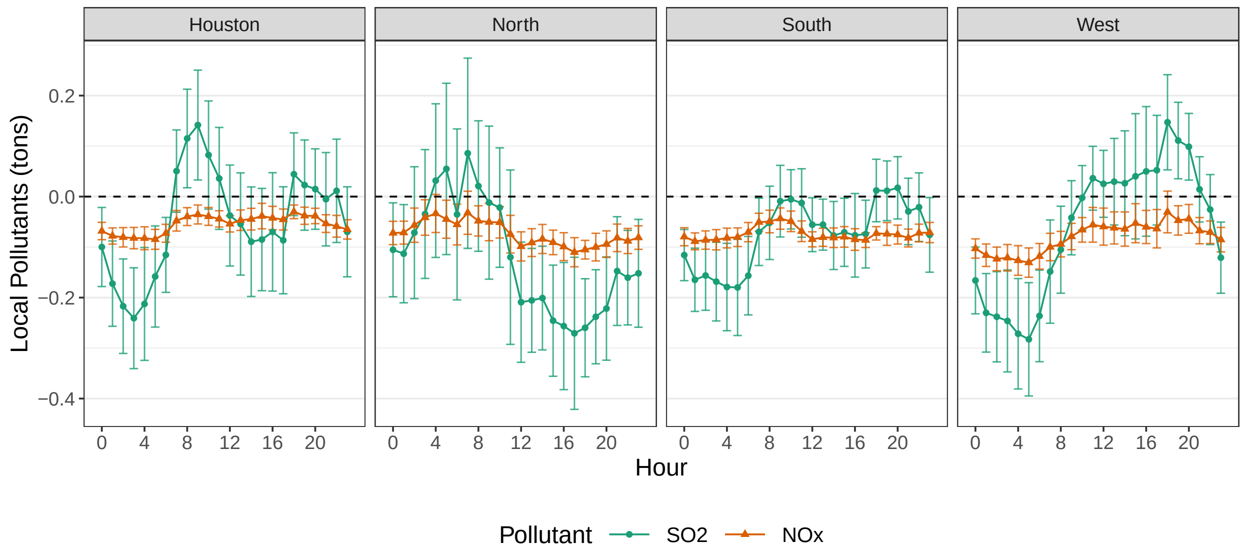
Since the health impacts of local pollutants vary across space due to differences in population, I use estimates of county-specific marginal damages due to SO_2 and NOx from Holland et al. (2016) to calculate the dollar value of damages due to emissions from each generator. I then aggregate these damages at the zonal level. Figure 11b shows the coefficient estimates using the sum of damages from SO_2 and NOx (2020 $) as the dependent variable in Equation 20. These coefficients measure the impact of additional wind generation on the damages from SO_2 and NOx across the four zones for each hour of the day.

Figure 11b shows evidence of significant heterogeneity in damages avoided from local pollutants across zones. For South and West zones, additional wind leads to decline in damages from SO_2 and NOx across all hours, whereas the effect is statistically insignificant for the North. In case of Houston, we see a significant rise in local emissions during the early hours of the day. This is similar to the rise in carbon emissions and is indicative of the ramping up of coal generators during the early hours of the day to meet the demand.

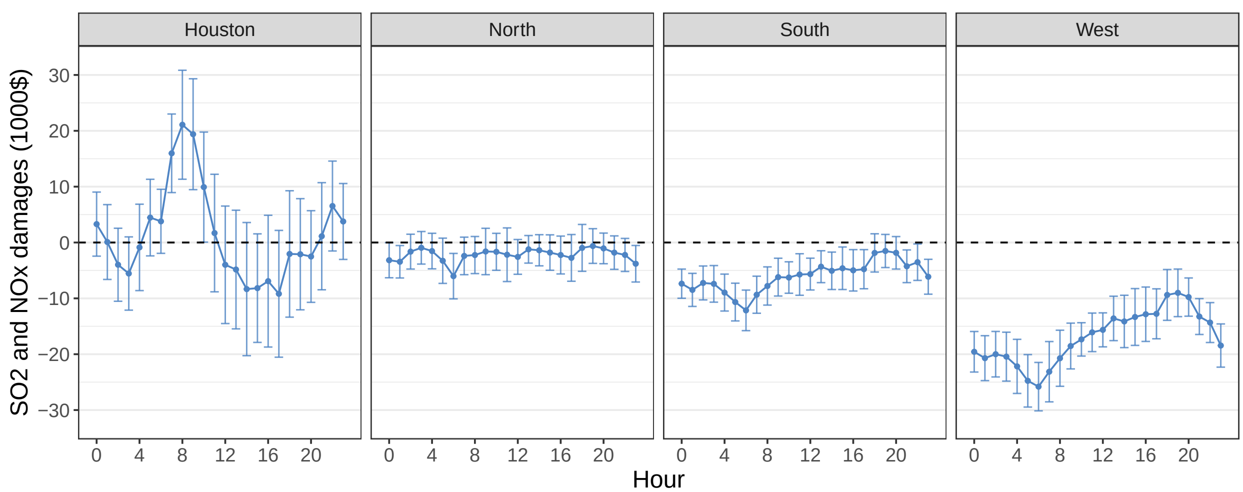
Zooming in on the marginal generators within West and Houston, we observe that the estimates for SO_2 and carbon emissions for coal are driven by the only coal power plants in these zones. In Houston, the coal emissions are due to W.A. Parish Coal Plant (four generators with total capacity of 2.7 GW) whereas in West the emissions are due to Oklaunion Power Plant (single generator with 720 MW capacity).[20] The increase in emissions in the day suggests that during the periods of low wind generation post 8:00, availability of transmission capacity tends to promote power ramping up of coal generators located near the population centers to meet the demand. These ramping

19. I present the estimates for the effect of wind generation on tons of SO_2 and NOx from coal and natural gas generators in Figure E5 in Appendix.

20. Figure E4 in Appendix E shows the location of various coal and natural gas fired marginal generators in the sample from 2011 to 2014 along with ERCOT Zones and CREZ counties.



(a) Impact of additional wind generation (GWh) on local pollutants (SO_2 and NOx) in physical units



(b) Impact of additional wind generation (GWh) on damages (2020 $) from local pollutants (SO_2 and NOx)

Figure 11: Short-run impact of wind generation on emissions from local pollutants (SO_2 and NOx).

Figure 11b uses county-specific marginal damage estimates from Holland et al. (2016) to reflect the $ value of damages from local pollution (SO_2 and NOx).

effects are shown to undercut the emissions reductions from wind, especially when operating at low levels of efficient generation (Lew et al. 2012). Furthermore, these emissions are a cause of concern from a policy perspective especially since the polluting generators are located near major population centers.

5.1.3 Value of damages avoided due to CREZ expansion

I calculate the value of marginal carbon emissions avoided due to wind integrated from CREZ expansion as:

$$D_z(\$) = \sum_{h=0}^{24} \tau \times \beta_h \times \rho_{zh} \tag{21}$$

where, D_z is the zonal daily average of damage (in 2020 \$) due to marginal carbon emissions in zone z. I assume social cost of carbon, τ as \$44 per ton of CO_2 emissions (US Interagency Working Group on Social Cost of Carbon 2014), β_h is the hourly average wind generation added due to CREZ in the short-run estimated in Equation 14, and ρ_{zh} is the impact of additional GWh of wind generation on marginal emissions. For local pollution, I simply multiply the coefficient estimates in Figure 11b with β_h and aggregate over the hours to get the value of average daily damage avoided.

Table 2 reports the daily value of damages avoided (2020 \$) from marginal emissions due to CREZ expansion for each zone.[21] We notice a decline in damages from carbon emissions across all the zones in the short-run. The total value of daily carbon emissions avoided is about \$55,000 with three fifths of the share coming from North and the South zone, and a fifth each from Houston and West. The story is a bit different for local pollutants. While the total daily damages avoided is about \$87,000, the reduction is mostly concentrated in the West followed by South and North. The increase in emissions from W.A. Parish Coal plant in Houston during early hours of the day is reflected as a net increase in damages for Houston. The total daily damages avoided from carbon, SO_2 and NOx emissions is \$144,751 which translates to about \$53 million annually.

21. The coefficient estimates of hourly averages of damages avoided for each zone due to CREZ are presented in Figure E6 in Appendix E. The pattern for carbon emissions and local pollution is similar to Figure 10a and Figure 11b respectively.

Table 2: Average daily damages (2020 $) avoided from marginal generators due to CREZ

	Damages Avoided (2020 $)			
Zone	CO_2	SO_2 + NOx	Total	Percent (%)
Houston	11,087	-6,070	5,017	3
North	17,814	8,307	26,121	18
South	15,133	23,210	38,343	26
West	11,271	63,999	75,270	52
Total	55,305	89,446	144,751	100

Notes: This table reports the daily average of damages from marginal carbon and local pollutants avoided due to additional wind integrated from CREZ expansion for each Zone.

6 Long-run: Impact of CREZ Expansion on Investment in Wind Energy

In this section I examine whether the counties announced to site CREZ transmission infrastructure saw higher levels of wind investment in the long-run. Following the enactment of the Texas Senate Bill 20, the Public Utilities Commission of Texas (PUCT) began the multi-year process of identifying the locations and cost of CREZ expansion project. In April 2008, ERCOT submitted a transmission optimization study that delineated four scenarios of transmission expansion (ERCOT 2008). Later that year, PUCT selected the second scenario which aimed to accommodate 18.5 GW of electric power by building 3,600 miles of 345 kV electricity transmission lines between existing and new substations throughout the Panhandle, West, and North zones of Texas (PUCT 2009).[22]

In the data I only see the counties where these substations were located and thus I refer to these counties as 'CREZ counties'.[23] I refer to August 2008 as the "announcement date" as it provides the first most accurate information of transmission siting in the

22. Electrical or Transmission substations typically serve as the terminal points for high voltage transmission lines as well as serve as the hub for transmission lines carrying electricity from nearby power generation plants. Therefore, wind farms and other power plants tend to locate near these substations to deliver their electricity.

23. I do not have access to exact location of these substations because it is restricted data for the purposes of national security.

CREZ transmission expansion. The specific technical details of the transmission expansion - the cost breakdown, expected completion dates, and various transmission service providers responsible for the expansion was released in October 2010 in CREZ Progress Report (RS&H 2010).

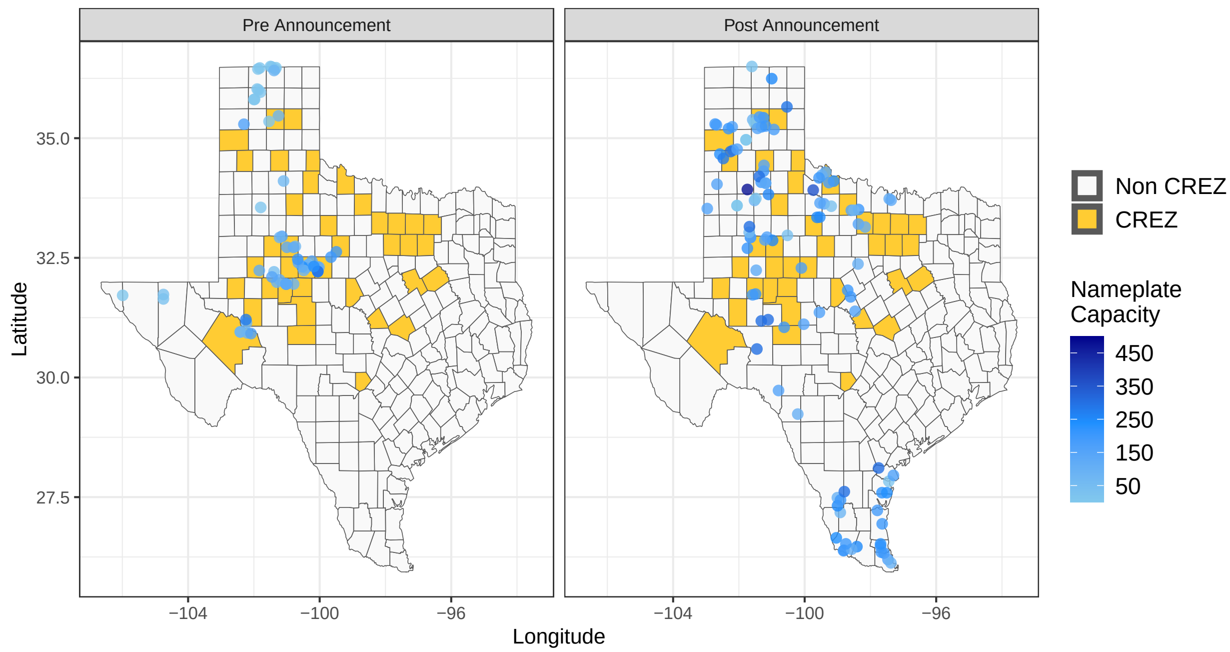


Figure 12: Location and Nameplate Capacity (MW) of wind projects Pre and Post CREZ announcement in August 2008.

Note: This figure shows wind farms in Texas Pre-CREZ Announcement (Jan 2001 - Aug 2008) and Post-CREZ Announcement (Sep 2008 - Dec 2019) samples.

From Figure 12 we can see a cluster of wind projects located within and near CREZ counties post 2008. However, there is also a cluster of wind farms in coastal Texas. This is because of a superior wind quality in this region which could be profitable for wind developers. To parse out whether CREZ counties saw higher levels of wind investment than other counties after accounting for wind quality and other confounding factors, I estimate the following specification:

$$y_{it} = \alpha + \beta \cdot crez_i + \mathbf{X}'\Pi + \epsilon_{it} \tag{22}$$

where, y_{it} is the outcome of interest. I use total wind capacity in county i in year t, average wind capacity of the project (total nameplate capacity/total number of projects in the county), and total number of turbines in county i in year t as the dependent variables for this analysis. The variable $crez_i$ is a binary variable that specifies whether county i is a CREZ county.

Since the location details for the CREZ project were announced in late 2008, the analysis is restricted to annual county level observations from 2012 through 2019. This excludes projects that were already in development or were perhaps sited in CREZ counties just prior to the grid expansion announcement. Since project planning and development typically takes a few years, this allows for the addition of wind capacity in response to transmission expansion.[24]

I use a battery of control variables and fixed effects summarized by vector **X** in Equation (22). I use wind turbine class, capacity factor, and cubic polynomial of average wind speed to flexibly control for wind resource quality of a county. These variables are aggregated at the county level from 2km by 2km grid data from NREL's Wind Integration National Datatset (WIND) toolkit (Draxl et al. 2015). I use average yearly wind project cost data from Lawrence Berkeley's Wind Technologies Report, land price data, and median land acreage compiled by the Real Estate Center at Texas A&M University to control for project costs.

To control for demographic factors that could influence CREZ transmission siting and location choice for wind projects I use median household income in 2007 and average population over 2007 to 2010. I use average farm size in a county to account for variation in wind investment due to turbine dis-amenities. This is because it is harder to site wind farms in areas with small farms (Winikoff and Parker 2019). This data comes from the USDA Census of Agriculture. The rationale behind these variables is that counties with higher household incomes, population, or average farm size could have a higher bargaining power in influencing the regulator's decision to site transmission infrastructure and wind projects (Cohn and Jankovska 2020).[25]

24. Generator interconnection is one of the first steps towards wind project development (AWEA 2019). The period between signing generator interconnection agreement and commercial operation is about 2-3 years for a typical wind project in Texas.

25. Household income, population, and average farm size for other years is highly correlated with the 2007 variables that I use in the analysis. Therefore, including values of these variables for other years in the sample does not change the results.

Cities and counties often enact regulations for wind projects that are sited in their jurisdiction. These regulations are commonly known as setbacks or wind ordinances. They usually specify limits on factors like the size of wind turbines, height of turbines, noise, maximum capacity. The presence of a wind ordinance could affect investment in wind capacity in a county and could also be correlated with siting of transmission infrastructure. I include an indicator variable specifying whether the county (or a city in the county) has a wind ordinance.[26] I use the publicly available dataset on wind ordinances from WINDExchange for this variable.

To control for Zone specific characteristics I use Zone fixed effect and cubic polynomial for time trend to control for increasing trend in wind generation across all counties. I use fixed effect for the years 2012 and 2013 to control for a sudden decline in wind installations due to Production Tax Credit (PTC) expiration in late 2012 and the subsequent extension in early 2013. Standard errors are clustered at the county level to account for serial correlation at the county level.

A key concern in Equation (22) is the endogeneity of $crez_i$ due to the selection of specific counties for CREZ expansion. That is, counties with superior wind quality and higher levels of wind capacity were selected to site CREZ substations and transmission lines. Therefore, estimating (22) using OLS would lead to biased estimates of β. To the extent that I attempt to account for these factors by including a rich array of control variables, the concern of a lack of common support amongst counties still remains. Thus, I address these concerns by implementing a matching strategy to obtain unbiased estimate of the impact of CREZ expansion on wind investment.

6.1 Matching Strategy

The objective of the matching exercise is to construct a control group of counties that are comparable to the CREZ counties on a wide set of observable characteristics. Comparing the counterfactual outcomes from the control group, conditional on confounding factors would provide the unbiased impact of transmission expansion. However, making a causal claim requires the validity of the conditional independence assumption (CIA). For the present context, CIA can be written as:

26. Most counties in Texas do not have wind ordinances for large wind projects (i.e. projects bigger than 10 MW). Out of 254 counties, I only find cities in five counties, namely Dallas, Ellis, Kleberg, Taylor, and Wichita to have enacted a wind ordinance for both smaller and bigger wind projects.

$$\mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 0) \quad (23)$$

where, ϵ_{it} is the unobserved component of dependent variable of interest (y_{it}) - wind capacity (MW), total turbines, and average project size. Under the assumption that the unobserved component (ν_i) of a county that affects the treatment status is time invariant, using county fixed effect would eliminate the selection bias. However, since the treatment variable is assigned at the county level and at the beginning of the sample, I cannot include county fixed effects.

Instead, I assume that ν_i can be approximated using some flexible function of observable county characteristics $\mathbf{Z}$, i.e. $\nu_i = f(\mathbf{Z})$. Therefore, validating CIA involves comparing counties with exactly the same combination of characteristics, such that $\mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 0)$. This would provide an estimate of the unbiased effect of the treatment. However, the presence of continuous variables in $\mathbf{Z}$ and a finite sample make it impossible to compare counties based on an exact fit of $f()$.

I use Coarsened Exact Matching (CEM) introduced by Iacus et al. (2012) to obtain the set of counties comparable on observable dimensions which includes both continuous and discrete variables. I divide the sample of counties across CREZ (treated) and non-CREZ (control) groups and then match the counties across the two groups based on observable characteristics using CEM. I use a wide variety of pre-treatment observable covariates to account for factors that could have influenced CREZ siting as well as investment in wind energy post 2012. These factors include historical wind capacity, wind resource quality, average land price and acreage, and county level demographic characteristics.

For wind resource quality, I use wind speed (m/s), capacity factor, and wind turbine class designation from NREL (Draxl et al. 2015). I use average land price over 2007-2010 and median land acreage to account for variation in project costs due to land prices across the counties. To account for citizen bargaining power and NIMBYism in citing of wind projects and transmission lines I use average farm size in 2007, median household income in 2007, and average population of a county over 2007-2010. Finally I match treated and counties exactly on ERCOT Zones to capture any regional differences due to counties being in different load zones in the Texas electricity market.[27]

Table 3 provides the balance table of these observable characteristics for pre- and post-matched samples. As evident, CEM provides a well balanced group of treated and control counties that look identical on all observable dimensions. From Table 3 we see that counties that do not lie in the common support of observable characteristics used in matching are discarded from the sample. Therefore, the control group comprises of 30 counties and the treated group comprises of 13 counties. Figure 13 shows the map of treated and control counties. Most of the control counties (light yellow) are adjacent to the treated counties (dark yellow).

For the regression analysis on the counties obtained by matching, I use the same set of control variables as described in Equation 22. The key assumption is that conditional on the vector of controls $\mathbf{X}$, there are no unobservables that affect the outcome variable and treatment status $(crez_i = 1)$. Therefore, conditional on matching the counties on $\mathbf{Z}$, selection into CREZ is "as-good-as" random. Thus, the OLS regression of Equation 22 on the matched sample gives the unbiased impact of CREZ on wind investment in the long-run.

27. Amongst the set of observable dimensions: wind capacity as of 2008, wind speed, capacity factor, average land price over 2007-2010, median land acreage, average farm size in 2007, median household income in 2007, and average population over 2007-2010 are continuous whereas Power Curve and Zone are discrete variables. Each category within Power Curve is matched exactly whereas I use the following structure for exact matching on Zone: {{Panhandle, West}, North, Coastal, Houston, South, None}.

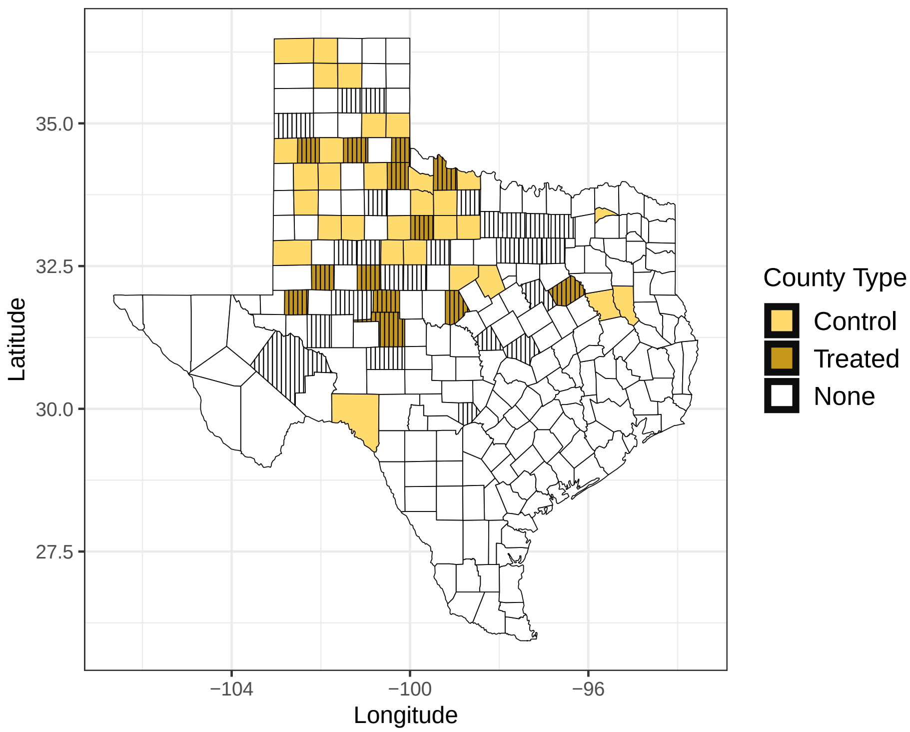


Figure 13: Treated and Control counties obtained using Coarsened Exact Matching.

Note: Hatched counties depict the CREZ counties in the overall sample. Unshaded hatched and non-hatched counties are discarded from the sample used in the analysis as they lie outside of the common support of observable characteristics.

Table 3: Balance Table of key observables for Pre- and Post-Matching Sample

Variables	Pre-Matching			Post-Matching		
	Means Treated [CREZ = 1]	Means Control [CREZ = 0]	p-value	Means Treated [CREZ = 1]	Means Control [CREZ = 0]	p-value
Wind Capacity as of 2008 (MW)	158.599	5.579	0.000	5.581	4.264	0.138
Wind Speed (m/s)	7.923	7.348	0.000	7.887	7.891	0.619
Capacity Factor	0.449	0.413	0.000	0.437	0.439	0.949
Wind Turbine Class = 1	0.000	0.005	–	0.000	0.000	–
Wind Turbine Class = 2	0.692	0.393	–	0.837	0.837	–
Wind Turbine Class = 3	0.308	0.603	–	0.163	0.163	–
Avg. Land Price (2007-2010)	284.684	424.427	0.000	228.424	231.216	0.929
Median Land Acreage	560.184	779.632	0.032	360.746	351.736	0.161
ERCOT Zone: Coastal	0.000	0.051	–	0.000	0.000	–
ERCOT Zone: Houston	0.000	0.028	–	0.000	0.000	–
ERCOT Zone: None	0.000	0.107	–	0.000	0.000	–
ERCOT Zone: North	0.308	0.220	–	0.163	0.163	–
ERCOT Zone: Panhandle	0.179	0.136	–	0.302	0.371	–
ERCOT Zone: South	0.026	0.252	–	0.000	0.000	–
ERCOT Zone: West	0.487	0.206	–	0.535	0.466	–
Avg. Farm Size in 2007	1,595.667	1,724.206	0.418	1,183.140	1,262.035	0.118
Median Income in 2007	43,133.130	39,739.930	0.000	35,789.190	35,574.620	0.837
Avg. Population (2007-2010)	171,282.000	83,280.770	0.002	28,917.870	20,612.030	0.026
Total Units	312	1,712		104	240	

Notes: This table presents balance test of key pre-treatment observable characteristics of a county. Each unit is a county-year observation. Pre-Matching sample includes all county-year observations. Post-Matching sample is selected using Coarsened Exact Matching (CEM).

6.1.1 Results

Table 4 reports the regression results of the baseline specification with total nameplate capacity (MW), total turbines, and average project capacity (MW) in a county as the dependent variables respectively. These regressions use the full set of control variables, i.e. cubic polynomial of time trend, controls for wind resource quality, project costs, county demographics, fixed effects for wind ordinance, ERCOT zone, and Production Tax Credit. I also include the interaction of group fixed effects with the time trend to allow for time varying unobserved factors that could affect specific matching groups.

The results for total nameplate capacity indicate a significant increase in wind capacity in CREZ counties. Column (1) in Table 4 shows that transmission expansion led to approximately 73 MW higher wind capacity in treated counties over 2012 - 2019. The semi-elasticity indicates a 202.3 percent increase in wind capacity for CREZ counties. In a similar vein, Column (2) shows that treated counties on average had about 40 more turbines than the control counties with a 'semi-elasticity' of 245 percent. Both of these results are statistically significant at 5 percent critical level.

Column (3) examines whether the size of a wind project varies differentially with county type. Everything else equal, we might expect wind developers to build bigger wind projects near sites that allow access to transmission capacity and therefore a positive coefficient. The coefficient estimate lends weak evidence in favor of this hypothesis. I find that CREZ counties were associated with 33 MW bigger wind projects, however, the coefficient estimate is only significant at 10 percent critical level.

In order to contextualize these estimates, I compute the value of carbon emissions avoided due to wind investment as a result CREZ expansion. I use an emissions rate of 0.601 tons of CO_2 avoided for each MWh of on-shore wind in Texas (EPA 2021). Assuming the capacity factor of wind as 34.57 percent as before, wind added due to CREZ avoided roughly 5.34 million tonnes of CO_2 emissions from the power sector in Texas in 2019. Using a social cost of carbon of \$51/ton for 2019 (citation), the value of total reduction in carbon emissions in 2019 in Texas is about \$271 million.[28]

28. The total value of damages prevented from emissions is likely to be much higher if we include local pollutants. However, calculating this will require computing the SO_2 and NOx offsets due to additional wind across space in 2019. Such an analysis is beyond the scope of this paper.

Table 4: Effect of CREZ expansion on wind investment - matching results

	Dependent variable		
	Total Nameplate Capacity (MW)	Total Turbines	Avg. Capacity of a project (MW)
	(1)	(2)	(3)
CREZ	72.640***	39.419***	32.756*
	(26.499)	(13.075)	(19.093)
Mean Dep. Variable	35.907	16.067	26.951
Semi-elasticity (%)	202.3	245.3	121.5
Controls	✓	✓	✓
Group × Trend FE	✓	✓	✓
Matching Weights	✓	✓	✓
Sample	Matched	Matched	Matched
Observations	344	344	344
R^2	0.467	0.476	0.426

Notes: This table reports the result of baseline regressions on the matching sample. The sample is a balanced panel of 13 treated (CREZ) and 30 control (non-CREZ) counties from 2012-2019 obtained using CEM. The independent variable is a binary variable indicating whether a county received CREZ transmission infrastructure or not. All specifications include cubic polynomial of time trend and controls for wind quality, project cost, county level regulation and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. I also include group by trend fixed effects to allow for time varying unobserved factors affecting matching groups. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

6.1.2 Anticipation effects, multi phase projects, and Robustness Checks

A potential source of bias in measuring the causal impact of CREZ on wind investment could be the anticipation amongst wind developers for CREZ expansion announcement in 2008. This would be reflected as a spike in investment in wind projects within CREZ counties in the years leading upto the transmission expansion announcement. Using the data on generator interconnection in ERCOT, I verify the existence of such an anticipation effect in Appendix C. The results allow me to rule out the anticipation effects both two and four years prior to the announcement of transmission expansion.

Another threat to identification could be if projects within CREZ counties prior to 2008 saw subsequent extensions shortly after 2012. This would be a selection issue in the sense that a county gets selected to site CREZ infrastructure because of the likely development of a project extension in the near future. To address this concern I examine the occurrence of post 2012 extensions for wind projects that started operating before 2008 within CREZ counties. Figure E7 in Appendix shows that existence of multi-phase wind projects and project extensions are not a cause of concern.

I conduct a series of robustness checks to see how the coefficient estimates change with different combinations of control variables and fixed effects. I also check how the results change when I exclude group fixed effects and the matching weights. The results for all the three dependent variables are similar to the estimates of baseline specification reported in Table 4. Appendix F presents the tables for these robustness checks.

7 Discussion and Policy Implications

Efforts to combat climate change in the US have largely focused on expanding the solar and wind generating capacity. A key factor in ensuring that the power generated through these sources is fully utilized is the availability of transmission lines that could carry the electricity to demand centers. Using the CREZ transmission expansion in Texas as the case study, this paper studies the short- and long-run impacts of large scale grid expansion. I examine the impact of grid expansion on markups and emissions from marginal fossil fuel generators in the short-run and wind investment in the long-run.

The short run analysis suggests that CREZ had a moderate effect on lowering markups- about 2.5% during peak demand hours and 7% during off-peak hours. These lower markups prevented about $44 million in annual rents accrued by marginal generators from power retailers in the short-run. Further, CREZ prevented about $53 million in annual damages from marginal emissions. About 60% of the damages avoided are from local pollutants (SO_2 and NOx) with the remaining share from carbon emissions. I find an increase in emissions as a result of ramping up of marginal coal generators due to wind intermittency during the early hours of the day and is most pronounced in the West and Houston region of Texas.

Next, I estimate the extent to which grid expansion spurs investment in wind energy in the long-run. I use coarsened exact matching to address the selection issue that locations with superior wind quality were selected to site CREZ lines and substations. OLS regressions on the matched sample suggests that counties with CREZ transmission infrastructure saw significant investment in wind capacity (+202%) over 2012 - 2019. A back of the envelope calculation shows that the wind capacity added due to CREZ prevented approximately $271 million worth of carbon emissions in Texas in 2019.

The results in this paper have several policy implications. First, the short-run results highlight the role of technologies like grid-scale energy storage in mitigating some of the negative impacts of wind intermittency. Grid-scale storage can reduce wind curtailment and use the stored electricity at peak demand when the wind generation is low. This can be useful in further lowering markups, in turn enhancing the benefits of transmission expansion. Energy storage can also address the rise in emissions due to wind intermittency by utilizing the stored electricity to meet the excess demand in periods of low wind in order to prevent the ramping-up of coal generators.

Second, the long-run analysis provides an estimate of the potential forgone wind capacity had there been no CREZ expansion. This speaks to the potential impacts of large scale grid expansion projects on renewable investment in other parts of the US like the Midwest. However, adequate transmission capacity would be required in order to avoid congestion due to growing wind investment in parts of Iowa that are far from the demand centers.

Third, the impacts of grid expansion highlight the need for investments in grid resilience and upgrades. This is because growing wind capacity can put a strain on the transmission system thereby undoing some of the gains of transmission expansion. The impacts of such investments are particularly salient during the periods of grid congestion and weather shocks that can lead to wind curtailments to maintain grid reliability. These curtailments can also undo some of the estimated short-run gains from transmission expansion.

These findings open up several avenues for future research. A key issue with renewable sources is their intermittent nature. As discussed before, grid-scale energy storage can be a useful step to address some of the concerns highlighted in this paper. Karaduman (2020) is an important first empirical analysis looking at the incentives for investment in energy storage. Using a dynamic equilibrium model, the author finds heterogeneous impacts of grid-level storage on the returns from renewable production and the need for a capacity market to incentivize private players to invest in storage. However, more research is needed to understand the impacts of the interaction between grid expansion and grid level storage.

Wind intermittency could also lead to dynamic effects on the costs of electricity generation from fossil fuel plants. Jha and Leslie (2021) study this in the context of the Australian electricity market and its transition to greater solar rooftop capacity. They find higher market power exercised by fossil fuel plants in the evening as some plants choose not to incur the ramp-up costs necessary to produce at sunset thereby lowering the competition in the evening. While the short-run analysis in this paper does not account for such dynamic changes in markups due to generator ramp-up, it is a worthwhile extension that would provide a more precise impact of grid expansion on generator markups.

References

AWEA. 2019. *US Wind Industry Annual Market Report.* Technical report. American Wind Energy Association, Washington DC (United States).

Billo, Jeff. 2017. *The Texas Competitive Renewable Energy Zone Process.* Electric Reliability Council of Texas.

Birge, John R, et al. 2018. "Limits to Arbitrage in Electricity Markets: A case study of MISO." *Energy Economics* 75:518–533.

Borenstein, Severin, et al. 2000. "The competitive effects of transmission capacity in a deregulated electricity industry." *The Rand Journal of Economics* 31 (2): 294.

Borenstein, Severin, et al. 2002. "Measuring market inefficiencies in California's restructured wholesale electricity market." *American Economic Review* 92 (5): 1376–1405.

Borenstein, Severin, et al. 2008. "Inefficiencies and market power in financial arbitrage: a study of California's electricity markets." *The Journal of Industrial Economics* 56 (2): 347–378.

Bushnell, James B, et al. 2008. "Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets." *American Economic Review* 98 (1): 237–66.

Cohn, Julie, and Olivera Jankovska. 2020. *Texas CREZ Lines: How Stakeholders Shape Major Energy Infrastructure Projects.* Technical report. Baker Institute for Public Policy, Rice University.

Cullen, Joseph. 2013. "Measuring the environmental benefits of wind-generated electricity." *American Economic Journal: Economic Policy* 5 (4): 107–33.

Davis, Lucas, and Catherine Hausman. 2016. "Market impacts of a nuclear power plant closure." *American Economic Journal: Applied Economics* 8 (2): 92–122.

Doshi, Gaurav, and Xiaodong Du. 2021. "Transmission Integration and the Market for Congestion Revenue Rights." *The Energy Journal* 42 (5).

Draxl, Caroline, et al. 2015. "The wind integration national dataset (wind) toolkit." *Applied Energy* 151:355–366.

EIA. 2019. "Electricity Data Browser." US Energy Information Administration. Data retrieved from: `https://www.eia.gov/electricity/data/browser/`.

EPA. 1998. *AP 42, Fifth Edition, Volume I Chapter 1: External Combustion Sources.* Technical report. Retrieved from: https://www.epa.gov/air-emissions-factors-and-quantification/ap-42-fifth-edition-volume-i-chapter-1-external-0.

———. 2021. *Avoided emissions and generation tool (AVERT).* U.S. Environmental Protection Agency (EPA) Office of Air and Radiation Climate Protection Partnerships Division, Washington D.C. Data retrieved from: https://www.epa.gov/avert/download-avert/.

ERCOT. 2008. *CREZ Transmission Optimization Study.* Electric Reliability Council of Texas, TX (United States), April. https://www.nrc.gov/docs/ML0914/ML091420467.pdf.

———. 2021. "Fact Sheet." Electric Reliability Council of Texas. Retrieved from: http://www.ercot.com/content/wcm/lists/219736/ERCOT_Fact_Sheet_1.19.21.pdf.

Fell, Harrison, and Jeremiah X Johnson. 2021. "Regional disparities in emissions reduction and net trade from renewables." *Nature Sustainability* 4 (4): 358–365.

Fell, Harrison, and Daniel T Kaffine. 2018. "The fall of coal: Joint impacts of fuel prices and renewables on generation and emissions." *American Economic Journal: Economic Policy* 10 (2): 90–116.

Fell, Harrison, et al. 2021. "Emissions, transmission, and the environmental value of renewable energy." *American Economic Journal: Economic Policy* 13 (2): 241–272.

Hernández, Cristián. 2018. "Resource Misallocation in an Oligopoly Environment: Evidence from Texas' Electricity Market." Working Paper.

Holland, Stephen P, et al. 2016. "Are there environmental benefits from driving electric vehicles? The importance of local factors." *American Economic Review* 106 (12): 3700–3729.

Hortacsu, Ali, and Steven L Puller. 2008. "Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market." *The RAND Journal of Economics* 39 (1): 86–114.

Iacus, Stefano M, et al. 2012. "Causal inference without balance checking: Coarsened exact matching." *Political analysis:* 1–24.

Ito, Koichiro, and Mar Reguant. 2016. "Sequential markets, market power, and arbitrage." *American Economic Review* 106 (7): 1921–57.

Jha, Akshaya, and Gordon Leslie. 2021. "Start-up Costs and Market Power: Lessons from the Renewable Energy Transition." Working Paper.

Joskow, Paul L, and Jean Tirole. 2000. "Transmission rights and market power on electric power networks." *The Rand Journal of Economics:* 450–487.

———. 2005. "Merchant Transmission Investment." *The Journal of Industrial Economics* 53 (2): 233–264.

Kaffine, Daniel T, et al. 2013. "Emissions savings from wind power generation in Texas." *The Energy Journal* 34 (1).

Karaduman, Omer. 2020. "Economics of grid-scale energy storage." Working Paper.

LaRiviere, Jacob, and Xueying Lu. 2020. "Transmission constraints, intermittent renewables and welfare." Working Paper.

Larson, Eric, et al. 2021. *Net-Zero America: Potential Pathways, Infrastructure, and Impacts, interim report.* Technical report. Retrieved from: `https://netzeroamerica.princeton.edu/the-report`. Princeton University, Princeton, NJ.

Lasher, Warren. 2014. "The Competitive Renewable Energy Zones Process." In *Quadrennial Energy Review Public Meeting: State, Local and Tribal Issues.* US Department of Energy, Santa Fe, NM.

Lew, Debra, et al. 2012. *Impacts of wind and solar on fossil-fueled generators.* Technical report. National Renewable Energy Lab.(NREL), Golden, CO (United States).

Mercadal, Ignacia. 2018. "Dynamic Competition and Arbitrage in Electricity Markets: The Role of Financial Traders." Available at SSRN 3281886.

Meyer, Robinson. 2021. "Unfortunately, I Care About Power Lines Now." Retrieved from: `https://www.theatlantic.com/science/archive/2021/07/america-is-bad-at-building-power-lines-lets-fix-that-transmission-climate/619591/`, *The Atlantic.*

New York Times. 2016. "Making Way for Wind Power." Editorial. Retrieved from: `https://www.nytimes.com/2016/04/05/opinion/making-way-for-wind-power.html`.

Novan, Kevin. 2015. "Valuing the wind: renewable energy policies and air pollution avoided." *American Economic Journal: Economic Policy* 7 (3): 291–326.

PUCT. 2007. *Interim Order on Reconsideration Docket No. 33672.* Technical report. Public Utility Commission of Texas, TX (United States).

———. 2009. *Docket 35665, PUC CADM: Final Order.* Public Utility Commission of Texas, Austin, TX (United States), March.

RS&H. 2010. *CREZ Progress Report No. 1.* Technical report. Public Utility Commission of Texas, TX (United States).

Ryan, Nicholas. 2021. "The Competitive Effects Of Transmission Infrastructure In The Indian Electricity Market." *American Economic Journal: Microeconomics* 13 (2): 202–242.

Saravia, Celeste. 2003. "Speculative Trading and Market Performance: The Effect of Arbitrageurs on Efficiency and Market Power in the New York Electricity Market." University of California Berkeley CSEM Working Paper 121.

Temple, James. 2017. "How to Get Wyoming Wind to California, and Cut 80% of U.S. Carbon Emissions." Retrieved from: `https://www.technologyreview.com/2017/12/28/146566/how-to-get-wyoming-wind-to-california-and-cut-80-of-us-carbon-emissions/`, *MIT Technology Review.*

The White House. 2021. *Fact Sheet: President Biden Sets 2030 Greenhouse Gas Pollution Reduction Target Aimed at Creating Good-Paying Union Jobs and Securing U.S. Leadership on Clean Energy Technologies.* The United States Government. Washington, D.C. Online; accessed 15-May-2021. `https://www.whitehouse.gov/briefing-room/statements-releases/2021/04/22/fact-sheet-president-biden-sets-2030-greenhouse-gas-pollution-reduction-target-aimed-at-creating-good-paying-union-jobs-and-securing-u-s-leadership-on-clean-energy-technologies/`.

Thill, David. 2020. "As bottleneck stymies projects, Midwest groups call for transmission reforms." Retrieved from: `https://energynews.us/2020/01/10/as-bottleneck-stymies-projects-midwest-groups-call-for-transmission-reforms/`, *Energy News Network.*

US Interagency Working Group on Social Cost of Carbon. 2014. *Technical support document: Technical update of the social cost of carbon for regulatory impact analysis under executive order 12866.* Technical report. American Council for an Energy Efficient Economy, Washington, DC.

Winikoff, Justin, and Dominic Parker. 2019. “Farm Size, Spatial Externalities, and Wind Energy Development.” Working Paper.

Woerman, Matt. 2019. “Market size and market power: Evidence from the Texas electricity market.” Energy Institute Working Paper 298.

Appendix

A Data Sources and Sample Construction

A.1 Data and sample for markup analysis

In this section, I describe the sample construction for the short-run analysis. The hourly generator level sample used in the short-run analysis on the effect of CREZ expansion on markups uses data from three sources - ERCOT Report 13029, EIA Form 860, and EPA's CEMS Data. A brief description of these data sources is as follows:

ERCOT Report 13029 This report includes the offer price and the name of the entity submitting the offer for the highest-priced offer selected or dispatched by the Security Constrained Economic Dispatch (SCED) two days after the applicable operating day. It identifies all the entities that submitted the highest-priced offers selected for each SCED run (in case of multiple entities). SCED is the market clearing process in ERCOT and occurs at every 15 minutes. Therefore, this data is at 15 minute intervals for August 2011 to December 2014. I aggregate this data at the hourly level and all the generators that appear in this data in a specific hour are regarded as marginal generators for that hour. Apart from the identity of the generation resource, this dataset also includes the Locational Marginal Price (LMP) resolved at the resource node for that generator. This acts as the wholesale price corresponding to the marginal generator.

EIA Form 860 This is an annual dataset of all the power plants and generators operating in the US. This data contains information like EIA code of the power plant and generator(s), plant name, location, generator technology, prime mover, main energy source, regulatory status of the power plant, nameplate capacity, operating month and year, planned retirement year, operating status etc.

CEMS Data This is an hourly level data of all the fossil fuel generators at least 25 MW in size. It contains information on hourly emissions (CO_2, NOx, and SO_2), hourly generation, and heat input. The generators are identified using ORISPL Code.

For my sample period, ERCOT Report 13029 contains about 300 fossil fuel generators that operate at the margin at some instance. Since I do not observe the EIA Plant Code or Generator ID in ERCOT Report 13029, I manually match each of the 300 fossil fuel generators to the corresponding generators in the EIA Form 860. I am able to successfully match most of the generators in the ERCOT data to EIA Data.

The next part of sample construction is to match the generator data in EIA to hourly generator data in CEMS. The generator identifiers in CEMS are the ORISPL Code and Unit ID. ORISPL Code corresponds directly to the EIA Plant Code for most cases. I verify and correct ORISPL Codes in case of any discrepancy. Similarly, Unit ID in CEMS data corresponds directly to generator id in EIA Form 860. However, I verify and correct all the cases where there is any discrepancy.

A.2 CREZ Transmission Expansion Data

I use Transmission Project Information Tracking (TPIT) Reports obtained from ERCOT to assemble the dataset on CREZ transmission expansion. These reports contain detailed information on various electricity transmission projects in Texas. I specifically focus on new transmission lines built as a part of CREZ project. These reports provide the length of each transmission line (in miles) along with their in-service dates. I also see the counties where the terminals of each specific line lies. These terminals are usually existing or new electrical substations. The data on the exact location of these substations is restricted since it is considered a matter of national security, thus, I only see the county where these substations are located.

Following counties are classified as 'CREZ' counties in my data: Archer, Bell, Borden, Briscoe, Brown, Carson, Castro, Childress, Coke, Collin, Cottle, Dallas, Deaf Smith, Denton, Dickens, Ector, Glasscock, Gray, Haskell, Hill, Jack, Kendall, Lampasas, Martin, Mitchell, Navarro, Nolan, Parker, Pecos, Schleicher, Scurry, Shackelford, Sterling, Tarrant, Taylor, Tom Green, Upton, Wilbarger, Wise.

B Institutional Details

B.1 Real-time electricity market

Real-time market operations mainly refers to the operating hour and the hour immediately preceding the operating hour. ERCOT collects the status of all the transmission infrastructure from Transmission Service Providers and identifies transmission constraints and forecasts demand at various points of the network for the operating hour. This information is made available to the supply side of the market that comprises of the generating firms.

To participate in the market, each firm submits offer curves for all the generators that it owns. These offer curves are monotonically increasing vectors of price-quantity pairs based on the demand and grid information provided by ERCOT. Firms enjoy great flexibility to specify and alter their offer curves which can be different for different hours of the day. They can input up to ten price-quantity pairs and alter their offer curve up to the hour preceding the operating hour. This allows a firm to update its strategy when more information on various factors like demand, transmission constraints, or strategies of competitors is available.

The demand side of the market is comprised of retailers and load serving entities who submit demand for energy at various locations in the operating hour. Equipped with the information on supply, demand, and transmission constraints, ERCOT deploys a market clearing process that occurs every 5 minutes. This process identifies least cost generating resources that would meet the electricity demand at various locations in the system while respecting transmission constraints and the capacity limits of the generating resources. Apart from matching supply to demand, a major task of this process is to prevent the system from exceeding operational limits thus maintaining the reliability of the network. This market clearing process generates market clearing prices called Locational Marginal Price which is the location specific wholesale price of electricity.

B.2 Transmission constraint and inefficient dispatch of generators

Presence of transmission constraints prevents ERCOT from dispatching the cost effective generating units to meet the demand at a particular location. To see this, consider a simple example. Say there are two regions A and B. Region A consists of low cost generators that can provide up to 100 MW of electricity and region B consists of high cost generators that can also provide 100 MW of electricity. However, Region A and B are connected by a transmission line that can carry only 50 MW of electricity. Suppose at some time *t* there is a demand for 80 MW of electricity in region B by households. ERCOT as the planner, would like to dispatch all of the 80 MW from low cost generators in Region A. However, due to the transmission limit it can only dispatch 50 MW. At this point, the transmission constraint between A and B is said to be binding or there is transmission congestion between A and B. To meet the remaining demand, ERCOT has to dispatch 30 MW of electricity from high cost generators located in region B. Thus, presence of transmission constraints leads to dispatch of higher cost generators when the demand could have been met by low cost generators.[29]

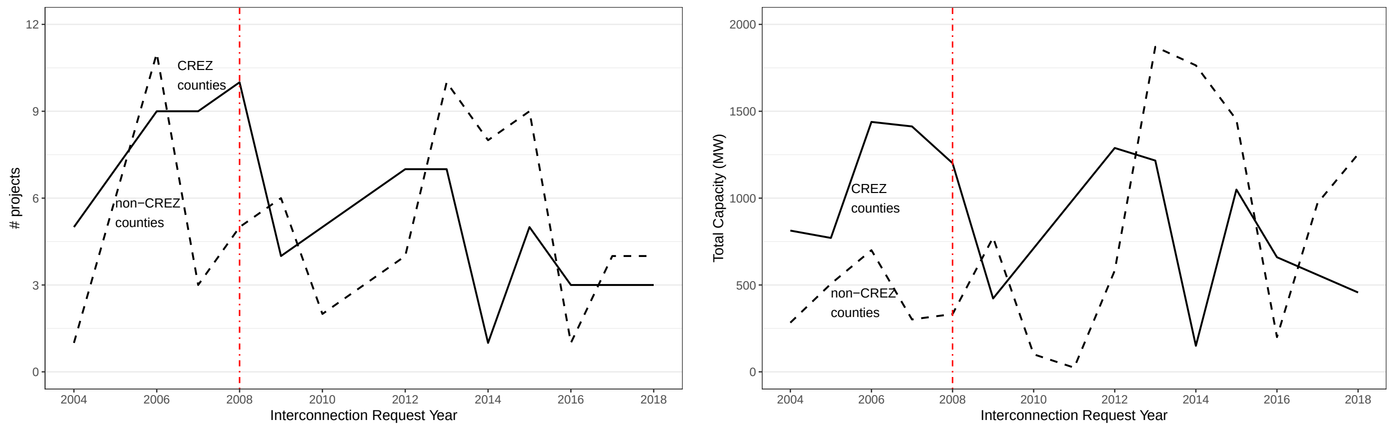
29. The dispatch of electricity in reality is more complicated since the flow of current follows Kirchhoff's Laws. This example abstracts from such real life aspects in order to illustrate the impact of transmission constrains on generator dispatch.

C Anticipation Effects

In this section, I examine whether there was an anticipation amongst wind developers to invest in wind projects in the period leading upto the announcement of CREZ transmission expansion in late 2008. Existence of such an anticipation could lead to biased estimates of the impact on CREZ announcement on wind investment in Section 6.1. The direction of the bias is expected to be downwards since the coefficient estimate would not capture the wind investment in periods before the announcement.

To examine the anticipation effects, I use information on generator interconnection as a measure of changes in wind project planning in ERCOT since the latter is usually unobserved or hard to measure. Wind developers usually sign the interconnection agreement if they expect to build a project at a particular site and this is usually one of the first steps in the process of building a wind project (AWEA 2019).I use interconnection data from EIA Form 860 for the years 2004 - 2012 and Generator Interconnection Status (GIS) Reports from ERCOT for the years 2013 - 2019 to get the date when a wind project signed the interconnection agreement. I match these data with the wind project data from EIA 860 and AWEA to get information on project level characteristics. The matched dataset comprises of 147 projects that signed the interconnection agreement between 2004 and 2018. In terms of successful matches, this represents about 87 percent of the existing wind projects in Texas between 2004-2018.

Figure 14 shows the number of projects and the total capacity of projects located in CREZ and non-CREZ counties that signed the interconnection agreements over 2004 - 2018. Any anticipation effect of transmission expansion would be marked by a spike in the number of projects signing the interconnection agreement mainly in CREZ counties in the years leading upto grid expansion announcement in 2008. We might also expect a rise in the total capacity of projects signing the interconnection agreements prior to 2008. However, from Figure 14a and Figure 14b we do not notice any clear pattern in either the number of project or the total capacity for the years leading upto 2008 in both CREZ and non-CREZ counties.



(a) Total # projects signing the interconnection agreement

(b) Total capacity (MW) of projects signing the interconnection agreement

Figure 14: # projects and wind capacity in the ERCOT interconnection queue over 2004 - 2018

Note: Solid line is corresponding to CREZ counties and dashed line is corresponding to non-CREZ counties. Dashed vertical line indicates the year of CREZ announcement.

I estimate specifications to test the existence of an anticipation effect after controlling for confounding factors that could influence generator interconnection. Specifically, I estimate versions of the following specification:

$$y_{it} = \beta \cdot \mathbb{1}\{year \in [k,\ 2008]\} + \alpha_i + \mathbf{X}'\Pi + \epsilon_{it} \tag{24}$$

where, y_{it} is the inverse hyperbolic sines (IHS) of number of projects or the total nameplate capacity of projects that signed the interconnection agreement in county i in year t. The independent variable of interest $\mathbb{1}\{year \in [k,\ 2008]\}$ is an indicator for the range of years from k to 2008, denoting the anticipation period. I consider two versions of this variable - $k = 2006$, i.e. $\mathbb{1}\{year \in [2006,\ 2008]\}$ and $k = 2004$, i.e. $\mathbb{1}\{year \in [2004,\ 2008]\}$ as the anticipation period. I estimate Equation 24 separately for CREZ and non-CREZ counties.

I use a rich set of covariates to control for confounding factors. I use county fixed effects denoted by α_i and a vector of county and demographic controls summarized by $\mathbf{X}$. This includes a linear time trend, cubic polynomial of county specific wind speed, capacity factor of wind generation, median land acerage, real price of land, indicator for whether the county has a wind ordinance, average farm size (acres) in 2007, median household income, and log of population. To account for correlation in interconnection queue across counties, I cluster the error ϵ_{it} at the county level.

Table C1 reports the results of OLS regression of Equation 24 with [2006, 2008] as the anticipation period. Column (5) is the baseline specification for the sample using CREZ counties and Column (6) is the baseline specification for the sample using non-CREZ counties. Panel A shows the results for IHS of the number of projects in interconnection as the dependent variable. The coefficient estimates suggest that anticipation effect for both CREZ and non-CREZ counties is positive but statistically and economically insignificant. Restricting the sample to counties obtained using matching (Panel A.2) in the long-run analysis does not change the results by much with the exception of the estimate for non-CREZ counties. I find a weak positive effect with an elasticity of 8 percent, however the coefficient is only significant at 10 percent critical level.

Panel B shows the results for IHS of the total capacity of projects in interconnection as the dependent variable. I find a positive anticipation effect for CREZ counties but it is not statistically significant in the baseline specification. Interestingly, the coefficient estimate for non-CREZ counties is negative but the magnitude is economically and statistically insignificant. Restricting to the counties in matching sample (Panel B.2) flips the pattern with CREZ counties showing a negative anticipation effect and non-CREZ counties showing a positive anticipation effect. However, none of these effects are statistically indistinguishable from zero.

Table C2 reports the results of OLS regression of Equation 24 with [2004, 2008] as the anticipation period. Column (5) and Column (6) are the baseline specifications for the samples using CREZ counties and non-CREZ counties respectively. Similar to Table C1, the coefficient estimates do not reveal any evidence of anticipation effects during the years 2004 to 2008 for both CREZ and non-CREZ counties. Therefore, based on the results from this analysis I rule out the possibility of an anticipation effect in the form of an increase in the number and capacity of wind projects in the ERCOT interconnection queue in the years leading upto CREZ announcement in late 2008.

Table C1: Anticipation of CREZ announcement for the years 2006 to 2008

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent variable: IHS (# projects in interconnection queue)					
	A.1 All counties in Texas					
Year $\in$ [2006, 2008]	0.102*	0.002	0.102*	0.002	0.066	0.002
	(0.054)	(0.008)	(0.055)	(0.009)	(0.059)	(0.010)
Elasticity	0.107	0.002	0.107	0.002	0.068	0.002
R^2	0.016	0.000	0.137	0.116	0.145	0.117
	A.2 Restricting to counties in the matching sample					
Year $\in$ [2006, 2008]	0.004	0.065	0.004	0.065	0.004	0.077*
	(0.054)	(0.044)	(0.055)	(0.045)	(0.056)	(0.044)
Elasticity	0.004	0.067	0.004	0.067	0.004	0.080
R^2	0.000	0.017	0.064	0.081	0.085	0.092
	B. Dependent variable: IHS (Total capacity (MW) in interconnection queue)					
	B.1 All counties in Texas					
Year $\in$ [2006, 2008]	0.454*	−0.021	0.454*	−0.021	0.304	−0.002
	(0.242)	(0.035)	(0.250)	(0.037)	(0.287)	(0.043)
Elasticity	0.575	−0.020	0.575	−0.02	0.356	−0.002
R^2	0.012	0.0001	0.137	0.123	0.145	0.124
	B.2 Restricting to counties in the matching sample					
Year $\in$ [2006, 2008]	−0.018	0.158	−0.018	0.158	−0.013	0.244
	(0.264)	(0.151)	(0.273)	(0.156)	(0.281)	(0.154)
Elasticity	−0.018	0.172	−0.018	0.172	−0.013	0.276
R^2	0.000	0.004	0.072	0.063	0.097	0.079
County FE			✓	✓	✓	✓
Time Trend					✓	✓
Wind Controls					✓	✓
County Controls					✓	✓
Sample	CREZ	non-CREZ	CREZ	non-CREZ	CREZ	non-CREZ

Notes: This table reports the results of regressions analyzing the anticipation effect of CREZ announcement for the years 2006 to 2008. Sample specifies whether the estimation sample is CREZ counties or non-CREZ counties. Panels A.1 and B.1 use all the counties in the data. Total observations in 'CREZ' and 'non-CREZ' Sample in A.1 and B.1 is 585 and 3,225 respectively. Panels A.2 and B.2 restrict the observations to the counties obtained in the matching sample. Total observations in 'CREZ' and 'non-CREZ' Sample in A.2 and B.2 is 195 and 450 respectively. The independent variable is an indicator variable for the years in 2006 to 2008. Time Trend is a linear time trend variable. Wind Controls include capacity factor and cubic polynomial of wind speed. County Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include median land acreage, real land price, indicator for the presence of wind ordinance, average farm size (acres) in 2007, median household income in 2007, and log of population. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table C2: Anticipation of CREZ announcement for the years 2004 to 2008

	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent variable: IHS (# projects in interconnection queue)					
	A.1 All counties in Texas					
Year ∈ [2004, 2008]	0.090*	−0.003	0.090*	−0.003	0.090	−0.003
	(0.046)	(0.006)	(0.048)	(0.006)	(0.065)	(0.009)
Elasticity	0.094	−0.003	0.094	−0.003	0.095	−0.003
R^2	0.017	0.0001	0.138	0.116	0.147	0.117
	A.2 Restricting to counties in the matching sample					
Year ∈ [2004, 2008]	−0.031	0.026	−0.031	0.026	−0.043	0.063
	(0.043)	(0.028)	(0.044)	(0.029)	(0.087)	(0.039)
Elasticity	−0.031	0.026	−0.031	0.026	−0.042	0.065
R^2	0.003	0.004	0.067	0.068	0.087	0.083
	B. Dependent variable: IHS (Total capacity (MW) in interconnection queue)					
	B.1 All counties in Texas					
Year ∈ [2004, 2008]	0.394*	−0.044*	0.394*	−0.044	0.443	−0.016
	(0.215)	(0.027)	(0.222)	(0.027)	(0.319)	(0.040)
Elasticity	0.482	−0.043	0.482	−0.043	0.557	−0.016
R^2	0.013	0.001	0.137	0.124	0.147	0.125
	B.2 Restricting to counties in the matching sample					
Year ∈ [2004, 2008]	−0.194	0.006	−0.194	0.006	−0.264	0.205
	(0.221)	(0.106)	(0.229)	(0.109)	(0.435)	(0.153)
Elasticity	−0.176	0.006	−0.176	0.006	−0.232	0.227
R^2	0.005	0.000	0.077	0.059	0.100	0.075
County FE			✓	✓	✓	✓
Time Trend					✓	✓
Wind Controls					✓	✓
County Controls					✓	✓
Sample	CREZ	non-CREZ	CREZ	non-CREZ	CREZ	non-CREZ

Notes: This table reports the results of regressions analyzing the anticipation effect of CREZ announcement for the years 2004 to 2008. Sample specifies whether the estimation sample is CREZ counties or non-CREZ counties. Panels A.1 and B.1 use all the counties in the data. Total observations in 'CREZ' and 'non-CREZ' Sample in A.1 and B.1 is 585 and 3,225 respectively. Panels A.2 and B.2 restrict the observations to the counties obtained in the matching sample. Total observations in 'CREZ' and 'non-CREZ' Sample in A.2 and B.2 is 195 and 450 respectively. Time Trend is a linear time trend variable. Wind Controls include capacity factor and cubic polynomial of wind speed. County Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include median land acreage, real land price, indicator for the presence of wind ordinance, average farm size (acres) in 2007, median household income in 2007, and log of population. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***$p<0.01$;**$p<0.05$;*$p<0.1$

D Instrumental Variable Strategy

D.1 Identification

In addition to matching, I use an instrumental variables strategy with the designation of counties into five Renewable Energy (RE) Zones in 2007 (z_{i2007}) as the instrument. The IV specification can be written as:

$$crez_i = \rho \cdot z_{i2007} + \mathbf{X}'\Gamma + \nu_i \tag{IV1}$$

$$y_{it} = \alpha + \beta_{IV} \cdot \widehat{crez}_i + \mathbf{X}'\Pi + \eta_{it} \tag{IV2}$$

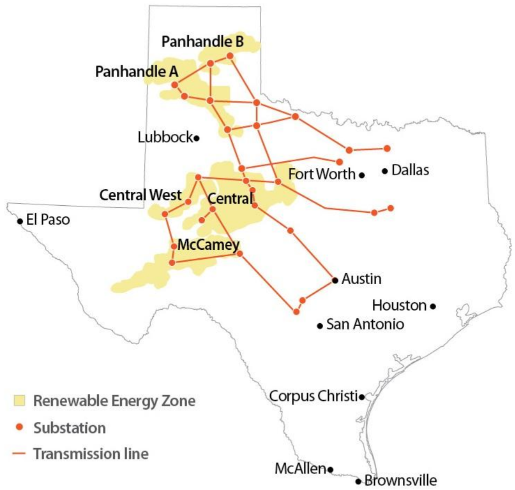
where, Equation (IV1) is the first stage and Equation (IV2) is the second stage equation.

In 2007, the Public Utility Commission of Texas (PUCT or Commission) in its Interim Order on Reconsideration in Docket 33672 designated five zones as Competitive Renewable Energy Zones (PUCT 2007). The five RE Zones - Panhandle A, Panhandle B, Central West, Central, and McCamey are depicted in Figure 15a. These zones were selected based on the existing and potential wind generation capacity, and wind developer interest in these areas. Thus, one of the objectives of the CREZ project was to integrate the wind capacity in these zones to the power grid.

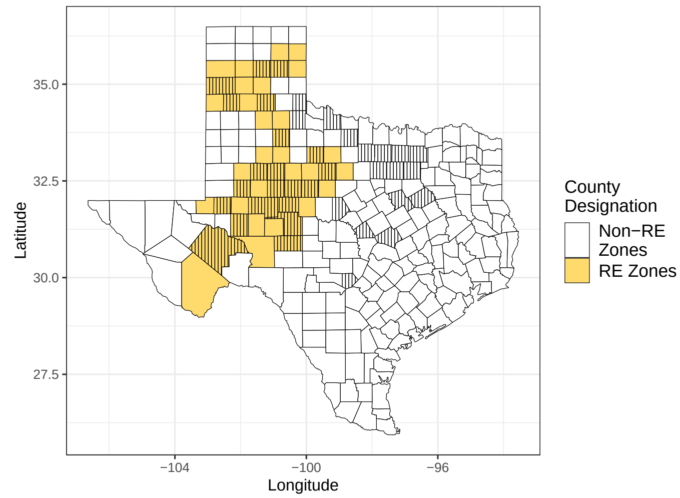
From Figure 15a, I identify the counties that were part of the five RE Zones and denote them as z_{i2007}. These counties are highlighted in yellow in Figure 15b. In April 2008, ERCOT released the CREZ Transmission Optimization Study which analyzed various scenarios for the placement of CREZ transmission lines, substations, and wind collection points within the five RE Zones(ERCOT 2008). Ultimately, scenario 2 was selected by the PUCT as the optimal scenario for CREZ Project. Over the next few years, exact locations of transmission infrastructure was determined based on cost considerations, inputs from county legislatures and wind developers. The hatched counties in Figure 15b are the counties that sited CREZ infrastructure - substations and wind collection points.

As evident from the Figure 15b, there is a significant overlap between the counties that were part of the RE Zones (z_{i2007}) and the CREZ counties $(crez_i)$. Therefore, using z_{i2007} as an instrument for $crez_i$ validates the relevance condition:

$$\mathbb{E}[z_{i2007}, crez_i|\mathbf{X}] \neq 0 \tag{25}$$



(a) RE Zones designation in 2007



(b) Overlap between counties in RE Zones and counties with CREZ substations (hatched).

Figure 15: Designation of counties in Renewable Energy (RE) Zones in 2007 and selection of counties siting CREZ substations in August 2008.

In words, conditional on the control variables $\mathbf{X}$, there is a non-zero correlation between a county being designated as a part of the RE Zones and siting transmission infrastructure. Another key aspect of the validity of the instrument z_{i2007} is the exclusion restriction:

$$\mathbb{E}[z_{i2007}, \eta_{it}|\mathbf{X}] = 0 \tag{26}$$

Equation 26 refers to the exogeneity of the instrument to the dependent variable y_{it}. In other words, it implies that conditional on $\mathbf{X}$, investment in wind energy in a county from 2012-2019 is only driven by its designation to one of the five RE Zones through its effect on the likelihood of that county being selected as a CREZ county.

The exclusion restriction is likely to be satisfied because selection of a site for wind development is a function of numerous factors other than whether it was designated as a CREZ zone. These factors include wind resource quality, location specific costs, wind siting regulations, availability of transmission, and support of the local community. The variables in $\mathbf{X}$ attempt to control for these confounding factors. Further, the main reason for RE Zones was to designate general areas with historically high wind quality where transmission expansion was needed. This was useful in narrowing ERCOT's choice set of identifying possible sites for transmission expansion.

D.2 Results

Table 3 reports the results for the OLS and 2SLS regressions of the baseline specification with total nameplate capacity (MW), total turbines, and average capacity per project as the outcome variables respectively. These regressions use the full set of control variables, i.e. time trend, vector of controls for wind resource quality, project cost, county level regulation, demographics, and fixed effects for zone and PTC.

The results for total nameplate capacity suggest a significant increase in wind capacity in CREZ counties. The OLS specification in Column (1) shows that CREZ counties saw an average 43 MW increase in wind capacity than non-CREZ counties. The first stage F-Statistic value for 2SLS regression suggests a strong correlation of the instrument in the First Stage. The coefficient estimate in Column (2) shows that after accounting for confounding factors, CREZ counties have a 62 MW higher wind capacity than non-CREZ counties. The story is the same when total turbines in a county is used as dependent variable. For the 2SLS estimate in Column (4), we see that on an average, CREZ counties have 30 more turbines than non-CREZ counties. All of these results are statistically significant at 5 percent critical level.

Columns (5) and (6) examine whether there is a difference in size of a wind project between CREZ v.s. a non-CREZ county? Controlling for confounding factors, we expect wind developers to build bigger wind projects in sites that have excess transmission capacity and therefore a positive coefficient. However, I find weak evidence in support for this hypothesis. The OLS specification in Column (5) shows that an average project in a CREZ county is about 11 MW bigger than a project in a non-CREZ county. The 2SLS estimate finds this difference to be about 14 MW. However, only the OLS estimate is statistically indistinguishable from a null effect.

D.2.1 Robustness Checks and Caveats

I conduct a series of robustness checks to see how the coefficient estimate of CREZ changes with different combinations of control variables and fixed effects for the OLS and 2SLS specifications. I also restrict the sample to counties in West, Panhandle, and North Zone to discern whether these results are driven primarily by investment in counties within these zones. The results for all the three dependent variables are similar to the estimates reported in Table 3. Appendix F presents the tables for these robustness checks.

Note that the 2SLS estimates in Table 3 are the Local Average Treatment Effect (LATE) of CREZ or the Average Treatment Effect (ATE) for the sub-population of compiler counties. The compiler counties in this case would be the counties that were selected as CREZ because they were part of the RE Zones in 2007, or the counties that were not selected as CREZ because they were not part of the RE Zones. These estimates do not speak to the wind investment in CREZ counties that were selected as CREZ (i.e. $crez_i = 1$) even though they were not part of the RE Zones or the counties that were not selected as CREZ even though they were part of the RE Zones.

Another concern with the instrumental variable strategy is that the counties within the RE Zones were inherently different than the other counties. Table F2 in Appendix shows that counties within RE Zones were statistically different than the other counties on almost all of the observable characteristics pre-transmission expansion announcement. I attempt to account for these factors by including a rich array of control variables that condition on these inherent differences in the counties. However, the concern of a lack of common support amongst the treated and control counties in the 2SLS regressions still remains. Thus, I further address these concerns by implementing a matching estimator which constructs a control group of identical counties based on observable characteristics.

Table 3: Effect of CREZ expansion on wind investment - IV results

	Dependent variable					
	Total Nameplate Capacity (MW)		Total Turbines		Avg. Project Capacity (MW)	
	(1)	(2)	(3)	(4)	(5)	(6)
CREZ	43.041***	62.181**	23.358***	30.285***	10.620**	14.089
	(10.051)	(26.786)	(5.149)	(13.596)	(4.745)	(16.136)
Time Trend	✓	✓	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓	✓	✓
Wind Controls	✓	✓	✓	✓	✓	✓
Project Cost Controls	✓	✓	✓	✓	✓	✓
Regulatory Controls	✓	✓	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓	✓	✓
Regression	OLS	2SLS	OLS	2SLS	OLS	2SLS
First Stage F-Stat		12.842		12.842		12.842
Mean Dep. Variable	32.939	32.939	15.866	15.866	19.911	19.911
Observations	2,024	2,024	2,024	2,024	2,024	2,024
R^2	0.218	0.216	0.206	0.205	0.196	0.195

Notes: This table reports the results of OLS and 2SLS regressions for the effect of CREZ expansion on wind investment. The independent variable is a binary variable indicating whether a county is CREZ or not. The instrument for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of time trend variable. Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***$p<0.01$;**$p<0.05$;*$p<0.1$

E Supplementary Figures

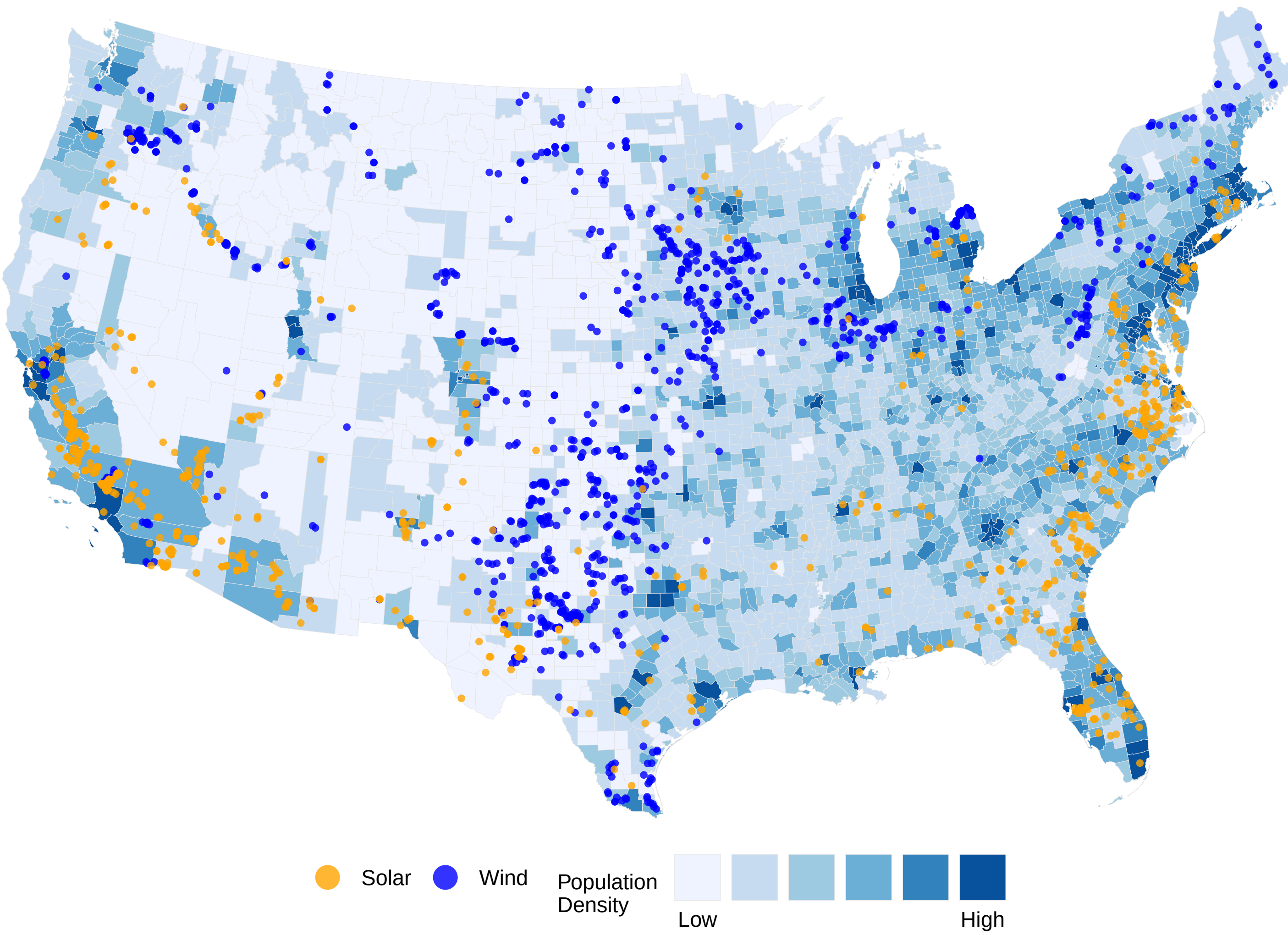
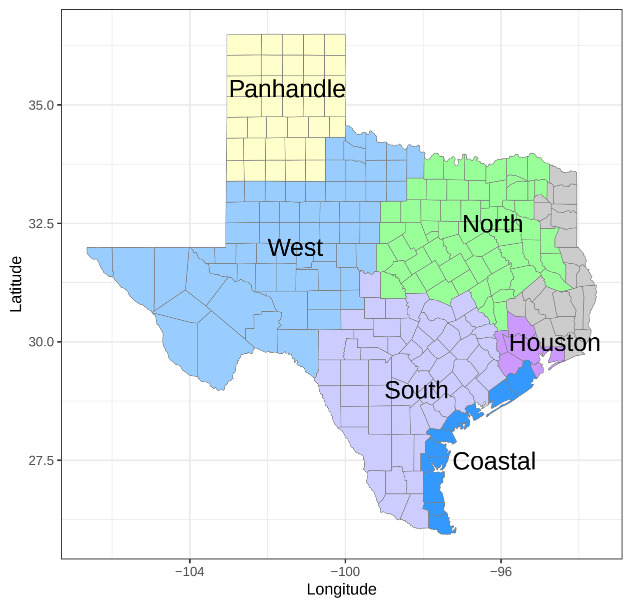
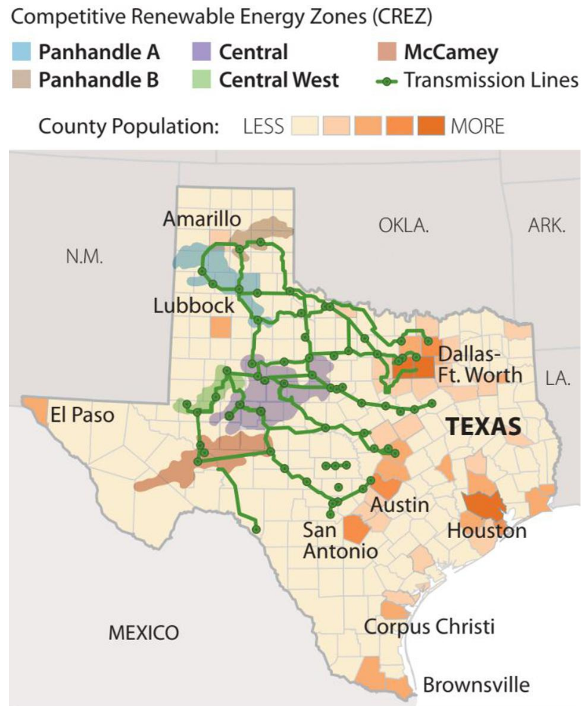


Figure E1: All solar and wind projects with $\geq$ 10 MW of nameplate capacity that started operation post 2001.

Note: The county level population density is based on 2014 data from US Census Bureau. Population density bins are: [0, 10], (10, 50], (50, 100], (100, 500], (500, 1,000], (1000, 72,000].



(a) ERCOT Zones Map



(b) CREZ lines and county level population. Source: US EIA

Figure E2: ERCOT Zones and CREZ transmission expansion

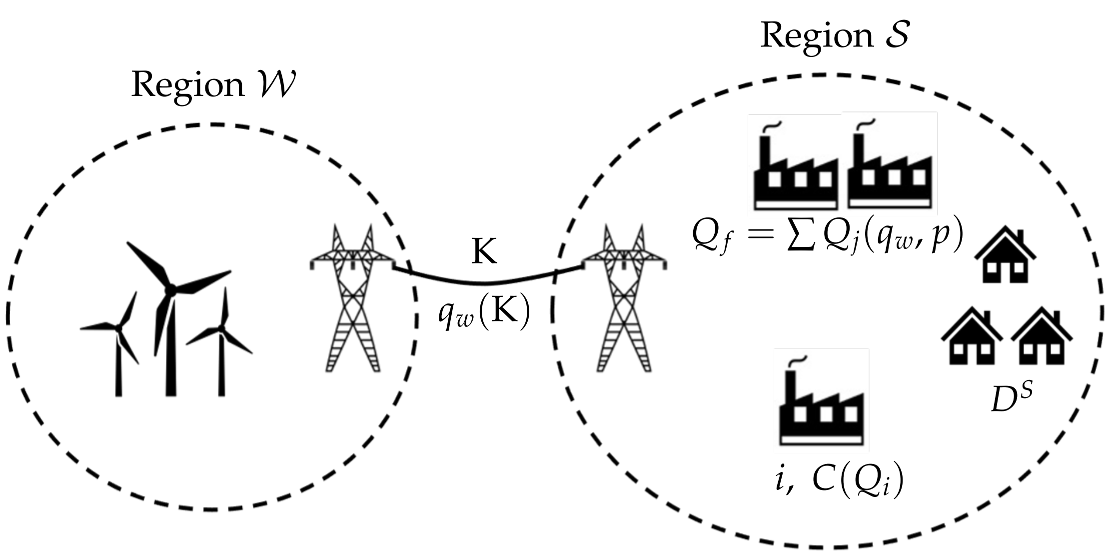


Figure E3: Theory model setup

Note: K denotes the transmission capacity between Regions $\mathcal{W}$ and $\mathcal{S}$, $q_w(\text{K})$ is the amount of wind generation transported into region $\mathcal{S}$. D^S is the inelastic demand for electricity, $C(Q_i)$ is generator i's cost of generating electricity, and Q_f is the total electricity generated by other fossil fuel generator's.

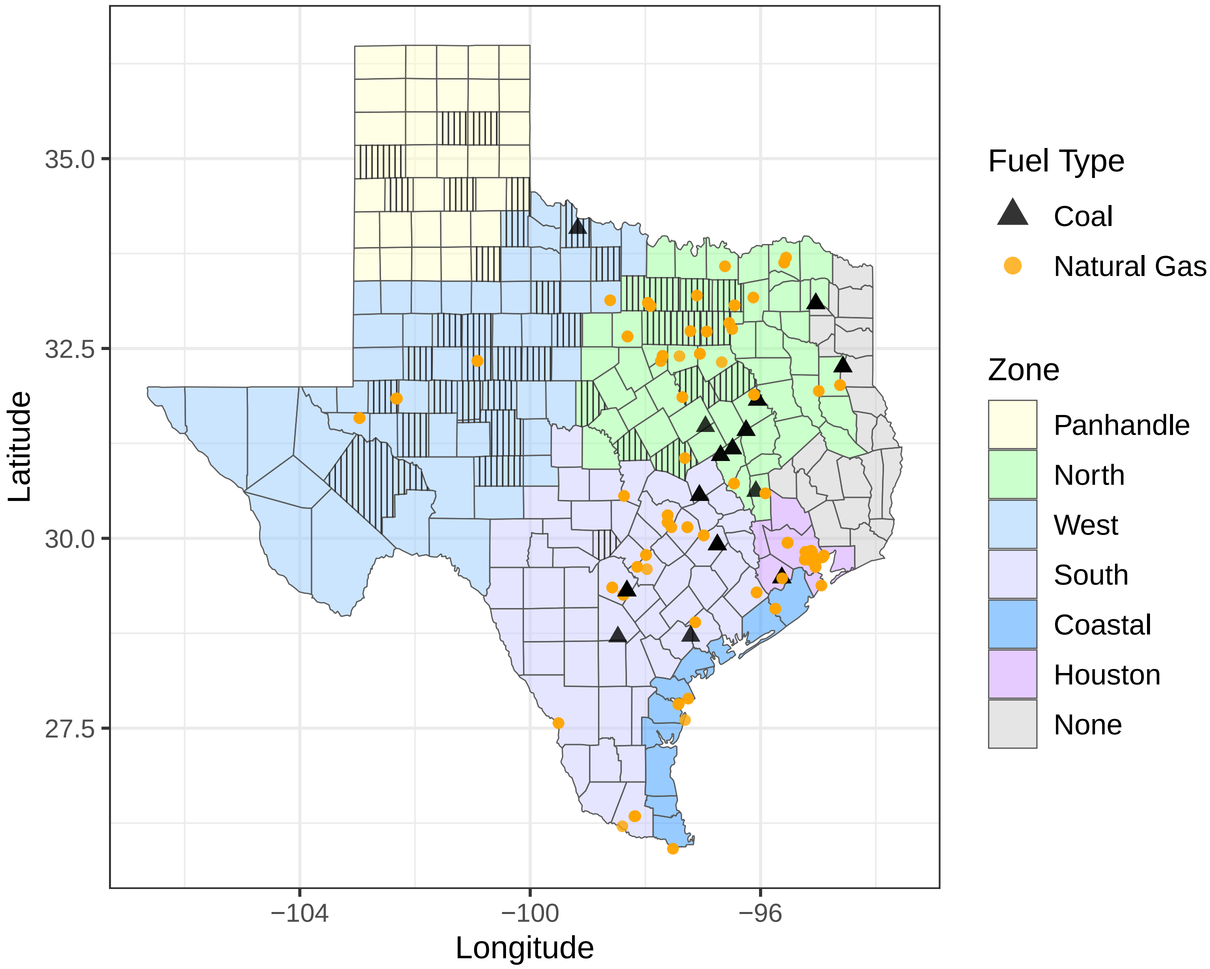
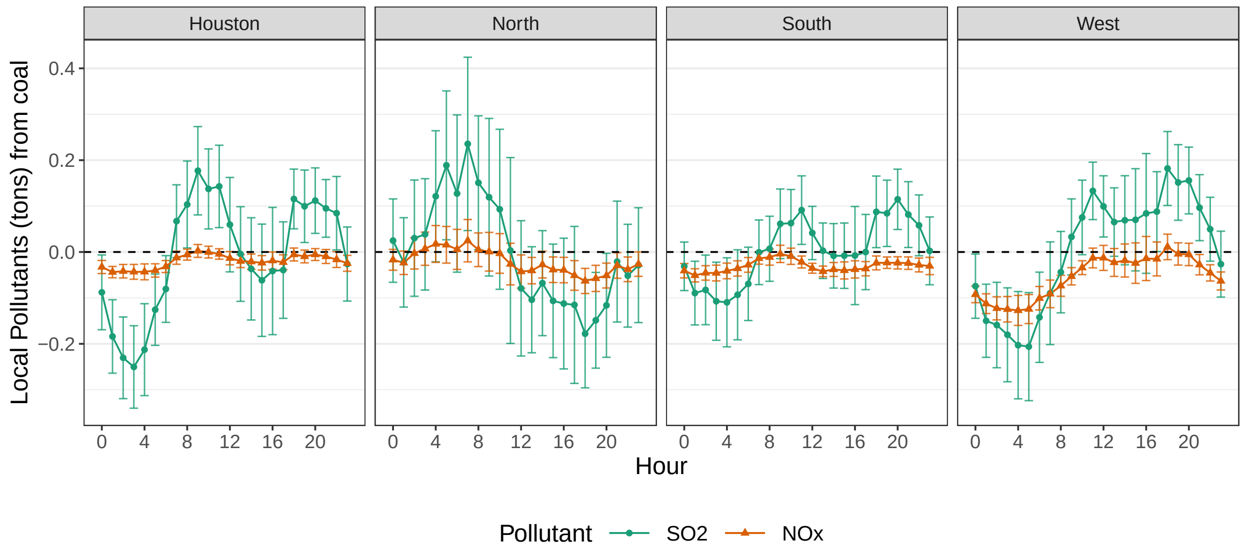
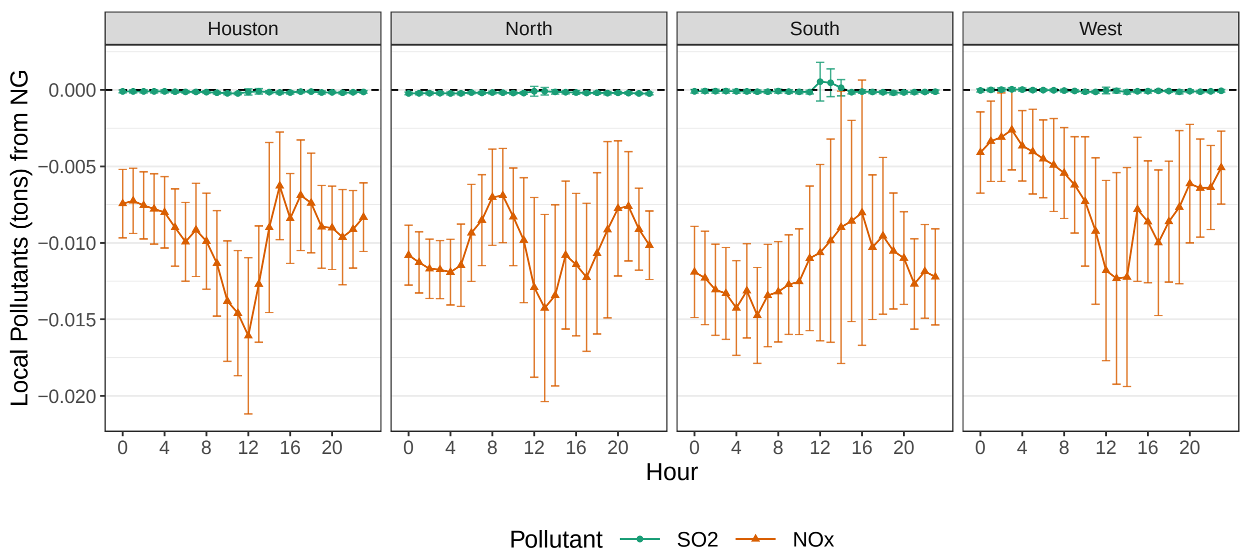


Figure E4: Coal and natural gas fired generators operating at the margin in sample from 2011 - 2014. Hatched counties denote the counties that received CREZ transmission expansion.

E.1 Wind generation and local pollution from marginal generators



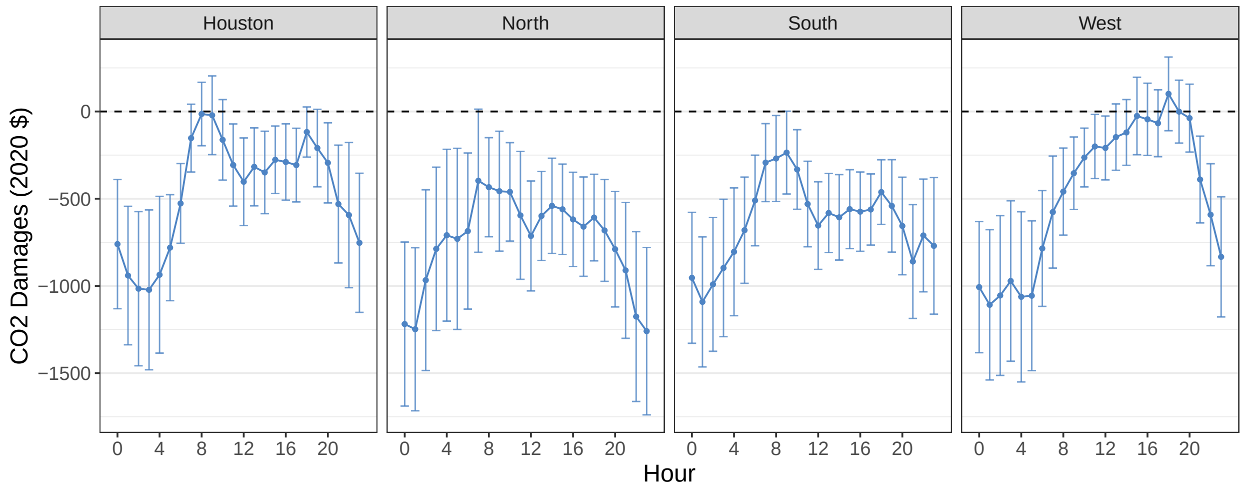
(a) Impact of wind generation on local pollutants (SO_2 and NOx) from coal generators



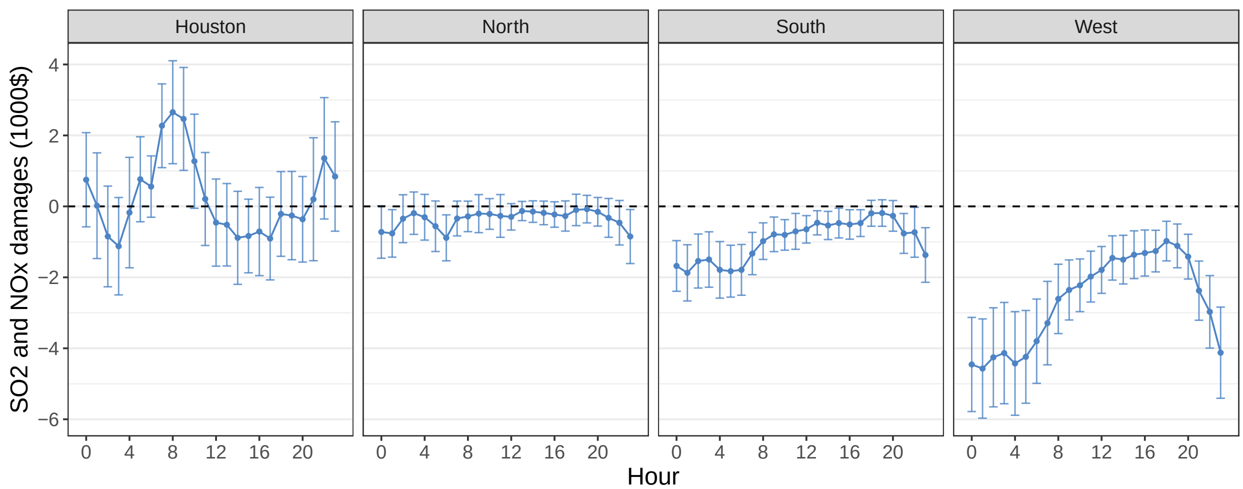
(b) Impact of wind generation on local pollutants (SO_2 and NOx) from natural gas generators

Figure E5: Short-run impact of wind generation on local pollutants (SO_2 and NOx) by generator type

E.2 Total damages from CO_2 and local pollutants for each hour



(a) Damages due to global pollution (CO_2)



(b) Damages due to local pollution (SO_2 and NOx)

Figure E6: Hourly averages of the marginal damages (2020 $) avoided due to CREZ expansion for each zone over 2011 - 2014.

E.3 Existence of multi-phase wind projects and project extensions

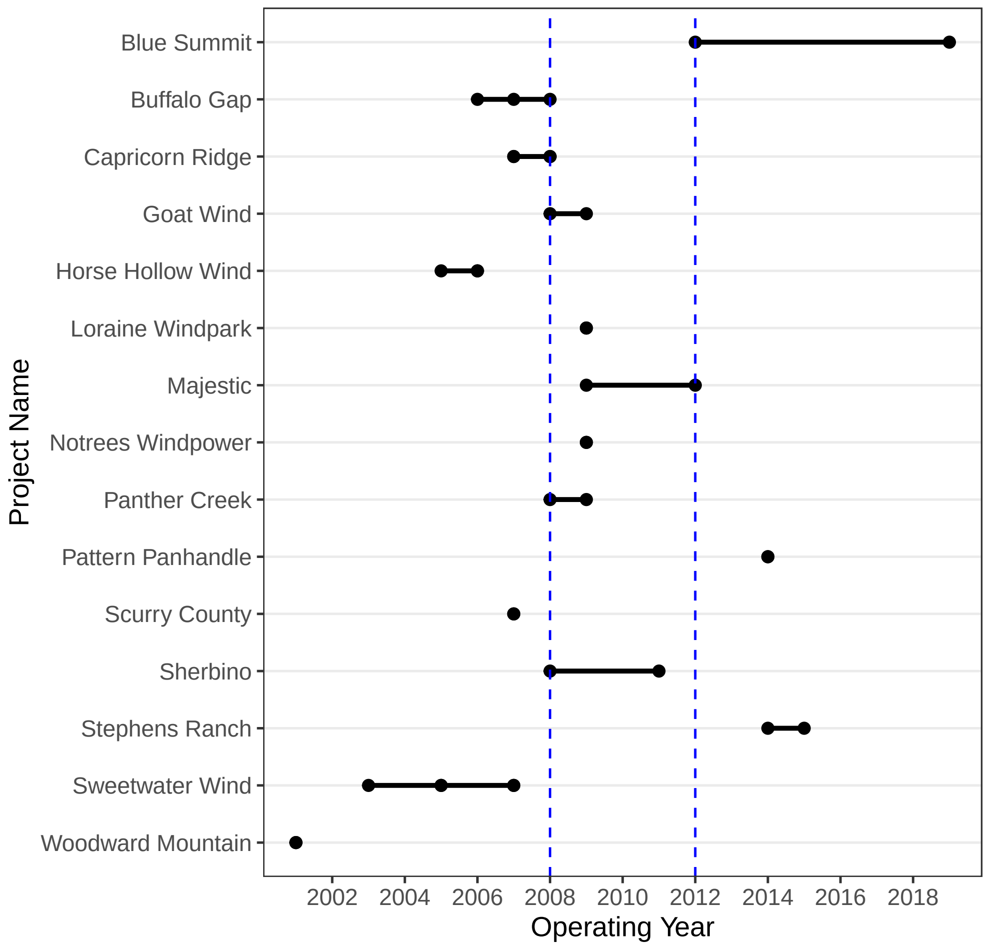


Figure E7: Wind projects with multiple phases and extensions

Note: This figure presents projects with multiple phases or extensions within CREZ counties. Each dot represents at least one phase. Projects with single dots (Loraine Windpark, Notrees Windpower, Pattern Panhandle, Scurry County, and Woodward Mountain) have multiple phases completed in the same year. There are 37 individual projects within 15 "main projects" shown in this figure. The selection issue arises if a line segment intersects both the dotted vertical lines for the years 2008 and 2012. From the figure, we do not see any instance of such a situation. However, wind projects under Majestic and Sherbino warrant more attention. The first phase of Majestic was completed in 2009 and the second one was completed in 2012. This is not a cause of concern since the first phase started operating post CREZ announcement in 2008 and only the second phase is counted in the dependent variable(s). In case of Sherbino, although the first phase was completed in 2008, the second phase was completed in 2011 and is therefore not included in the dependent variable(s).

F Supplementary Tables

F.1 Descriptive Statistics

Table F1: Descriptive statistics of key variables used in the short-run analysis of the effect of CREZ on fossil fuel generator markups

	2011		2012		2013		2014	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Marginal Cost ($/MWh)	17.19	14.13	12.97	12.16	18.48	17.88	20.71	19.39
Markups ($/MWh)	12.20	33.10	12.76	41.79	12.15	76.53	12.59	36.09
Load (GWh)	35.30	8.24	36.16	7.59	37.02	7.62	38.05	7.93
Wind Generation (GWh)	2.47	1.59	2.96	1.73	3.30	1.95	3.88	2.43
CREZ progress	0.06	0.01	0.11	0.06	0.61	0.22	1	0

This table reports the annual means and standard deviation (SD) of key variables used in the short-run analysis in Section 4. Note that marginal cost and markups is at the hourly generator level, load and wind generation are the hourly system level. CREZ progress is the ratio of cumulative addition of transmission lines (at the daily level) to total length of transmission lines built in CREZ.

Table F2: Difference in means test of key observable characteristics of counties within and outside Renewable Energy (RE) Zones

Variable	Not in RE Zones ($z_{i2007} = 0$)	Within RE Zones ($z_{i2007} = 1$)	p-value
Wind Resource Quality			
Wind Speed (m/s)	7.281	8.085	0.000
Capacity Factor	0.413	0.442	0.000
Power Curve	2.672	2.061	0.000
Project Cost Variables			
Median Land Acreage (acres)	724.720	993.122	0.038
Real Land Price	313.567	119.592	0.000
Avg. Annual Project Cost	1,759.884	1,759.884	1.000
Wind Regulation			
Wind Ordinance	0	0	–
Demographic Variables			
Avg. farm Size 2007	1,487.549	2,607.163	0.000
Population	105,193.300	20,974.870	0.000
Median Income in 2007	40,461.870	39,435.020	0.011

This table reports the difference in means test of key observable characteristics of counties within ($z_{i2007} = 1$) and outside ($z_{i2007} = 0$) the five Renewable Energy (RE) Zones. To highlight the pre-existing differences in the counties, all the observations are pre-transmission expansion announcement. Avg. Annual Project Cost is the capacity weighted average project cost in $/kW and is common across all counties for a given year. Wind Ordinance is a dummy equal to 1 if a county had a wind ordinance in that year.

F.2 Robustness check results for markup analysis

Table F3: Effect of integration of 1 GWh of wind energy on fossil fuel generator markups ($/MWh)

	Dependent variable: Hourly markups ($/MWh)		
	(1)	(2)	(3)
Wind Generation (GWh)			
$\times$ 𝟙{hour = 0}	-1.875 $(0.087)^{***}$	-1.594 $(0.071)^{***}$	-1.487 $(0.068)^{***}$
$\times$ 𝟙{hour = 1}	-2.211 $(0.096)^{***}$	-1.831 $(0.080)^{***}$	-1.139 $(0.055)^{***}$
$\times$ 𝟙{hour = 2}	-2.449 $(0.101)^{***}$	-2.022 $(0.086)^{***}$	-1.130 $(0.061)^{***}$
$\times$ 𝟙{hour = 3}	-2.532 $(0.103)^{***}$	-2.078 $(0.089)^{***}$	-1.112 $(0.063)^{***}$
$\times$ 𝟙{hour = 4}	-2.469 $(0.106)^{***}$	-2.001 $(0.089)^{***}$	-1.063 $(0.063)^{***}$
$\times$ 𝟙{hour = 5}	-2.252 $(0.102)^{***}$	-1.767 $(0.082)^{***}$	-1.315 $(0.071)^{***}$
$\times$ 𝟙{hour = 6}	-1.363 $(0.104)^{***}$	-0.883 $(0.085)^{***}$	-3.583 $(0.249)^{***}$
$\times$ 𝟙{hour = 7}	-1.822 $(0.093)^{***}$	-1.396 $(0.071)^{***}$	-1.309 $(0.076)^{***}$
$\times$ 𝟙{hour = 8}	-1.786 $(0.095)^{***}$	-1.389 $(0.074)^{***}$	-1.233 $(0.095)^{***}$
$\times$ 𝟙{hour = 9}	-1.546 $(0.094)^{***}$	-1.278 $(0.079)^{***}$	-1.657 $(0.078)^{***}$
$\times$ 𝟙{hour = 10}	-1.364 $(0.098)^{***}$	-1.288 $(0.080)^{***}$	-1.640 $(0.066)^{***}$
$\times$ 𝟙{hour = 11}	-1.474 $(0.110)^{***}$	-1.519 $(0.099)^{***}$	-1.447 $(0.069)^{***}$
$\times$ 𝟙{hour = 12}	-1.340 $(0.126)^{***}$	-1.460 $(0.122)^{***}$	-1.551 $(0.107)^{***}$
$\times$ 𝟙{hour = 13}	-1.380 $(0.119)^{***}$	-1.553 $(0.122)^{***}$	-1.799 $(0.131)^{***}$
$\times$ 𝟙{hour = 14}	-1.308 $(0.118)^{***}$	-1.655 $(0.137)^{***}$	-2.452 $(0.222)^{***}$
$\times$ 𝟙{hour = 15}	-0.932 $(0.107)^{***}$	-1.380 $(0.123)^{***}$	-3.329 $(0.297)^{***}$
$\times$ 𝟙{hour = 16}	0.018 (0.144)	-0.374 $(0.125)^{***}$	-8.639 $(0.882)^{***}$
$\times$ 𝟙{hour = 17}	-0.527 $(0.101)^{***}$	-1.117 $(0.132)^{***}$	-3.571 $(0.244)^{***}$
$\times$ 𝟙{hour = 18}	-0.605 $(0.102)^{***}$	-1.383 $(0.134)^{***}$	-2.621 $(0.146)^{***}$
$\times$ 𝟙{hour = 19}	-1.206 $(0.099)^{***}$	-1.781 $(0.113)^{***}$	-1.631 $(0.063)^{***}$
$\times$ 𝟙{hour = 20}	-1.157 $(0.093)^{***}$	-1.607 $(0.098)^{***}$	-1.620 $(0.063)^{***}$
$\times$ 𝟙{hour = 21}	-1.284 $(0.099)^{***}$	-1.721 $(0.103)^{***}$	-1.110 $(0.079)^{***}$
$\times$ 𝟙{hour = 22}	-1.375 $(0.089)^{***}$	-1.519 $(0.084)^{***}$	-1.576 $(0.074)^{***}$
$\times$ 𝟙{hour = 23}	-1.668 $(0.090)^{***}$	-1.579 $(0.079)^{***}$	-1.500 $(0.069)^{***}$
Generator FE	✓	✓	✓
Load and Load^2		✓	✓
Hour $\times$ Month $\times$ Year FE			✓
Number of FE	284	284	1244
Observations	619,864	619,864	619,864
R^2	0.141	0.152	0.253

This table reports the coefficient estimates of Equation 13. The dependent variable is the markup ($/MWh) set by the marginal generator i at hour t. The coefficient of interest are the interaction of hourly wind generation (GWh) and indicator of the hour. All regressions use hourly data from August 2011 to December 2014. Standard Errors clustered by generator reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table F4: Effect of CREZ completion on hourly wind generation (GWh)

	Dependent variable: Hourly Wind Generation (GWh)		
	(1)	(2)	(3)
crez			
$\times\ \mathbb{1}\{\text{hour} = 0\}$	1.722 (0.206)***	0.170 (0.024)***	0.228 (0.027)***
$\times\ \mathbb{1}\{\text{hour} = 1\}$	1.706 (0.207)***	0.165 (0.024)***	0.221 (0.027)***
$\times\ \mathbb{1}\{\text{hour} = 2\}$	1.621 (0.208)***	0.160 (0.025)***	0.213 (0.028)***
$\times\ \mathbb{1}\{\text{hour} = 3\}$	1.489 (0.207)***	0.156 (0.024)***	0.202 (0.028)***
$\times\ \mathbb{1}\{\text{hour} = 4\}$	1.336 (0.209)***	0.166 (0.023)***	0.200 (0.025)***
$\times\ \mathbb{1}\{\text{hour} = 5\}$	1.163 (0.208)***	0.171 (0.022)***	0.171 (0.021)***
$\times\ \mathbb{1}\{\text{hour} = 6\}$	0.992 (0.207)***	0.169 (0.021)***	0.147 (0.020)***
$\times\ \mathbb{1}\{\text{hour} = 7\}$	0.827 (0.207)***	0.169 (0.020)***	0.142 (0.020)***
$\times\ \mathbb{1}\{\text{hour} = 8\}$	0.505 (0.210)***	0.161 (0.019)***	0.126 (0.018)***
$\times\ \mathbb{1}\{\text{hour} = 9\}$	0.286 (0.213)**	0.151 (0.018)***	0.127 (0.018)***
$\times\ \mathbb{1}\{\text{hour} = 10\}$	0.216 (0.215)*	0.146 (0.018)***	0.128 (0.019)***
$\times\ \mathbb{1}\{\text{hour} = 11\}$	0.047 (0.214)	0.138 (0.018)***	0.123 (0.018)***
$\times\ \mathbb{1}\{\text{hour} = 12\}$	−0.113 (0.211)	0.130 (0.017)***	0.115 (0.018)***
$\times\ \mathbb{1}\{\text{hour} = 13\}$	−0.193 (0.207)	0.127 (0.017)***	0.107 (0.017)***
$\times\ \mathbb{1}\{\text{hour} = 14\}$	−0.147 (0.203)	0.129 (0.017)***	0.106 (0.017)***
$\times\ \mathbb{1}\{\text{hour} = 15\}$	−0.059 (0.200)	0.131 (0.017)***	0.102 (0.016)***
$\times\ \mathbb{1}\{\text{hour} = 16\}$	0.038 (0.200)	0.134 (0.017)***	0.102 (0.016)***
$\times\ \mathbb{1}\{\text{hour} = 17\}$	0.105 (0.201)	0.138 (0.018)***	0.099 (0.015)***
$\times\ \mathbb{1}\{\text{hour} = 18\}$	0.313 (0.206)**	0.149 (0.018)***	0.104 (0.016)***
$\times\ \mathbb{1}\{\text{hour} = 19\}$	0.597 (0.205)***	0.160 (0.019)***	0.124 (0.018)***
$\times\ \mathbb{1}\{\text{hour} = 20\}$	0.844 (0.198)***	0.168 (0.020)***	0.145 (0.020)***
$\times\ \mathbb{1}\{\text{hour} = 21\}$	1.177 (0.196)***	0.179 (0.021)***	0.179 (0.023)***
$\times\ \mathbb{1}\{\text{hour} = 22\}$	1.513 (0.199)***	0.188 (0.023)***	0.207 (0.025)***
$\times\ \mathbb{1}\{\text{hour} = 23\}$	1.676 (0.204)***	0.182 (0.023)***	0.224 (0.026)***
High System Limit		✓	✓
Hour × Month FE			✓
Observations	29,205	29,205	29,205
R^2	0.051	0.992	0.992

This table reports the coefficient estimates of Equation 14. The dependent variable is total wind generation (GWh) at hour t. The coefficient of interest are the interaction of CREZ progress (*crez*) and indicator of the hour. All regressions use hourly data from August 2011 to December 2014. Newey-West Autocorrelation corrected standard errors with a 7 day lag structures reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

F.3 Robustness check results for OLS and IV regressions

Table F5: Effect of CREZ expansion on total wind capacity (MW) - IV results

	Dependent variable: Total Nameplate Capacity (MW)			
	(1)	(2)	(3)	(4)
CREZ	51.429***	43.041***	124.817***	62.181**
	(10.293)	(10.051)	(26.919)	(26.786)
Time Trend	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓
Wind Controls		✓		✓
Project Cost Controls		✓		✓
Regulatory Controls		✓		✓
Demographic Controls		✓		✓
Regression	OLS	OLS	2SLS	2SLS
First Stage F-Stat			15.838	12.842
Mean Dep. Variable	32.939	32.939	32.939	32.939
Observations	2,032	2,024	2,032	2,024
R^2	0.137	0.218	0.088	0.216

Notes: Columns (1) and (2) show OLS estimation results for Equation 22 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is total nameplate capacity (MW) of wind projects in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table F6: Effect of CREZ expansion on total turbines in a county - IV results

	Dependent variable: Total Turbines in a County			
	(1)	(2)	(3)	(4)
CREZ	27.550***	23.358***	59.022***	30.285**
	(5.313)	(5.149)	(13.382)	(13.596)
Time Trend	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓
Wind Controls		✓		✓
Project Cost Controls		✓		✓
Regulatory Controls		✓		✓
Demographic Controls		✓		✓
Regression	OLS	OLS	2SLS	2SLS
First Stage F-Stat			15.838	12.842
Mean Dep. Variable	15.866	15.866	15.866	15.866
Observations	2,032	2,024	2,032	2,024
R^2	0.136	0.206	0.098	0.205

Notes: Columns (1) and (2) show OLS estimation results for Equation 22 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is the total wind turbines in a county in year *t*. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table F7: Effect of CREZ expansion on average capacity of a wind project - IV results

	Dependent variable: Average project capacity (MW)			
	(1)	(2)	(3)	(4)
CREZ	16.252***	10.620**	43.684***	14.089
	(4.368)	(4.745)	(14.172)	(16.136)
Time Trend	✓	✓	✓	✓
Zone FE	✓	✓	✓	✓
Wind Controls		✓		✓
Project Cost Controls		✓		✓
Regulatory Controls		✓		✓
Demographic Controls		✓		✓
Regression	OLS	OLS	2SLS	2SLS
First Stage F-Stat			15.838	12.842
Mean Dep. Variable	19.911	19.911	19.911	19.911
Observations	2,032	2,024	2,032	2,024
R^2	0.122	0.196	0.097	0.195

Notes: Columns (1) and (2) show OLS estimation results for Equation 22 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is the total wind turbines in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

F.4 Robustness check results for matching regressions

Table F8: Effect of CREZ on total wind capacity (MW)

	Dependent variable: Total Nameplate Capacity (MW)			
	(1)	(2)	(3)	(4)
CREZ	43.041*	60.425**	71.670***	72.640***
	(22.676)	(27.899)	(26.194)	(26.499)
Time Trend	✓	✓	✓	✓
Wind Controls	✓	✓	✓	✓
Project Cost Controls	✓	✓	✓	✓
Regulatory Controls	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓
FE	Zone	Group	Group	Group × Trend
Sample	OLS	Matched	Matched	Matched
CEM Weights			✓	✓
Mean Dep. Variable	33.069	35.907	35.907	35.907
Observations	2,024	344	344	344
R^2	0.218	0.339	0.390	0.467

Notes: The dependent variable is total turbines in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. OLS Sample is a balanced panel of 253 Texas counties from 2012-2019. Matched Sample is a balanced panel of 13 treated (CREZ) and 30 control (non-CREZ) counties from 2012-2019 obtained using CEM. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed. Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table F9: Effect of CREZ on Total Turbines in a County

	Dependent variable: Total Turbines in a County			
	(1)	(2)	(3)	(4)
CREZ	23.358**	34.451**	39.826**	39.419***
	(11.855)	(15.718)	(15.467)	(13.075)
Time Trend	✓	✓	✓	✓
Wind Controls	✓	✓	✓	✓
Project Cost Controls	✓	✓	✓	✓
Regulatory Controls	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓
FE	Zone	Group	Group	Group × Trend
Sample	OLS	Matched	Matched	Matched
CEM Weights			✓	✓
Mean Dep. Variable	15.928	16.067	16.067	16.067
Observations	2,024	344	344	344
R^2	0.206	0.347	0.408	0.476

Notes: The dependent variable is total turbines in a county in year *t*. The independent variable is a binary variable indicating whether a county is CREZ or not. OLS Sample is a balanced panel of 253 Texas counties from 2012-2019. Matched Sample is a balanced panel of 13 treated (CREZ) and 30 control (non-CREZ) counties from 2012-2019 obtained using CEM. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed. Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table F10: Effect of CREZ on Average Capacity (MW) of a wind project in a County

	Dependent variable: Average Capacity (MW) of a project			
	(1)	(2)	(3)	(4)
CREZ	10.620*	26.500	32.722*	32.756*
	(10.721)	(19.046)	(18.832)	(19.093)
Time Trend	✓	✓	✓	✓
Wind Controls	✓	✓	✓	✓
Project Cost Controls	✓	✓	✓	✓
Regulatory Controls	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓
FE	Zone	Group	Group	Group $\times$ Trend
Sample	OLS	Matched	Matched	Matched
CEM Weights			✓	✓
Mean Dep. Variable	19.99	26.951	26.951	26.951
Observations	2,024	344	344	344
R^2	0.196	0.313	0.345	0.426

Notes: The dependent variable is the average capacity of a wind project in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. OLS Sample is a balanced panel of 253 Texas counties from 2012-2019. Matched Sample is a balanced panel of 13 treated (CREZ) and 30 control (non-CREZ) counties from 2012-2019 obtained using CEM. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed. Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1