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

Gaurav Doshi†

Job Market Paper

This version: October 10, 2022

Click here for the most recent version

Abstract

Expanding the power grid is a key factor in achieving decarbonization and fully utilizing the benefits of renewable energy. This paper examines the impact of large-scale grid expansion on price-cost markups and emissions from fossil fuel generators in the short run and wind investment in the long run. I focus on the rollout of a grid expansion project that linked windy areas in West Texas to population centers in the east. Results suggest moderate declines in markups and emissions, with total annual benefits of roughly $285 million. Counties that received investment in transmission infrastructure saw higher wind capacity in the long run, avoiding $271 million worth of carbon emissions in 2019. These findings highlight the potential to unlock large economic benefits from transmission expansion across the US.

JEL Classifications: L11, Q40, Q41, Q53.

Keywords: Electricity Markets, Emissions, Market Power, 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 greatly benefited from comments and suggestions from Paul Dower, Joseph Eggers, Corbett Grainger, Varan Kitayaporn, Jacob LaRiviere, Arik Levinson, Itzel De Haro López, Laura Schechter, Andrew Stevens, Matt Woerman, and seminar participants at the CU Environmental and Resource Economics workshop, Camp Resources, ASSA Annual Meeting, AERE Summer Conference, EMEE Summer Workshop, OS-WEET, and UW Madison. I am also grateful to Ted Hailu and ERCOT Client Services for their assistance in accessing and using the data from ERCOT. All errors are my own.

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

1 Introduction

The US is aiming toward rapid electrification and decarbonization in the coming decades to combat climate change (National Academies of Sciences, Engineering, and Medicine 2021). A critical factor in achieving these goals is massive investment in expanding the power grid. Because most wind and solar farms in the US are located far from demand centers, high-capacity transmission lines are necessary to move this electricity over long distances. Thus, investment in transmission lines is crucial to fully realize the benefits of renewable energy and achieve ambitious energy policy targets.[1]

Inadequate transmission capacity impedes the integration of electricity from renewable sources and 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 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 and emissions from fossil fuel generators. The main innovation of my approach is to provide an empirical framework to study both the market and non-market impacts of increased transmission capacity, with the advantage of comparing the potential benefits from both channels.

Grid expansion can also speed up the transition to green energy in the long run. Ignoring these effects understates the economic benefits of transmission expansion. However, any analysis to quantify this response is complicated due to endogeneity from non-random siting of electricity transmission. I provide one of the first causal estimates in the economics literature on the magnitude of long-term investment in wind energy in response to transmission expansion. To do so, I exploit rich spatial and temporal

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 ambitious energy goals (New York Times 2016; Temple 2017; Meyer 2021). Figure E1 shows that most utility-scale solar and wind projects ($\geq$10 MW) in the US are located far from densely populated counties.

data from the rollout of a large-scale transmission expansion project called Competitive Renewable Energy Zones (CREZ) in Texas.

For the short-run analysis, I construct a model of optimal bidding to understand how transmission expansion affects a marginal fossil fuel generator's incentives in setting markups. This model is most closely related to the one developed by 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 this model by including a renewable sector connected to the demand centers through high-capacity transmission lines. I develop this model in the context of a uniform auction wherein the generator participates by bidding on the price and quantity of electricity. I focus on the case of a marginal generator because its optimal bid determines the wholesale price. The corresponding markup set by the generator(s) is the 'realized markup.'

My model yields insights on how transmission expansion affects realized markups. In the short run, transmission expansion integrates wind energy into the grid, which affects the marginal fossil fuel generator in two ways: first, by displacing the energy production from the generator, and second, by changing the slope of the net-demand curve (net of electricity from wind) it faces. The overall impact of transmission expansion on markups is driven by the extent to which grid expansion integrates electricity from wind and the impact of additional wind on markups.

The above finding from the theoretical model motivates the empirical strategy for the short-run analysis. I use a fixed effects model to estimate the empirical analogs of the relationship between transmission expansion and markups. In the first step, I estimate the effect of transmission expansion on hourly wind generation, followed by the impact of wind generation on hourly markups. The empirical specifications flexibly control for confounding factors like electricity demand and seasonality that could be correlated with wind generation and markups. I find that CREZ expansion led to moderate decreases in markups, with the magnitude of reduction strongest during periods of high wind generation. A counterfactual analysis suggests a $227 million annual reduction in rents collected by fossil fuel generators from consumers of electricity. These transfers are policy relevant, as they can lead to lower retail prices with potential distributional and equity implications, as well as welfare consequences in the medium term.

I use the same empirical framework to study the impact of CREZ expansion on hourly emissions across different regions of Texas. I find a decline in emissions on the order of $60 million annually, with about 57 percent of these benefits from the decline in local pollutants (SO_2 and NOx) and the remaining share from lower carbon emissions. The decline in emissions is dampened by ramping up of marginal coal generators as a result of wind intermittency during early hours of the day. These ramp-up effects are characterised by a spike in emissions, and are a cause of concern as they occur mainly from coal generators located near densely populated regions in Texas.

Next, I estimate the magnitude of long-run investment in wind generation in response to investment in electricity transmission. The identification challenge here is that locations with superior wind quality were selected to site CREZ lines. I implement Coarsened Exact Matching to address the selection issue (Iacus et al. 2012). 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 pre-grid expansion wind capacity, wind resource quality (wind speed, capacity factor, wind turbine class), land price (average land price and median land acreage), terrain ruggedness, and county demographics (average farm size, population, and median household income).

Regressions using the matched sample suggest that counties that received transmission infrastructure saw 74 MW (+205%) higher wind capacity, 40 more turbines (+249%), and about 29 MW (+109%) bigger wind projects over 2012 to 2019. A back of the envelope calculation shows that this wind capacity prevented approximately $271 million in damages from carbon emissions in Texas in 2019. This research adds empirical evidence on the long-run value of investment in transmission expansion, an understudied topic in economics.

The long-run specifications include control variables that account for unobserved characteristics of transmission line expansion that might be correlated with wind investment. These variables include matching characteristics, indicators for county-level wind ordinances, and Production Tax Credit expiration, as well as matching groups by time trend fixed effects. Moreover, these results are robust to a battery of tests to address various threats to identification in the matching exercise. These tests check for selection on unobservables, including county-level lobbying efforts for or against CREZ expansion, anticipation of CREZ announcement, spillovers to control counties adjacent to CREZ counties, and project extensions near the announcement date which could influence location selection.

Finally, I explore how these responses affect wind curtailment, defined as the reduction in electricity generated from a project below the level it could have produced given available resources. While CREZ reduced wind curtailment in the short run by integrating wind into the grid, I provide descriptive evidence of rising curtailments in wind farms near CREZ counties over the past few years as a result of localized long-run investment in wind. These curtailments point to inadequate transmission capacity, which could dampen the short-run benefits.

While transmission expansions are expensive endeavors, the benefits accrue over time. However, my analysis indicates shorter payback periods than the ones reported in the literature (LaRiviere and Lyu 2022). The CREZ project cost about \$6.8 billion and my estimates imply a payback period of as short as 12 - 15 years.[2] These estimated benefits are in conjunction with many additional benefits, such as enhanced grid reliability, reduced transmission congestion,[3] and less local pollution. Therefore, these estimates are

2. The benefits from lower carbon emissions are computed using the Social Cost of Carbon (SCC) of \$51/ton-CO_2 as set by the US government (US Interagency Working Group on Social Cost of Carbon 2021). Rennert et al. (2022) have proposed an updated value of \$185/ton-CO_2 that incorporates the latest research in climate models and other socioeconomic factors. Using this value, the payback periods are as short as 5 - 9 years.

3. Transmission lines are said to be congested when they operate at maximum capacity. Some of the main reasons for transmission congestion are insufficient transmission capacity and spike in demand due to weather conditions.

conservative. The findings from this paper also provide insights for grid expansion in other parts of the US. The theoretical model and the empirical strategy can be applied to regions such as the Midwest and the Southwest, where transmission expansion would integrate renewable resources into the grid and lead to reductions in both emissions and market power associated with the fossil fuel sector.

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 post-deregulation electricity markets have found that market power contributes to high wholesale prices (Borenstein et al. 2002) and misallocation of generating resources due to sub-optimal bidding behavior (Hortacsu and Puller 2008; Hernández 2018). The existence of market power in sequential electricity markets causes a lack of arbitrage, which results in price premia across markets (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 electricity markets. Theoretical studies in this area employ Cournot models and simulations to show how expansion in transmission capacity leads to more 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). I make theoretical and empirical contributions to this literature by developing an auction-based model of the marginal fossil fuel generator and estimating the empirical analogs of the comparative statics derived from this model.

Third, I add to the recent literature looking at the link between transmission expansion, wind energy, and 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). Recent papers find that CREZ led to a significant reduction in wholesale market prices (LaRiviere and Lyu 2022), congestion risk, and the cost of hedging (Doshi and Du 2021). Fell et al. (2021) study how CREZ expansion enhanced the environmental value of wind measured by emissions avoided. Finally, along with Gonzales et al. (2022), this study is among the first in economics to quantify the investment in renewable energy in response to transmission expansion.

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 summary statistics in Section 3. The short-run analysis of markups and emissions is presented in Section 4 and Section 5, respectively. Section 6 is the long-run analysis followed by the implications of short- and long-run responses on wind curtailment in Section 7. Section 8 provides a concluding discussion.

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. ERCOT also schedules the dispatch of generators in order to meet demand for electricity at all times. ERCOT oversees more than 46,500 miles of electricity transmission and 700 generators serving electricity demand from over 26 million consumers over the state of Texas.[4] 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, the power sector in Texas contributed about 212.4 million metric tonnes of carbon emissions, about 12.3 percent of the total carbon 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.

Figure 1a shows the distribution of all the utility scale 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, while most of the fossil fuel capacity and major demand centers are located in the East and South. The Texas electricity market is connected by a network of transmission lines that carried about 74,820 MW of electricity at a record peak demand on August 12, 2019 (ERCOT 2021).[5]

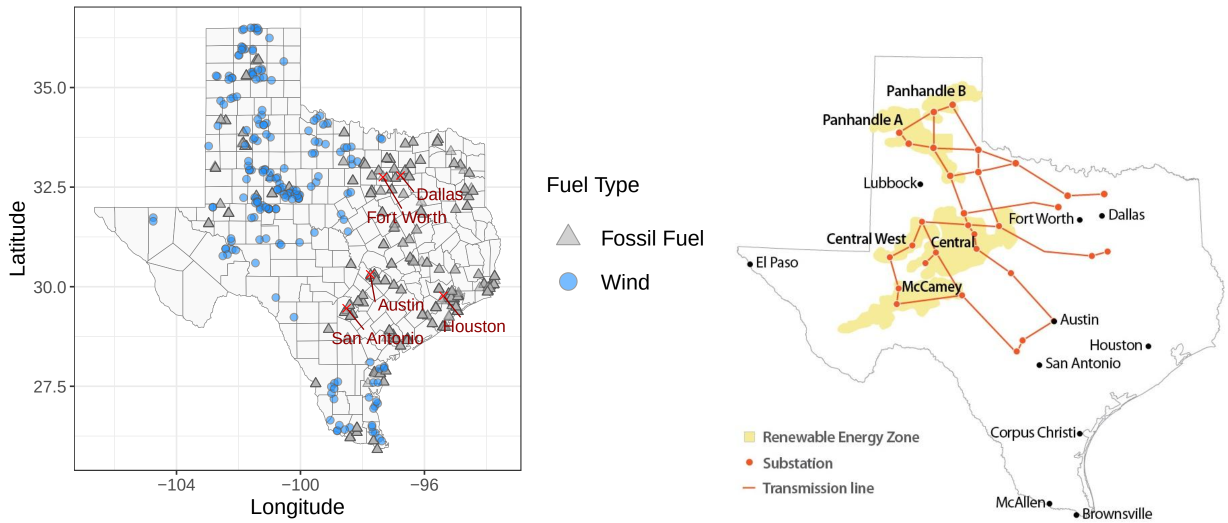
2.2 Competitive Renewable Energy Zones

Competitive Renewable Energy Zones (CREZ) was a large-scale transmission expansion project aimed at integrating electricity generation from wind farms located in the West to the major demand centers in the North, South, and Houston zones (Figure 1b). The project, commissioned in 2008 by the Public Utilities Commission of Texas, was aimed at accommodating 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 wind generators (Billo 2017). These lines were built over a period of 2011 through 2013 with a total cost of approximately $6.8 billion. All

4. The ERCOT interconnection is comprised of six zones within Texas - Panhandle, West, North, South, Houston, and Coastal. Figure E3a provides a sense of the geographic distribution of counties in these zones.

5. 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).

of the CREZ-based transmission lines were placed in service by December 2013 (Lasher 2014).



(a) Utility scale wind farms and fossil fuel generators

(b) CREZ lines and substations

Figure 1: ERCOT Zones and CREZ transmission expansion

Note: Figure 1a shows the geographic distribution of various electricity generators in Texas. Fossil Fuel generators include coal, natural gas, petroleum, and other gas based generators. Petroleum and other gas based generators are only 2 percent of total generators in Texas. Red triangles mark the locations of the five biggest population centers in Texas. Figure 1b shows the location of the CREZ transmission lines.

3 Data and Descriptive Statistics

I assemble multiple datasets with varying levels of temporal resolution. For the short-run analysis of generator markups, I assemble an 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, including ERCOT, the Energy Information Administration (EIA), and the Environmental Protection Agency (EPA).

3.1 Data

3.1.1 Markups

Markups are defined as $p - c$, where p is the Locational Marginal Price (LMP) and c is the marginal cost of production. LMP is the price of supplying one MWh of electricity at a particular location. I use data that is publicly available from ERCOT to identify the price-setting (marginal) generators and the corresponding LMP at each hour of the sample. The other component of markup is the marginal cost of production. The generating technology assumes a constant marginal cost of generation since fuel costs are constant in the short run. This assumption is common in the literature. Marginal cost is the sum of two main components: fuel costs and emissions permit costs for SO_2 and NOx.[6]

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 by the heat rate (HR_i) of the generator.[7] I use hourly electricity generation data at the generator level from ERCOT and heat input data from EPA's Continuous Emissions Monitoring system (CEMS). Finally, I compute emissions permit costs using daily data on NOx and SO_2 allowance prices ($p_t^{\text{SO}_2/\text{NO}_x}$) 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. The marginal cost c_{it} of generator i in period t is:[8]

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

6. Under the US 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.

7. EIA defines heat rate as the amount of energy used by a power plant to produce 1 KiloWatt hour (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.

8. Figure E2 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 fuel types.

3.1.2 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. Because the impact of local pollutants varies across space due to differences in population densities, I use estimates of county-specific marginal damages due to an additional ton of SO_2 and NOx from Holland et al. (2016).[9] I combine these county-specific damage estimates with SO_2 and NOx emissions from each generator to compute the dollar value of damages from these pollutants.

3.1.3 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 ease of interpretation. As shown in Figure 2a, the CREZ started in 2010, and over 80 percent of the project was completed in 2013.

3.2 Descriptive Statistics

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 and the remaining 30 percent are coal units. Coal generators are much larger in size than natural gas generators. The average coal generator is almost three times the capacity of an average natural gas generator.

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

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

	Coal		Natural Gas	
	Mean	Std. Dev.	Mean	Std. Dev.
Nameplate Capacity (MW)	602.37	200.99	189.93	86.53
Marginal Cost (\$/MWh)	21.83	21.04	15.50	14.22
Realized Markups (\$/MWh)	4.18	31.97	16.58	60.40
CO_2 damages (\$/MWh)	79.02	79.71	24.77	27.90
SO_2 & NOx damages (\$/MWh)	102.40	138.37	0.76	2.87

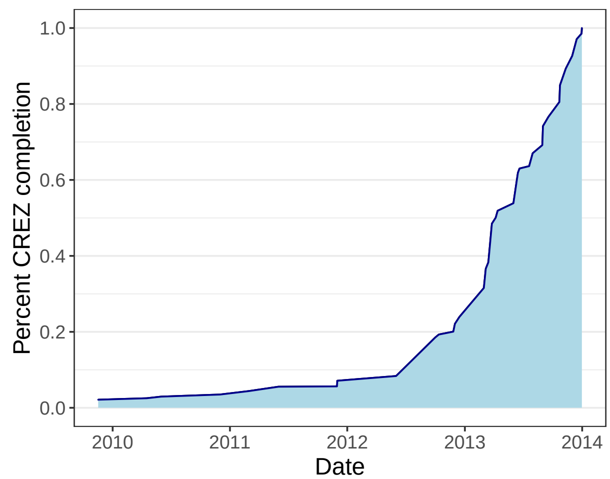
Notes: This table presents descriptive statistics of key variables by generator fuel type. Sample is marginal generator-hour observations from August 2011 - December 2014. Total # generator-hour observations (N) is 619,864. Frequency of coal generators is 33.12% and natural gas generators are 66.88%. Damages (in 2020 \$) 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 also much more polluting than natural gas. For ease of comparison, I present damage estimates (2020\$) for emissions per MWh of electricity generated. Damages from carbon emissions from coal generators are about \$79/MWh, compared to \$25/MWh 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 are on average \$100 higher than natural gas generators.

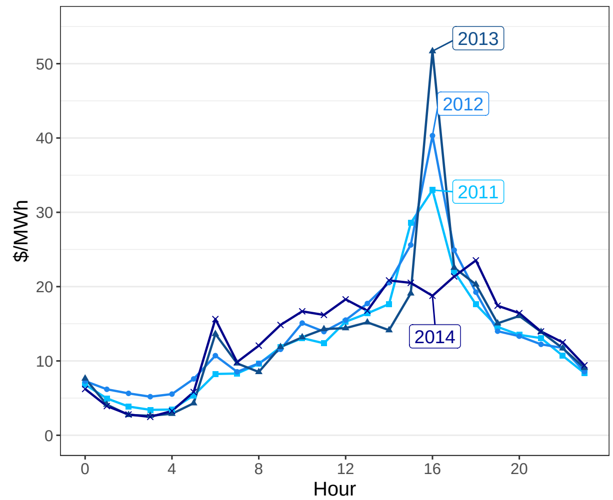
While the average marginal cost of coal generators is about \$6/MWh higher than the marginal cost of natural gas generators, the average markup set by marginal natural gas generators is about four times that of the coal generators. This is because coal generators tend to operate at the margin during the night, 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.

Figure 2b shows the hourly variation in markups, which is not apparent from Table 1. 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 (after CREZ expansion) across the peak hours of 14:00 to 17:00, perhaps most significantly at

16:00. Average markups show substantial hourly variation, suggesting seasonality and generator idiosyncrasies.



(a) Daily progress of CREZ expansion.



(b) Average hourly realized markups ($/MWh)

Figure 2: Daily CREZ progress and generator markups over the years

Note: Figure 2a shows the cumulative share of CREZ lines (miles) completed each day from 2010 to 2014. Figure 2b shows the average hourly realized price-cost markups set by fossil fuel generators (2011 - 2014, N = 619,864).

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 the participation of various market participants, including generators, retailers, transmission service providers, and distributors. ERCOT manages the efficient operation of the real-time market, including scheduling the dispatch of generators to meet the demand at all times using a series of sequential auctions.

Electricity transmission enables the flow of electricity from the generating units to the demand centers. Generators are scheduled to dispatch in an increasing order of electricity generating costs. Thus, renewable generators are always scheduled to dispatch first, followed by fossil fuel units. Natural gas generators are typically dispatched to meet any sudden surge in demand at peak hours.[10]

Transmission lines operate under certain capacity limits that need to be maintained. Inadequate transmission capacity between the West and other parts of Texas can lead to congestion, thereby preventing the export of electricity from the wind-rich West to demand centers in the East and South.[11] The presence of transmission constraints would cause ERCOT to schedule electricity from local generating units that are typically fossil fuel fired fired generators. This not only results in emissions that could have been offset by clean wind-based energy but also incentivizes local fossil fuel generators to charge markups over their marginal cost of production.[12] Transmission expansion is a key public policy investment aimed at relieving transmission congestion and integrating renewable generators into the grid. As I show in the theoretical model below, transmission expansion affects the markup charged by fossil fuel generators.

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

10. ERCOT defines peak hours as 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.

11. Transmission lines are said to be congested when they operate at maximum capacity. This is another way of saying that transmission constraints between two points A and B are binding. Some of the reasons for transmission congestion or binding transmission constraints are increase in demand due to weather conditions, outages, and insufficient transmission capacity, to name a few.

12. Please refer to Appendix B for an example that illustrates this phenomenon.

based on Ryan (2021), but I differ from it in two key ways. First, I introduce a wind-generating sector that 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 in 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: West, denoted by $\mathcal{W}$ and East, denoted by $\mathcal{E}$. West ($\mathcal{W}$) is comprised of wind farms and East ($\mathcal{E}$) is comprised of fossil fuel generators that serve a large demand center. Electricity transmission capacity (K) enables export of electricity generated by wind in $\mathcal{W}$ to demand centers in $\mathcal{E}$.[13]

In this model, I focus on the pricing decision of a profit-maximizing fossil fuel generator i located in $\mathcal{E}$. 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 $\mathcal{E}$.

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

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

13. Figure E4 in the Appendix illustrates this setup graphically.

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

Generator i faces a downward-sloping residual demand curve $D_i^r(p, q_w; \mathrm{K})$ comprised of three elements: market demand $D^{\mathcal{E}}$; electricity generated from wind q_w; and the total electricity generated from competitor fossil fuel generators, $Q_f(q_w, p) = \sum\limits_{j \neq i, j \in \mathcal{S}} Q_j(q_w, p)$. 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.[14] $Q_f(q_w, p)$ is strictly increasing in p and strictly decreasing in q_w.[15] Mathematically, D_i^r can be written as,

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

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:

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

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

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

14. Wind-based electricity generation incurs zero marginal cost and is always scheduled to dispatch first. I assume $D^{\mathcal{E}} > q_w$, otherwise there wouldn't be any need to schedule electricity from fossil fuel generators, because all of the market demand could be met by wind.

15. The interpretation of these assumptions is as follows:

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

Equation (4) is the optimal pricing rule for generator i, which sets price equal to marginal cost plus a markup term. $\frac{\partial p}{\partial b_i}$ is the slope of the market-clearing 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., $C_i'(D_i^r(p, K)) = c_i$, as well as 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 'realized 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 a 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 less than the marginal cost for the electricity generated.

The denominator is the slope of residual demand curve, which 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

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

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

I express Equation (6) as a percentage change in markups by multiplying both sides by the inverse of Equation (5). I split the resulting expression into two terms, Δ*Displacement* and Δ*Slope*, which measure the effect of transmission line expansion on markups:

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

ΔDisplacement. Changes in transmission capacity K can affect the production by generator i, causing a *displacement* of its residual demand curve.

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

I summarize the intuition behind the directions of the two terms on the right-hand side of Equation 8 in Propositions 1 and 2.

Proposition 1 *Transmission expansion leads to integration of electricity from wind into the grid.*

The second term in Equation (8) measures the magnitude of wind power that would be integrated into the grid in the short run as a result of transmission expansion. Because the generating capacity of wind remains fixed in the short run, transmission expansion would enable higher imports of wind generation to demand centers, i.e., $\frac{\partial q_w}{\partial \mathrm{K}} \geq 0$.

Proposition 2 *Integration of wind due to transmission expansion leads to a displacement of a marginal generator's residual demand curve.*

The first term in Equation (8) measures the extent to which the production decision of generator i is affected by additional wind energy. Consider the hypothetical electricity dispatch curve shown in Figure 3a. 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 of generators arranged in increasing order of the offer price. The dotted vertical line (D) is the demand for electricity and is assumed to be fixed in the short run. Generators are dispatched in increasing order of the offer price until the demand is met. The generator(s) dispatched with the highest offer price is the marginal generator, which determines the wholesale price of electricity.[16] In the scenario below, generator i submits the highest offer price c_4 and is thus the marginal generator.

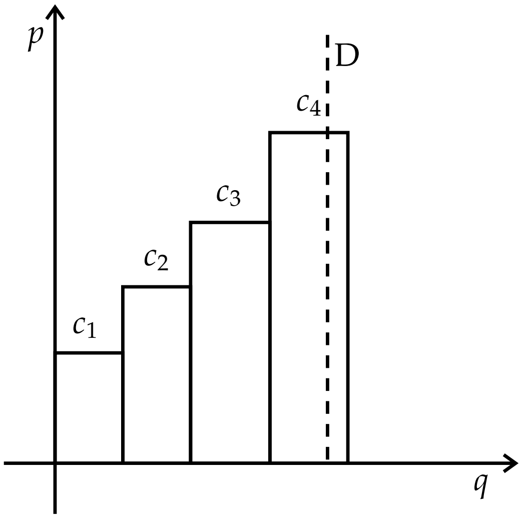
Next, consider the scenario in Figure 3b, wherein additional wind (W) displaces electricity generated from i, shown as the hatched area. This can be written as:

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

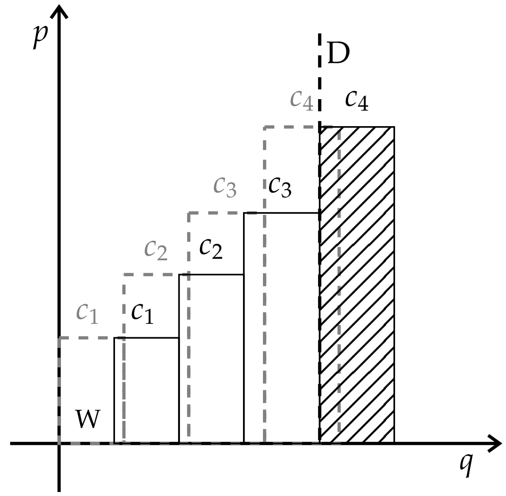
Thus, electricity from wind shifts the dispatch curve to the right, displacing power generated by the marginal generator i. This is reflected as an inward shift of i's residual demand curve, which in turn reduces its ability to set higher markups. This is shown in Figure 3c, with the generator moving from point A to point B of its offer curve after 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.

ΔSlope. This term measures the impact of transmission capacity on the slope of 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 ($D^{\mathcal{E}}$) and wind generation (q_w) are invariant to changes in p, the slope

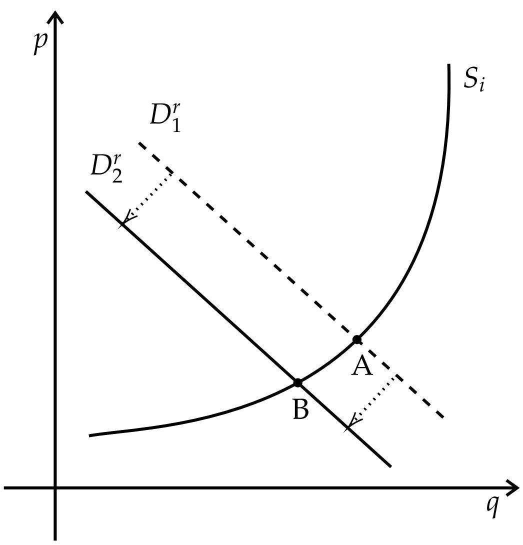
16. If there are multiple generators that have the highest offer price and are dispatched, all of them are referred to as marginal generators.



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

Notes: c_i denotes generator i's offer/bid price to supply electricity. The vertical dotted line in Figure 3a and Figure 3b denotes the demand for electricity (D), which is inelastic in the short run. W is the wind integrated into the grid due to transmission expansion, and S_i denotes the supply curve of generator i.

depends only on the production decisions of other fossil fuel generators. Therefore,

$$
\begin{aligned}
\frac{\partial^2 D_i^r(p, q_w; \mathrm{K})}{\partial p \partial \mathrm{K}} &= -\frac{\partial^2 Q_f(q_w, p)}{\partial p \partial q_w} \cdot \frac{\partial q_w}{\partial \mathrm{K}} \\
\frac{\partial^2 D_i^r(p, q_w; \mathrm{K})}{\partial p \partial \mathrm{K}} &= -\frac{\partial \eta_f}{\partial q_w} \cdot \frac{\partial q_w}{\partial \mathrm{K}}
\end{aligned} \tag{10}
$$

where $\eta_f = \frac{\partial Q_f}{\partial p}$ (> 0) is the slope of the aggregate supply curve of (marginal) fossil fuel generators. The value η_f determines the slope of the dispatch curve at the margin. Equation (10) shows that changes in the slope of the dispatch curve due to additional wind $\left(\frac{\partial \eta_f}{\partial q_w}\right)$ affects generator i's residual demand curve. This leads to the following proposition:

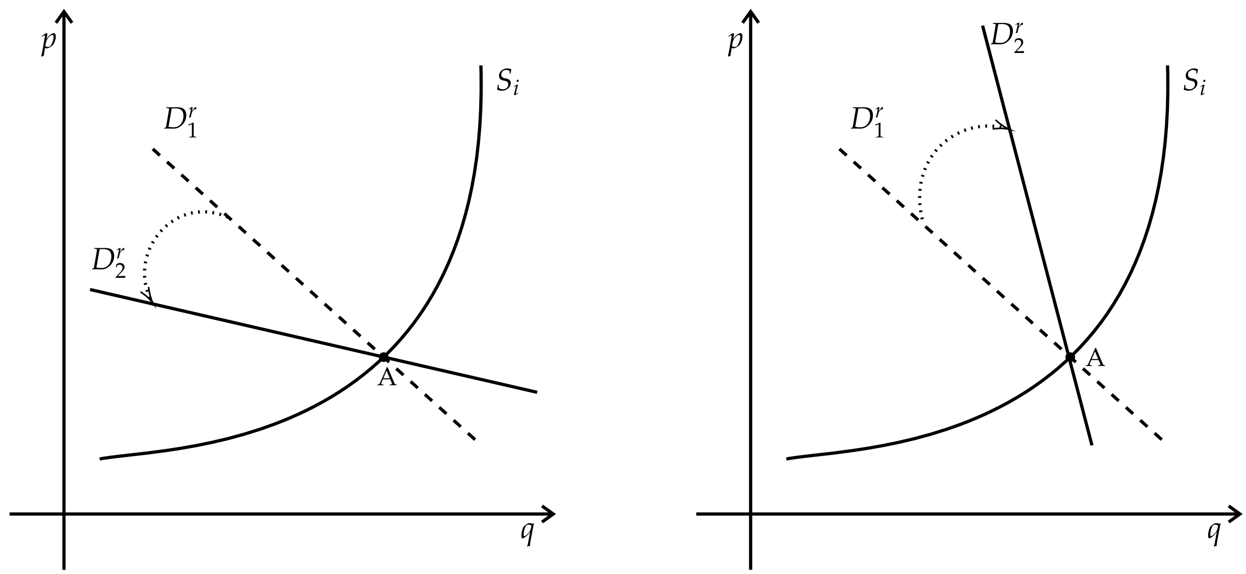
Proposition 3 *The impact of transmission expansion on the slope of the marginal generator's residual demand curve is ambiguous.*

As shown in Figure 3b, additional wind shifts the dispatch curve to the right. Heterogeneity in the cost of electricity generation by fossil fuel producers could lead to steeper or flatter electricity dispatch curve at the margin. For example, during periods of low demand, the dispatch curve tends be more elastic (flatter) whereas during periods of high demand, the dispatch curve is typically inelastic (steeper). Figure 4 illustrates changes in the slope of generator i's residual demand curve as a result of steeper or flatter dispatch curve at the margin.

Figure 4a shows that a flatter dispatch curve at the margin, post integration of wind, results in a more elastic residual demand curve. In this case, $\frac{\partial \eta_f}{\partial q_w}$ is negative which implies a more elastic residual demand curve (from Equation 10), thereby reducing the generator's ability to set higher markups. Figure 4b shows the opposite case, wherein the dispatch curve is steeper, i.e. $\frac{\partial \eta_f}{\partial q_w}$ is positive, resulting in a more inelastic residual demand curve. This in turn enhances the generator's ability to set higher markups.

Therefore, $\frac{\partial \eta_f}{\partial q_w}$ could be weakly positive or negative, which implies the right-hand side of Equation (10) to be weakly negative or positive. Substituting the expressions for ΔDisplacement from Equation (8) and ΔSlope from Equation (10) in Equation (7) shows the overall effect of transmission expansion on generator i's markups:

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



(a) Counterclockwise 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 4: Rotation of fossil fuel generator's residual demand curve after integration of wind due to transmission expansion.

Notes: 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.

4.2.3 Summary of main findings

Equation (11) can alternatively be expressed as:

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

Equation (12) summarizes the findings from the theoretical model. It shows that the overall effect of transmission expansion on realized markups in the short run is driven by two factors. The first is the effect of wind generation on markups, measured by $\frac{\partial(p-c_i)}{\partial q_w}$. The second is the extent to which transmission expansion integrates the electricity generated from wind into the grid, measured by $\frac{\partial q_w}{\partial K}$. In the empirical strategy below, I estimate the empirical analogues of each of the two components of Equation (12). The overall effect of grid expansion on markups is simply the product of the two components.

4.3 Empirical Strategy

The findings from the theoretical model motivate the empirical strategy for the short-run analysis. I run a set of fixed effects models to estimate the empirical analogues of Equation 12.

4.3.1 Impact of wind generation on markups

I use the following specification to estimate how additional wind generation affects markups:

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

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) at hour t is denoted by w_t. The parameter of interest is α_h, which measures the change in markups as a result of additional wind generation. Thus,

$$\alpha_h \approx \frac{\partial(p-c_i)}{\partial q_w}$$

I use a wide variety of controls to account for potential confounding factors in Equation 13. I use a quadratic polynomial of system-wide electricity demand D_t to account for variation in markups driven by spikes in electricity demand.[17] I use generator fixed effects (κ_i) to control for any generator-specific heterogeneity in markups. Finally, 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 patterns 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 the winter and summer months.

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

The identifying variation for α_h 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 identified from deviations in markups from generator-specific averages across all 16:00 hours (or 4 PM) within a month, in a given year. Standard errors are clustered at the generator level to account for correlation in markups at the unit level.

4.3.2 Impact of CREZ expansion on integrating wind generation

I use the following specification to estimate the impact of CREZ expansion on wind generation:

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

where w_t is the wind generation (GWh) in hour t and $crez_d$ is the percentage completion of CREZ transmission project at day d of the sample. The parameter of interest is β_h, which measures the integration of wind energy into the grid as a result of CREZ expansion. Thus,

$$\beta_h \approx \frac{\partial q_w}{\partial K}$$

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.[18] This variable incorporates the generating capacity and technology of the wind generator and also takes into account the real-time meteorological conditions that could affect the amount of power generated through wind farms.

As shown in Figure 5, 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

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

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.

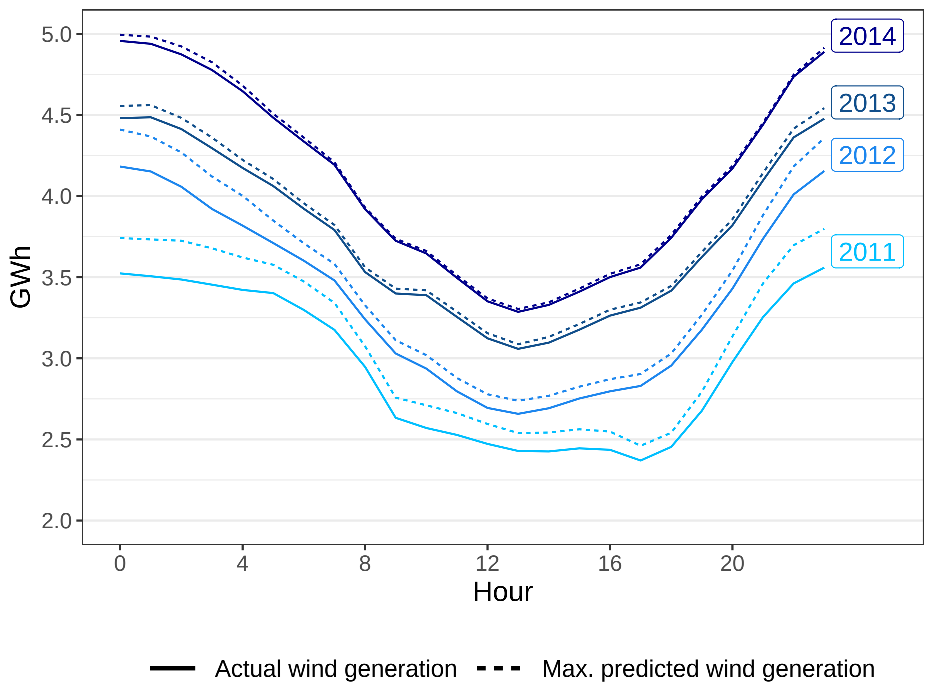


Figure 5: Hourly averages of actual wind generation (w_t) and maximum predicted wind generation (max_t) from 2011 - 2014.
Notes: max_t is the maximum energy production capability of the generator at period t. It is established by the generator itself and is continuously updated in real time.

I use hour-by-month fixed effects (η_{hm}) in Equation 14 to control for seasonality in wind generation. Thus, conditional on predicted wind generation (max_t) and η_{hm}, β_h identifies the additional wind energy integrated into the grid as a result of transmission expansion. The identifying variation comes from changes in wind generation caused by daily transmission expansion across the same hours in a given month. I use Newey West auto-correlation corrected standard errors with a seven-day lag structure in Equation 14.

Under the identifying assumption that the fixed effects and controls 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.

4.4 Results

Figure 6 shows the coefficient estimates of $\hat{\alpha}_h$ from Equation 13, i.e., the change in fossil fuel markups due to additional wind in the grid. We see that, on average, the drop in markups is strongest in magnitude at the peak demand hour at 4 PM, 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 on-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 during the on-peak hours than during the off-peak hours.[19]

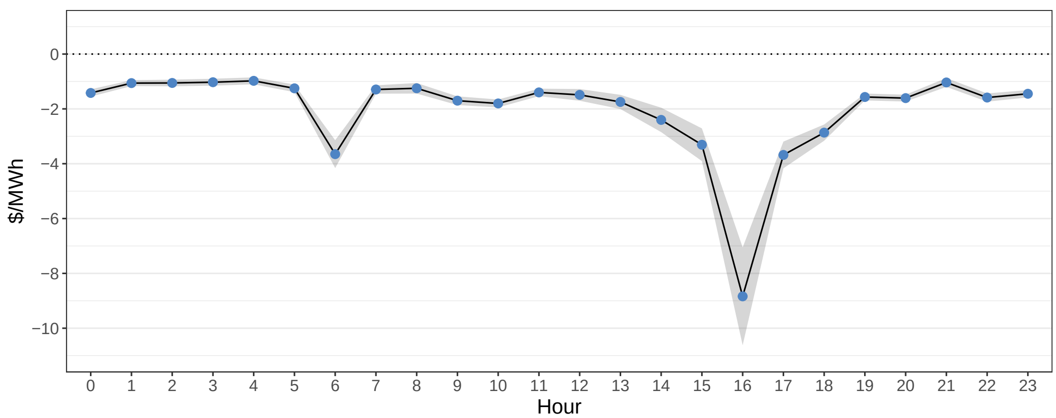


Figure 6: Effect of additional wind energy on realized markups.

Notes: This figure shows the coefficient estimates ($\hat{\alpha}_h$) from Equation 13. Each point estimate is the average impact of additional GWh of wind energy on generator markups (\$/MWh) for each hour. 95 percent confidence intervals constructed from standard errors clustered at the generator level.

Figure 7 presents the coefficient estimates of $\hat{\beta}$ Equation 14, i.e., the effect of CREZ expansion on wind generation. The coefficient estimates imply that keeping the stock of generating capacity fixed, CREZ integrated about 0.22 GWh of wind at midnight, and about 0.10 GWh during the peak demand hours between 3:00 and 6:00 PM. 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 night compared to the day.

19. The on-peak hours in ERCOT are defined as the hours between 7:00 AM and 10:00 Central Daylight Time from Monday through Friday.

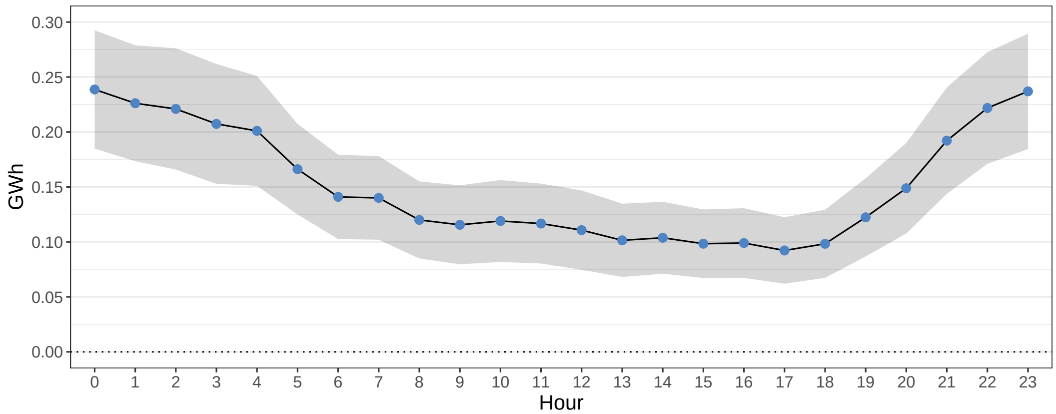


Figure 7: Impact of CREZ expansion on integrating wind energy into the grid

Notes: This figure shows the coefficient estimates ($\hat{\beta}_h$) from Equation 14. Each point estimate measures the average effect of CREZ expansion ($crez_d = 1$) on integrating wind generation (GWh) at each hour. 95 percent confidence intervals constructed from Newey-West auto-correlation corrected standard errors with a 7-day lag structure.

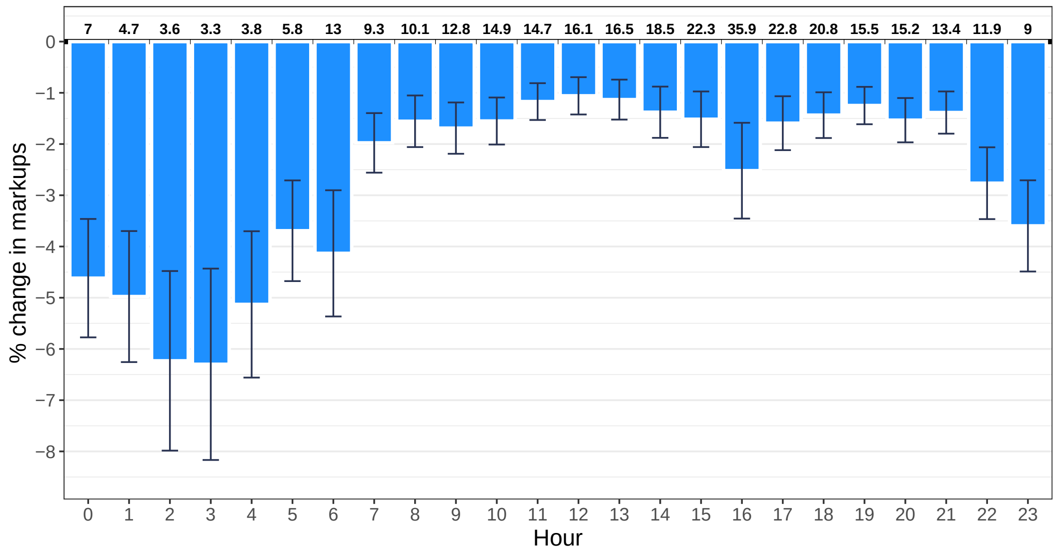
The overall impact of CREZ expansion on markups (θ) is given by the product of the effect of wind generation on markups (α_h) and the integration of wind energy as a result of CREZ (β_h),

$$\underbrace{\frac{\partial(p - c_i)}{\partial K}}_{\hat{\theta}} = \underbrace{\frac{\partial(p - c_i)}{\partial q_w}}_{\hat{\alpha}} \times \underbrace{\frac{\partial q_w}{\partial K}}_{\hat{\beta}} \tag{15}$$

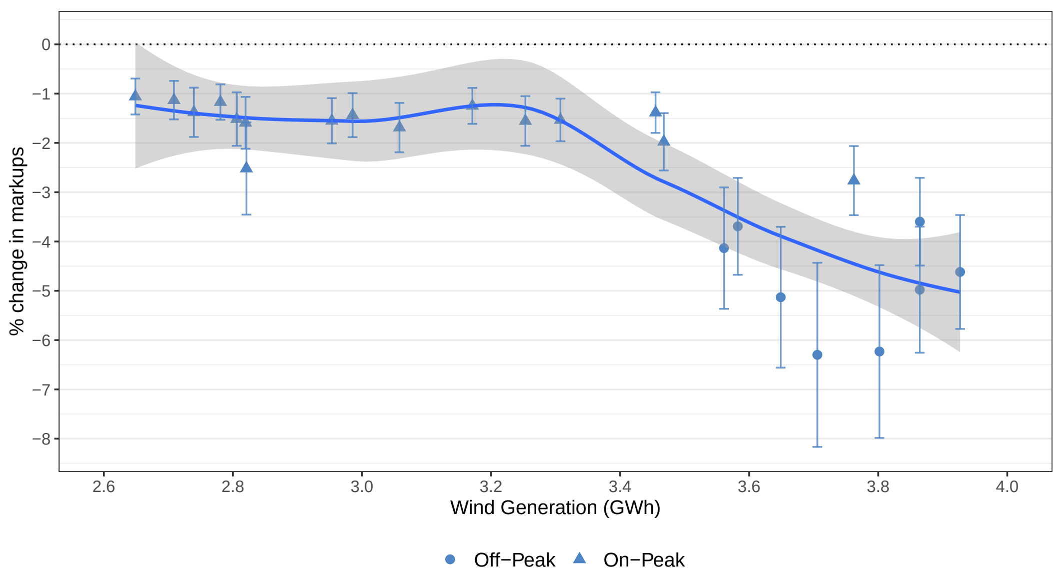
To provide a better sense of the magnitude of $\hat{\theta}$, I present the semi-elasticity of markups or the percentage change in markups in response to CREZ expansion in Figure 8a.[20] We see a clear distinction between the semi-elasticity of markups between off-peak vs. on-peak hours. The magnitude of the percent decline is highest for hours before 7:00 AM, with the maximum decrease of 6.3 percent at 3:00 AM. However, the percentage drop in markups for the on-peak hours (7:00 AM to 10:00 PM) is less than 3 percent, mainly because of the lower wind generation in these hours.

I arrange the semi-elasticity estimates in Figure 8a in increasing order of average wind generation for each hour in Figure 8b. The trend line of the estimates shows that the percentage decline in markups as a result of CREZ is strongest at high wind hours.

20. Figure E7 in the appendix shows the bar plot of magnitudes of $\hat{\theta}$. I calculate semi-elasticity values by dividing the θ estimates with the average markup at each hour.



(a) Percentage change (semi-elasticity) in markups due to CREZ completion ($crez_d = 1$)



(b) Semi-elasticity of markups for CREZ completion ($crez_d = 1$) for each hour.

Figure 8: Short-run impact of CREZ expansion on realized markups.

Notes: Figure 8a shows the percentage change (semi-elasticity) and the 95 percent confidence intervals for $\hat{\theta}_h = \hat{\alpha}_h \times \hat{\beta}_h$, where $\hat{\alpha}_h$ is the hourly impact of wind generation on markups from Figure 6 and $\hat{\beta}_h$ is the hourly impact of CREZ expansion on wind integration from Figure 7. Average markups for the sample are shown above the x axis. Figure 8b shows the semi-elasticity values in Figure 8a, arranged by increasing wind generation in that hour. On-Peak hours: 7:00 AM to 10:00 PM.

This is in line with the theoretical model, wherein additional wind leads to an inward shift in the fossil fuel generator's net-demand curve, thereby reducing its ability to set high markups. The percentage decline is highest at off-peak hours and peak hours with high wind generation.

4.5 Change in surplus from grid expansion

How do these changes in markups translate to gains or losses of producer surplus? In the short run, producers of electricity earn excess rents from the purchasers of electricity by exercising market power. While the retail rates of electricity paid by end-use consumers remain fixed in the short run, these excess rents are passed down from the retailers to the consumers, as the demand for electricity is downward sloping in the medium to long run (Deryugina et al. 2020).

I conduct a counterfactual exercise to estimate the changes in annual rents collected by 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. $crez = 0$). Next, I substitute $\widetilde{w}_t$ in the estimated Equation 13 to compute the counterfactual markups in the absence of CREZ expansion.

Rents accrued by the fossil fuel generators, or more simply the change in surplus (ΔS) from the absence of CREZ expansion, are:

$$\Delta S \approx \Delta(p - c) \times \widetilde{Q} \tag{16}$$

where $\Delta(p - c)$ is the change in markups and $\widetilde{Q}$ is the electricity from fossil fuel producers 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 fossil fuel generators. In other words, the electricity from wind integrated by CREZ would have been supplied by the fossil fuel generators in the absence of the expansion; therefore, $\widetilde{Q}_t = Q_t + (w_t - \widetilde{w}_t)$. Second, I assume constant marginal costs; thus, lower markups are reflected in a lower wholesale price of electricity (or LMP) for each generator. Therefore, $\Delta(p - c) = p_{(\text{crez}=0)} - p_{(\text{crez}=1)}$.[21]

The counterfactual analysis finds that generators would have accrued about $753 million (2020 $) over the sample period of my analysis in the absence of CREZ expansion. This is about a $227 million annual reduction in transfers from retailers to generators in the short run and from consumers of electricity in the long run. While these transfers due to market power do not indicate welfare loss in the short run, they may have distributional effects if they lead to higher retail prices and sub-optimal levels of electricity consumption in the medium run.

Note that this analysis does not include welfare gains due to more efficient dispatch of electricity generators as a result of transmission expansion. Further, I do not compute lower markups due to lower grid congestion. Thus, these figures are likely the lower-bound estimates of the decline in producer surplus due transmission expansion.

5 Short-run: Impact of CREZ expansion on emissions

As shown in Section 4, the addition of wind to the grid shifts the electricity dispatch curve to the right. In this section, I examine how integration of wind due to transmission expansion affected the emissions from the marginal fossil fuel generator(s) throughout the day. The closest empirical study in this regard is Fell et al. (2021), where the authors study how lower grid congestion as a result of CREZ enhanced the value of total wind generation as measured by lower emissions. By contrast, I focus on how CREZ integrated more wind into the grid (keeping the generating capacity fixed) and the subsequent impact on emissions from marginal fossil fuel generators in the short run. The findings from this section are therefore complementary to Fell et al. (2021).

21. 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.

Marginal generators typically respond to changes in demand for electricity over the course of the day by ramping up or down. Variation in the fuel types of generators at the margin over the course of a day makes such an analysis informative. As mentioned, 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 and thereby reduce emissions. I run the following regression to estimate the impact of additional wind capacity on marginal emissions:

$$E_{zt} = \rho_{zh} \cdot w_t + f(D_{zt,t-1}|\lambda) + \alpha_z + \delta_{hmy} + \epsilon_{zt} \tag{17}$$

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 effect of an additional GWh of system-wide wind generation on the marginal emissions in zone z at hour h.

I use a cubic polynomial of contemporaneous and lagged demand for electricity $D_{zt,t-1}$ at the zone level to control for the variation in marginal emissions due to changes in 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 patterns in wind generation that could also be correlated with variation in emissions. To account for baseline level of emissions across the zones, I use zone fixed effects α_z. Standard errors are clustered at the daily level to account for serial correlation. I restrict my analysis to the four main load zones in Texas: West, North, South, and Houston.[22]

22. Figure E3b shows these zones along with the geographic distribution of marginal coal and natural gas generators in the sample.

5.1 Results

5.1.1 Impact on marginal carbon emissions

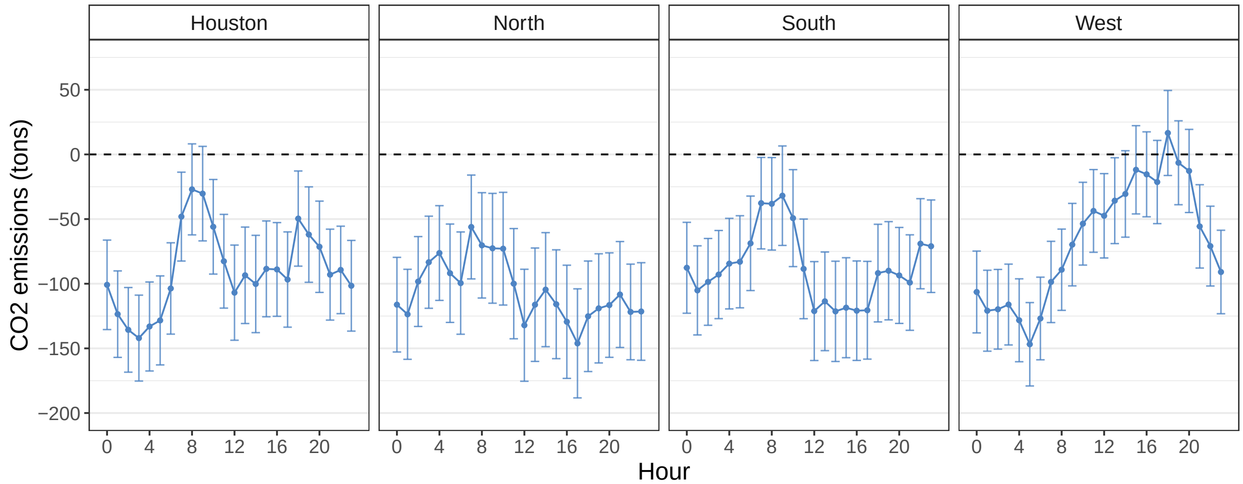
Figure 9a shows the estimates of ρ_{zh} from Equation 17, i.e., the effect of wind generation on carbon emissions. We see a clear decline in carbon emissions from generators across all the zones throughout the day in response to additional GWh of wind energy. The magnitude of decline in emissions is highest between noon and 10 PM in the North, South, and Houston zones. However, the drop in emissions from generators in the West is highest at night during periods of high wind.

To explore whether the pattern in Figure 9a is due to heterogeneity in generator fuel type, I estimate Equation 17 separately for the sample of marginal emissions from coal and natural gas generators. Two key insights emerge from the coefficient estimates in Figure 9b. First, the hourly pattern for coal is similar to the pattern in Figure 9a, 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 throughout the day across all four zones. This shifts the aggregate fuel type estimates in Figure 9a downward.

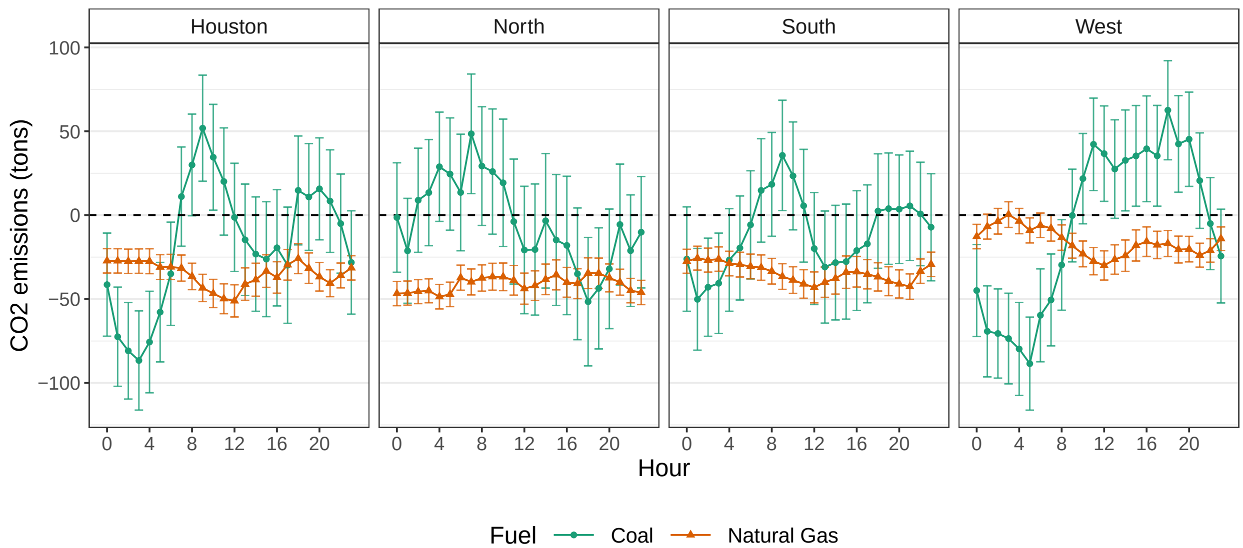
The coefficient estimates for coal generators suggest that electricity from wind has a significant effect in lowering emissions from coal generators at the margin during the night. However, there is a spike in emissions from coal units during the early hours of the day, especially in Houston and the 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 17. Figure 10 shows the coefficient estimates. The pattern of coefficient



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



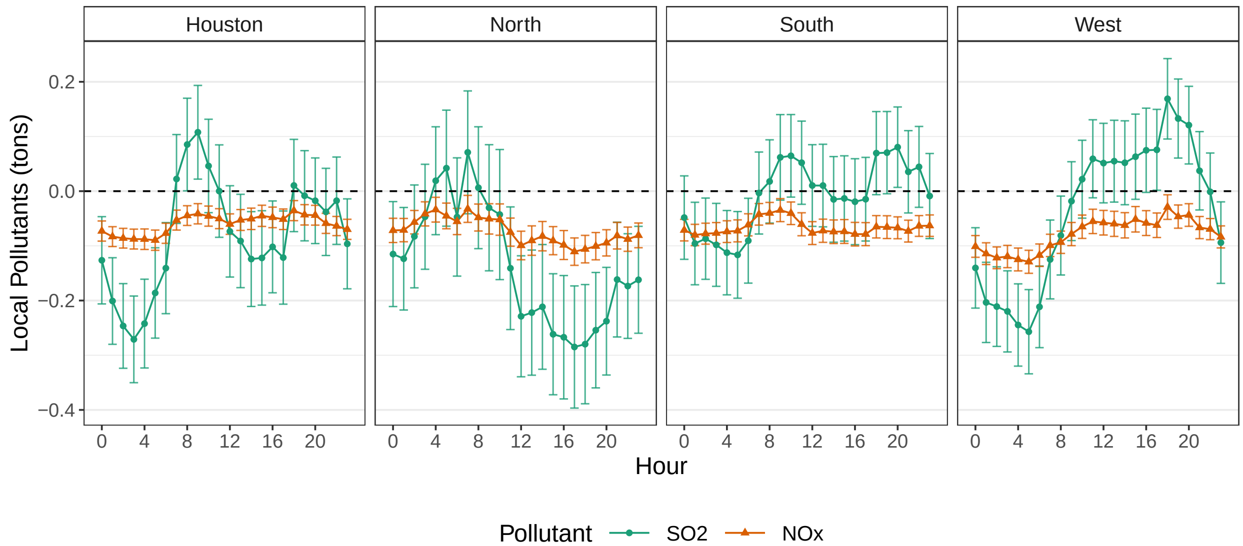
(b) Impact of additional wind generation on CO_2 emissions by generator fuel type

Figure 9: Short-run impact of wind generation on CO_2 emissions.
Note: This figure shows the coefficient estimates of the regression of hourly zonal carbon emissions on wind generation from Equation 17. 95 percent confidence intervals constructed from standard errors clustered at the daily level.

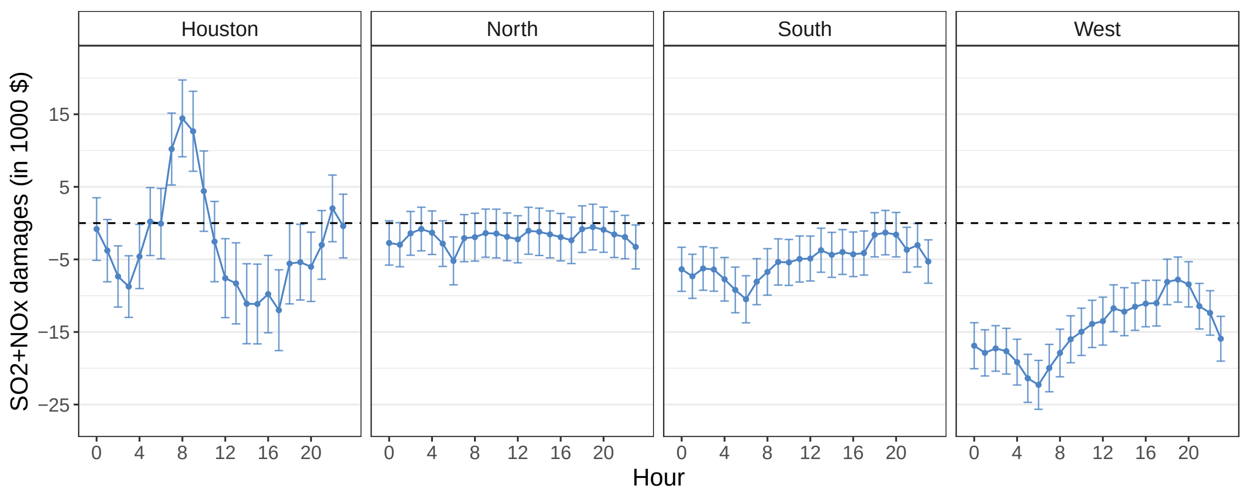
estimates of SO_2 in Figure 10a is similar to carbon emissions from coal generators in Figure 9b.[23] The presence of sulphur impurities leads to SO_2 emissions as a byproduct of burning coal in power plants, giving rise to the finding in Figure 10a. SO_2 emissions from natural gas power plants are low because of low amounts of sulphur in pipeline-

23. 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 E8 in the Appendix.

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).



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



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

Figure 10: Short-run impact of wind generation on emissions from local pollutants (SO_2 and NOx).

Note: This figure shows the coefficient estimates of the regression of hourly zonal local emissions (SO_2 and NOx) on wind generation from Equation 17. 95 percent confidence intervals constructed from standard errors clustered at the daily level. Figure 10b uses county-specific marginal damage estimates from Holland et al. (2016) to reflect the $ value of damages from local pollution (SO_2 and NOx).

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 aggregate these damages at the zonal level and estimate Equation 17 with the damage values as the dependent variable.

Coefficient estimates in Figure 10b show evidence of significant heterogeneity across zones with respect to damages avoided from local pollutants as a result of additional wind generation. For the South and West, additional wind leads to declines in damages from SO_2 and NOx across all hours, whereas the effect is statistically insignificant for the North. In the 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 West and Houston, we observe that the estimates for SO_2 and CO_2 emissions are driven by the only coal power plants in these zones. In Houston, the emissions are due to the W.A. Parish Coal Plant (four generators with total capacity of 2.7 GW), whereas in the West the emissions are due to the Oklaunion Power Plant (a single generator with 720 MW capacity). The spike in emissions is the result of ramping up these units to meet demand during periods of low wind generation after 8:00 AM. These ramping effects are shown to undercut the emissions reductions from wind, especially with generators operating at low levels of efficient generation (Lew et al. 2012). The availability of excess transmission capacity tends to promote power from these generators and is a cause of concern, especially since the 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:

$$D_z(\$) = \sum_{h=0}^{24} \tau \times \beta_h \times \rho_{zh} \tag{18}$$

where D_z is the zonal daily average of damage (in 2020 $) due to marginal carbon emissions in zone z. Taking the social cost of carbon, τ as $51 per ton of CO_2 emissions (US Interagency Working Group on Social Cost of Carbon 2021), 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 10b with β_h and aggregate over the hours to get the value of average daily damage avoided.

Table 2: Average daily damages avoided from marginal generators due to CREZ

	Damages Avoided (2020 $)			
Zone	CO2	SO2 + NOx	Total	Percent (%)
Houston	17,692	8,710	26,402	17
North	19,388	7,173	26,561	16
South	15,982	20,044	36,026	23
West	14,338	55,267	69,605	44
Total	67,400	91,194	158,594	100

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

Table 2 shows a decline in the estimates of daily damages from carbon emissions from generators across all the zones in the short run.[24] The total value of daily carbon emissions avoided is about $67,000, with three-fifths of the share due to generators from the North and the South, and a fifth each from Houston and the West. In the case of local pollutants, the total daily damages avoided are about $91,000, with 60 percent of the reduction concentrated in the West. The total daily damages avoided from CO_2, SO_2 and NOx emissions are approximately $160,000, which translates to about $58 million annually.

24. The coefficient estimates of hourly averages of damages avoided for each zone due to CREZ are presented in Figure E9 in Appendix E. The pattern for carbon emissions and local pollution is similar to Figure 9a and Figure 10b respectively.

6 Long-run: Impact of CREZ announcement on investment in wind energy

The CREZ transmission project was selected by the Public Utilities Commission of Texas (PUCT) in consultation with ERCOT after a multi-year process in July 2008 (NREL 2008).[25] It was aimed at accommodating 18.5 GW of total wind power – 6.9 GW by the end of 2008 and a projected 11.5 GW by 2012 – by building 3,600 miles of 345 kV electricity transmission lines between existing and new substations throughout the Panhandle, West, and East of Texas at a projected cost of $4.95 billion (PUCT 2009).[26]

Wind developers site their projects in regions with availability and access to transmission capacity and locate near the electrical substations to deliver their power to the grid.[27] In the data, I only see the counties where these substations were located and thus I call them 'CREZ counties'.[28] I refer to July 2008 as the "announcement date" because it provides the most accurate information about transmission siting in the CREZ project. The technical details of the transmission expansion – the cost breakdown, expected completion dates, and the transmission service providers responsible for the expansion – were released in October 2010 in the CREZ Progress Report (RS&H 2010).

Figure 11 shows a cluster of wind projects located within and near CREZ counties post 2008. This could be indicative of a long-run response to transmission expansion beyond the project capacity that was planned for 2012. [29] To parse out whether certain counties saw higher levels of wind investment in the long run as a result of CREZ

25. Refer Appendix B for a discussion on the planning behind CREZ expansion.

26. Electrical or transmission substations typically serve as the terminal points for high-voltage transmission lines, as well as serving as the hub for nearby generating plants to deliver their power to the grid.

27. Appendix C presents a simple conceptual model to build intuition about a wind developer's choice of siting its project.

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

29. There is also a cluster of wind farms in coastal Texas. This is because of superior wind quality in this region, which could be profitable for wind developers.

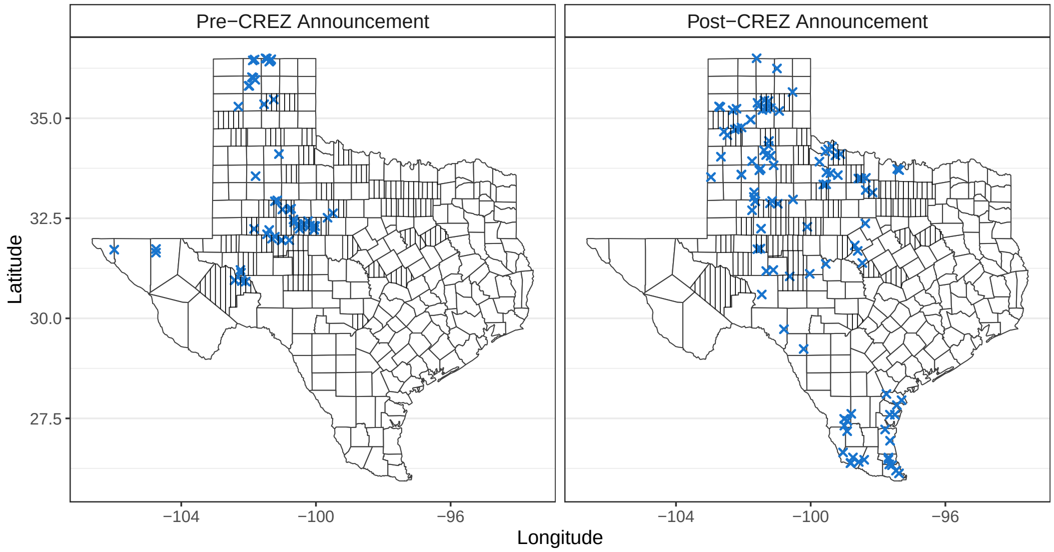


Figure 11: Location of wind projects pre- and post-CREZ announcement in July 2008
Note: This figure shows wind farms in Texas pre-CREZ announcement (Jan 2001 - Jul 2008) and post-CREZ announcement (Aug 2008 - Dec 2019). Counties announced as sites for substations for CREZ lines are shown with hatching.

expansion, I estimate the following specification:

$$y_{it} = \alpha + \beta \cdot crez_i + \mathbf{X}'\Pi + \epsilon_{it} \tag{19}$$

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 a substation for CREZ lines was sited in county i.

The analysis is restricted to annual county-level observations from 2012 through 2019 to estimate the wind capacity added beyond the projected period of CREZ planning. This excludes wind projects that were already in development or perhaps sited in CREZ counties just prior to the grid expansion announcement in late 2008. Because project planning and development typically takes a few years, this allows for the addition of wind capacity in response to transmission expansion.[30]

I use a battery of control variables and fixed effects summarized by vector $\mathbf{X}$ in Equation (19). I use wind turbine class, capacity factor, and cubic polynomial of average wind speed to flexibly control for a county's wind resource quality. These variables are aggregated at the county level from 2km $\times$ 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, and 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 siting and wind investment, I use median household income in 2007 and average population from 2007 to 2010. I use average farm size in a county to account for variation in wind investment due to turbine dis-amenities.[31] This data comes from the USDA Census of Agriculture. Cities and counties often enact regulations for wind projects that are sited in their jurisdiction. These regulations, commonly known as setbacks or wind ordinances, specify limits on factors such as the size of wind turbines, height of turbines, noise, and maximum capacity. I include an indicator variable specifying whether the county (or a city in the county) has a wind ordinance.[32] I use the publicly available dataset on wind ordinances from WINDExchange to construct this variable.

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

31. The rationale behind these variables is that urban areas tend to have higher opposition toward transmission and wind project siting (Andrade and Baldick 2016). Further, it is harder to site wind farms in areas with small farms (Winikoff and Parker 2019). 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.

32. Most counties in Texas do not have wind ordinances for wind projects. Out of 254 counties, I find that cities in only five counties – Dallas, Ellis, Kleberg, Taylor, and Wichita – have enacted a wind ordinance for both smaller and bigger wind projects. 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.

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

The main threat to identification in Equation (19) is the endogeneity of $crez_i$ due to the selection of regions with superior wind quality and historically higher levels of wind capacity as preferred locations for siting substations for CREZ lines. I attempt to account for these factors by including a rich array of control variables, but the concern about a lack of common support amongst counties still remains. I address these concerns by implementing a matching strategy to obtain an 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 treated 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):

$$\mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 0) \tag{20}$$

where ϵ_{it} is the unobserved component of the dependent variables of interest (y_{it}) – wind capacity, 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 effects 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 the CIA involves comparing counties with exactly the same combination of characteristics, such that $\mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 0)$. 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 that include 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 after 2012. These factors include pre-grid expansion wind capacity, wind resource quality, land price and ruggedness, ERCOT load zones, 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. I also match counties by terrain ruggedness, which I define as the standard deviation of elevation within a county using 30m $\times$ 30m elevation data from the National Elevation Dataset (NED). To account for bargaining power and community opposition (mainly from urban areas) in siting 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 perform exact matching on ERCOT load zones to capture regional differences across load zones in the Texas electricity market. [33]

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. Counties that do not lie in the common support of observable characteristics are discarded from the sample. Thus, the control group comprises 30 counties and the treated group comprises 13 counties. Figure 12 shows the map of treated and control counties. Most of the control counties (light yellow) are adjacent to the treated counties (dark yellow).

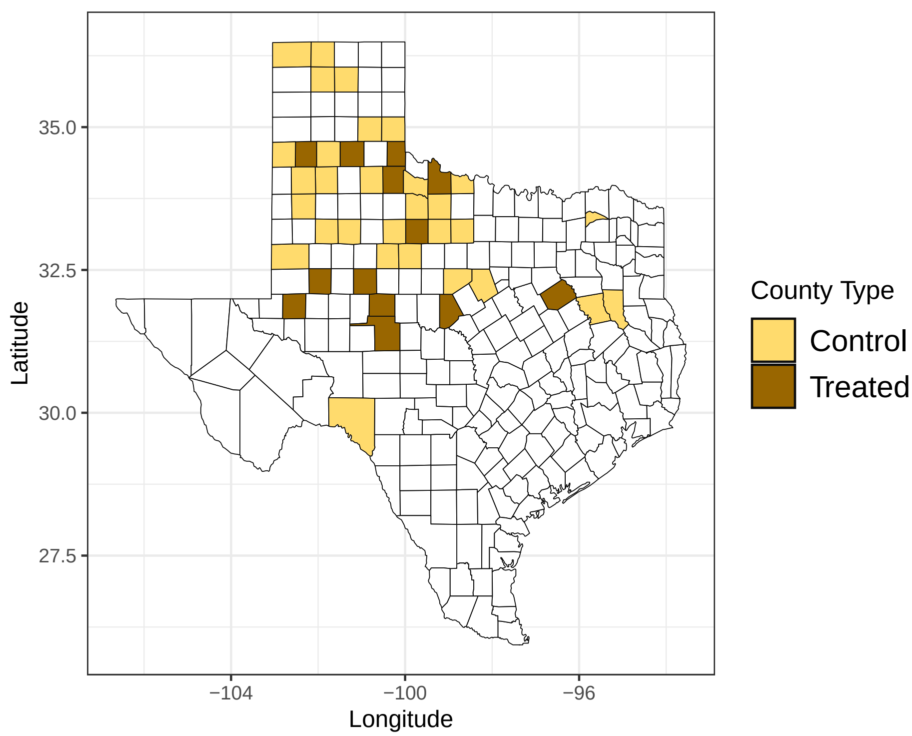


Figure 12: Treated and control counties obtained using Coarsened Exact Matching.
Notes: Total number of control counties are 30, total number of treated counties are 13. Unshaded counties are discarded from the sample used in the regression analysis because they lie outside of the common support of observable characteristics.

33. 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 wind turbine class and zone are discrete variables. Each category within wind turbine class is matched exactly, whereas I use the following structure for exact matching on zone: {{Panhandle, West}, North, Coastal, Houston, South, None}.

For the regression analysis on the counties obtained by matching, I use the same set of control variables as described in Equation 19. The key assumption is that, conditional on the vector of controls **X**, there are no unobservables that affect both the outcome variable and treatment status $(crez_i = 1)$. Thus, the coefficient estimate of '$crez_i$' in Equation 19 using the matched sample is the unbiased effect of CREZ on wind investment in the long run.

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

Type	Variables	Pre-Matching			Post-Matching		
		Means Treated [CREZ = 1]	Means Control [CREZ = 0]	p-val	Means Treated [CREZ = 1]	Means Control [CREZ = 0]	p-val
Wind Resource Quality	Pre-CREZ wind capacity	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	–
Land price and ruggedness	Avg. Land Price (2007-2010)	284.684	424.427	0	228.424	231.216	0.929
	Median Land Acreage	560.184	779.632	0.032	360.746	351.736	0.161
	Terrain Ruggedness (m)	22.238	20.033	0.001	21.073	18.648	0.268
ERCOT Load Zones	Coastal	0.000	0.051	–	0.000	0.000	–
	Houston	0.000	0.028	–	0.000	0.000	–
	None	0.000	0.107	–	0.000	0.000	–
	North	0.308	0.220	–	0.163	0.163	–
	Panhandle	0.179	0.136	–	0.302	0.371	–
	South	0.026	0.252	–	0.000	0.000	–
	West	0.487	0.206	–	0.535	0.466	–
Demographic characteristics	Avg. Farm Size (2007)	1,595.667	1,724.206	0.418	1,183.140	1,262.035	0.118
	Median Income (2007)	43,133.130	39,739.930	0	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 Counties		39	214		13	30	

Notes: This table presents balance test of key pre-treatment observable characteristics of a county. Pre-CREZ wind capacity is the total installed capacity (MW) in a county as of 2008. Terrain ruggedness is the standard deviation of elevation (metres) in a county. Each unit is a county-year observation. Wind turbine class is the indicator specifying the turbine class model most suited for projects in the county. Pre-Matching sample includes all county-year observations. Post-Matching sample is selected using Coarsened Exact Matching (CEM). Exact matching is implemented on factor variables like wind turbine class and ERCOT load zones.

6.1.1 Results

Table 4 reports the coefficient estimate of $crez_i$ from Equation 19 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, land price, terrain ruggedness, county demographics, fixed effects for wind ordinance, load zone, and an indicator for Production Tax Credit expiration. I also include interaction of group fixed effects with a linear time trend to account for time-varying unobserved factors that could affect specific matching groups.

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	73.73**	40.13***	29.33
	(29.40)	(14.44)	(17.68)
Mean dependent variable	35.9	16.1	26.9
Semi-elasticity (%)	205.4	249.2	109.0
Controls	✓	✓	✓
Group $\times$ Trend FE	✓	✓	✓
Sample	Matched	Matched	Matched
Observations	344	344	344
R^2	0.467	0.476	0.425

Notes: This table reports the estimate from Equation 19. 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 sited a substation for CREZ lines. All specifications include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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

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 74 MW higher wind capacity in treated counties over 2012 - 2019. The semi-elasticity indicates a 205.4 percent increase in wind capacity for CREZ counties. In a similar vein, Column (2) shows that treated counties had about 40 more turbines on average than the control counties, with a 'semi-elasticity' of 249 percent. Both of these results are statistically significant at the 5 percent 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 29 MW larger wind projects; however, the coefficient estimate is not statistically significant.

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 in Texas is 34.57 percent, 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-CO_2 (US Interagency Working Group on Social Cost of Carbon 2021), the value of total reduction in carbon emissions is about \$271 million.[34]

6.1.2 Threats to identification

1. Selection on unobservables:

The key threat to identification in matching is the selection of CREZ counties on unobservable characteristics. This would violate the Conditional Independence Assumption

34. 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.

(CIA) and the estimates would lose their causal interpretation. While I cannot test CIA directly, I provide institutional evidence and robustness checks to support its validity in this context.

The CREZ planning process involved discussions with various stakeholders, including wind developers, county officials, transmission service providers, and interested landowners. The final locations were selected based on their wind energy potential and to accommodate the existing stock of wind capacity (Lasher 2008, 2014). Several of the wind quality variables account for the wind energy potential of a county. Pre-CREZ wind capacity matches counties based on the existing stock of wind capacity, which was a key factor in selection of CREZ counties.

One of the unobservable factors is whether certain counties lobbied for or against siting of CREZ lines. While opposition is likely not a major concern in West Texas due to low land costs and minimal community opposition, it is certainly a concern for East and South Texas, where some of the lines were closer to urban areas (Andrade and Baldick 2016). In contrast, certain counties in the Panhandle region expressed interest to the PUCT for CREZ investment. This was in part due to an already declining population and economic loss in these counties in the years preceding CREZ expansion (Cohn and Jankovska 2020).

I construct a set of 'opposing' and 'enthusiastic' counties by reviewing individual cases filed by counties to PUCT and information from Cohn and Jankovska (2020).[35] These filings led to hearings and negotiations between county officials and PUCT regarding CREZ locations. I run the matching algorithm by excluding these two sets of counties separately from the original sample. The regression results for the new matched samples are reported in Appendix F.1 and are qualitatively similar to the baseline estimates in Table 4. I also conduct a series of robustness checks in Appendix F.3 to explore

35. The 'opposing' counties are: Kendall, Gillespie, Newton, Kimble, Kerr, Mason, and Schleicher. The 'enthusiastic' counties are: Dallam, Sherman, Oldham, Swisher, Lipscomb, Parmer, Lamar, Hall, and Deaf Smith.

how the coefficient estimates change when excluding some control variables, group fixed effects, and matching weights. The results are similar to the estimates in Table 4.

2. Anticipation of CREZ announcement:

A potential source of bias in measuring the causal impact could be the anticipation amongst wind developers of the CREZ announcement in 2008. This would be reflected as a spike in investment in wind projects within CREZ counties in the years leading up to the transmission expansion announcement. Using the data on generator interconnection in Texas, I examine the existence of such an anticipation effect in Appendix D. The analysis does not show the existence of any anticipation of the announcement of grid expansion two and four years prior to the announcement date.

2. Investment spillover to control counties adjacent to CREZ:

As shown in Figure 12, some of the control counties are adjacent to the treated counties. A potential concern could be that these counties saw higher wind investment as a result of being adjacent to the CREZ counties. I address this concern by estimating the baseline specification in Equation 19 along with an indicator for control counties that are adjacent to CREZ. Coefficient estimates of the indicator measuring this spillover effect in Appendix F.2 shows that while there is a small positive spillover effect, it is statistically indistinguishable from zero.

4. Selection of CREZ locations based on multi-phase wind projects and extensions:

Another threat to identification could exist 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 was selected as a site for CREZ infrastructure because of the likely development of a project extension within that county in the near future. To address this concern, I examine the occurrence of post-2012 extensions of wind projects that started operating before 2008 within CREZ counties. Figure E10 in the Appendix shows that the existence of multi-phase wind projects and project extensions are not a cause of concern.

7 Implications of short- and long-run effects of grid expansion on wind curtailment

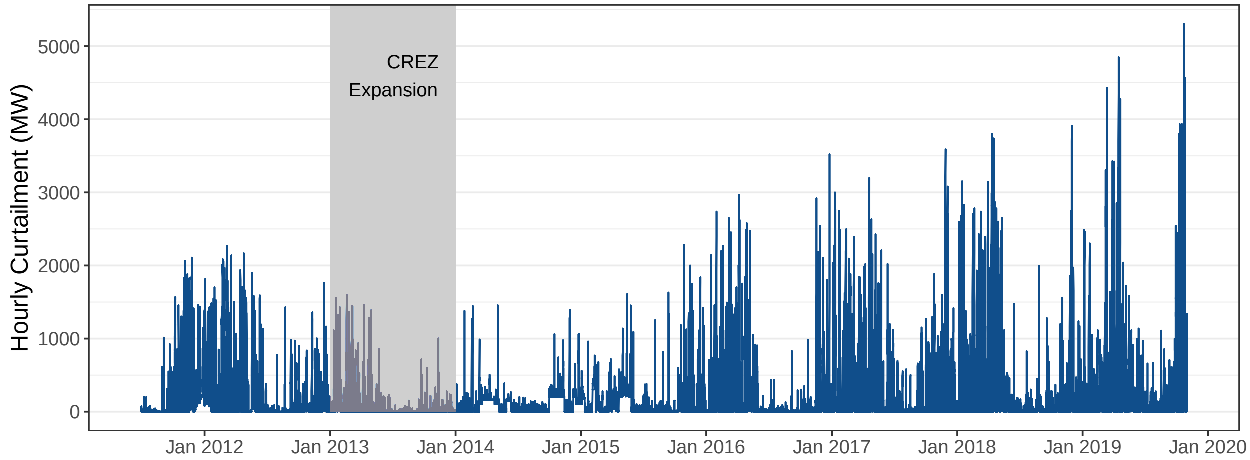
Electricity market operators typically curtail renewable resources during periods of congestion to maintain grid stability.[36] In Texas, the lack of adequate transmission capacity to transport electricity from wind farms in the West has been the primary source of wind curtailment, reaching about 17 percent of total wind generation in 2009 (Bird et al. 2014). Section 4 shows that, with generation capacity fixed in the short run, the availability of transmission capacity led to the integration of wind that would have been curtailed.

Figure 13a shows that CREZ expansion led to a significant decline in wind curtailments post-2014, but a steady rise since 2016. Further, Figure 13b shows that average hourly curtailments in 2019 were higher than pre-grid expansion levels in 2011 and 2012. In the long-run analysis (Section 6), locations that received investment in CREZ infrastructure saw higher levels of wind investments in the long run. Even though wind capacity in Texas has been increasing, there have not been any significant grid expansion projects post-CREZ.

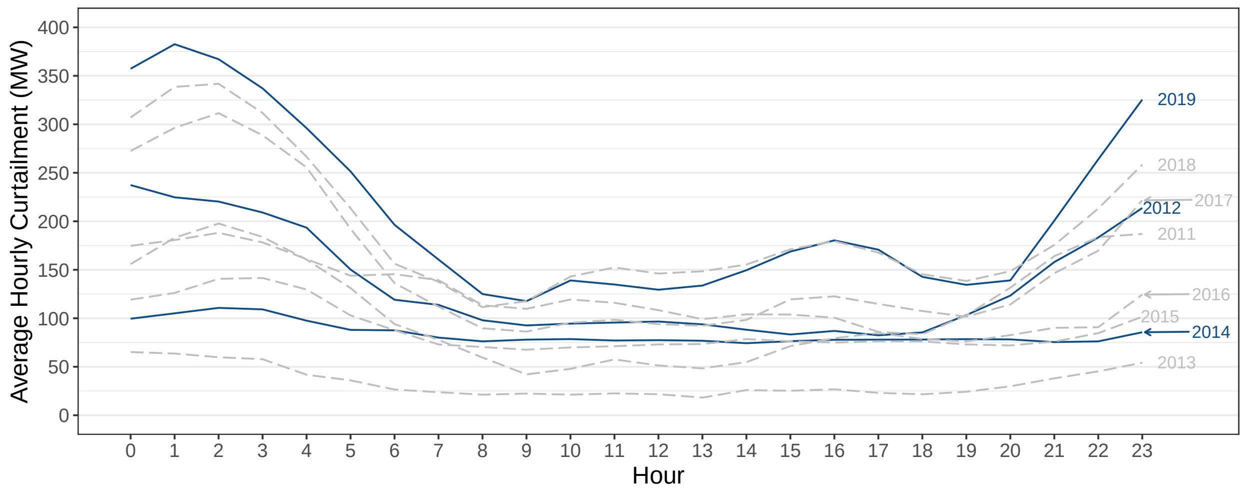
The rise in curtailment could be an outcome of localized investment in wind in the West and inadequate transmission capacity. I provide descriptive evidence by comparing curtailment in wind farms near CREZ counties to those farther away. I estimate the following two-way fixed effects specification at the quarterly level:

$$y_{it} = \sum_{\substack{k=Q2/2011 \\ (\neq Q4/2013)}}^{Q4/2019} \gamma_k \cdot \mathbb{1}\{\text{in/adjacent to CREZ}\} + \alpha_i + \delta_{qy} + \epsilon_{it} \tag{21}$$

36. As noted above, wind curtailment is the reduction in electricity generated from a wind generator below the level it could have produced given available resources (Bird et al. 2014). For example, suppose a wind generator is estimated to produce 100 MW of electricity in a period t but is finally scheduled to produce 80 MW. In that case, the corresponding wind curtailment is 20 MW. Curtailment typically is involuntary on the part of the generator. ERCOT determines the extent of curtailments based on transmission limits.



(a) Total hourly wind curtailment



(b) Average hourly wind curtailment

Figure 13: Wind curtailment in Texas from 2011 to 2019

Note: Figure 13b shows average hourly wind curtailments for each hour from 2011 to 2019. For clarity, solid lines highlight the curtailment pattern pre-CREZ expansion (2012), post-CREZ expansion (2014), and for the most recent year (2019) in the sample.

where y_i is the curtailment in wind farm i in hour t.[37] The parameter of interest γ_k measures the percentage difference in curtailment in wind farms within or adjacent to CREZ counties, compared to those located elsewhere, for each quarter in 2011 to 2019, with the fourth quarter of 2013 as the reference.[38] I include wind farm (α_i) and quarter

37. I use inverse hyperbolic sine (IHS) transformation of the dependent variable to account for the significant mass of zeros in the dependent variable.

38. A potential concern in the identification of γ_k is the variation in curtailments in wind farms in regions outside CREZ over the years. To explore this, I compare the trend of average monthly curtailment in regions near vs. outside CREZ for off-peak and peak hours in Figure E11. I show that the average

of the year (δ_{qy}) fixed effects. I estimate Equation 21 separately for off-peak [22:00 - 7:00) and on-peak [7:00 - 22:00) hours.

Figure 14 shows the estimates of γ_k from Equation 21. Curtailment was significantly higher in wind farms near CREZ counties in the years leading up to transmission expansion in 2014, especially in off-peak hours. For instance, in 2012, curtailment in wind farms in these regions reached about 1.5 times that of wind farms elsewhere. We notice a decline in curtailments after transmission expansion in 2014.

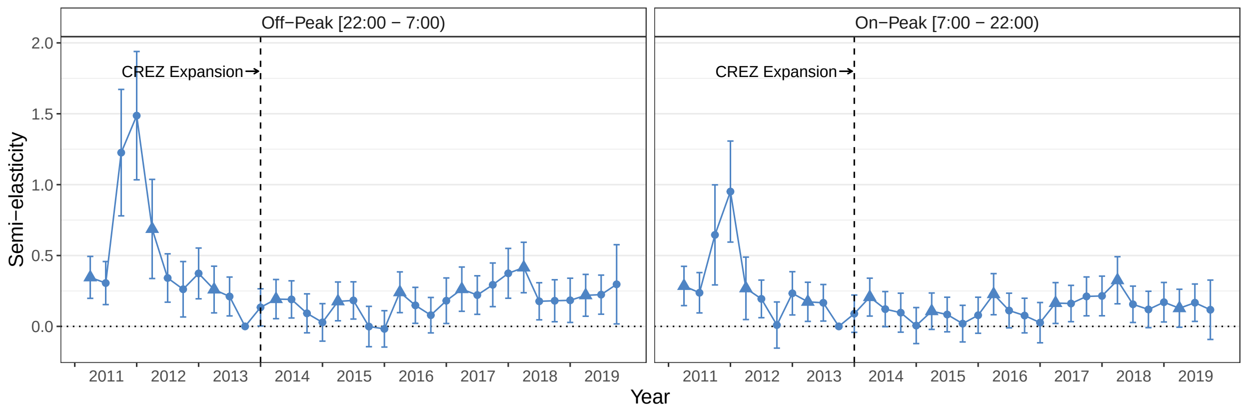


Figure 14: Percentage difference in curtailments between wind farms near CREZ counties and those in other regions

Note: This figure shows the estimates of γ_k from Equation 21. Each coefficient estimate shows the percentage change in curtailment between wind farms near CREZ counties to those in other regions for off-peak and on-peak hours over 2011 to 2019. Triangles highlight the coefficient estimates corresponding to the windier spring quarter (April - June) in Texas.

However, since 2017, wind farms near CREZ counties have seen a steady rise in curtailments, upward of 25 percent in the off-peak hours. This effect is both economically and statistically significant. Rising wind investment but inadequate transmission capacity could erode some of the short-run benefits from CREZ expansion. For instance, grid congestion during periods of high demand can incentivize fossil fuel firms to set higher markups. Similarly, inability to transport low-cost electricity from wind during high wind generation could lead to negative wholesale prices in the wind-rich West, thereby reducing the value of renewable investment in these regions.

wind curtailment in regions outside CREZ has remained almost stable over the years, while the average curtailments near CREZ counties exhibit a U-shaped trend. Therefore, γ_k is identified by the variation in curtailments in wind farms near CREZ counties.

8 Conclusion

Efforts to combat climate change in the US have primarily focused on expanding solar and wind generating capacity. A critical factor in fully utilizing the benefits of renewable energy is the availability of electricity transmission lines. Using the CREZ transmission expansion in Texas as a case study, this paper studies the short- and long-run impacts of large-scale grid expansion. I examine the effect of grid expansion on price markups and emissions associated with marginal fossil fuel generators in the short run and the effect on wind investment in the long run.

The short-run analysis shows that CREZ expansion led to lower market power and emissions from marginal fossil fuel producers. The decline in market power prevented about $230 million worth of annual transfers from consumers to producers and $60 million worth of emissions from marginal generators. These short-run effects are complementary to several other benefits estimated in the literature. These include gains from the trade of low-cost electricity (LaRiviere and Lyu 2022), emission benefits due to lower congestion (Fell et al. 2021), and enhanced grid reliability, to name a few.

In the long run, counties with investment in CREZ transmission saw significant investment in wind capacity (+202%) from 2012 to 2019. The added wind capacity prevented approximately $271 million worth of carbon emissions in Texas in 2019. While CREZ reduced wind curtailment in the short run by integrating wind into the grid, growing wind investment near CREZ counties has led to a steady rise in wind curtailment, indicating inadequate transmission capacity. This can have several market impacts, such as higher market power due to grid congestion, which could erode some of the estimated short-run benefits.

The benefits of CREZ expansion are multifaceted and dynamic. While the cost of this expansion was $6.8 billion incurred over three years, the benefits, as shown in this paper, are spread over a much longer time horizon. Assuming the estimated benefits to be static and additive indicates a payback period of 12 years. Alternatively, the payback period is about 15 years when accounting for long-run effects to accrue post-2019 and then remain static. The dollar value of some of these benefits is also dependent on the specific social cost of carbon. Using the more recent Social Cost of Carbon (SCC) estimate of \$185/ton-CO_2 (Rennert et al. 2022) suggests a payback period as small as 5 - 9 years. Moreover, these payback periods are conservative, given that I only consider a specific set of benefits, and some of the benefits are likely to evolve over time.

Grid expansion is a long-term public investment, wherein the benefits accrue over time. Because these are costly undertakings that take several years of planning and execution, quantifying both the short- and long-run effects is crucial to accurately assess the economic value of these investments. Finally, these findings also provide insights about transmission expansion in regions beyond Texas. For instance, grid expansion can unlock significant investment in renewable resources in the Midwestern and Southwestern US, while lowering market power and emissions from the fossil fuel sector in the short-run.

References

Andrade, Juan, and Ross Baldick. 2016. "Estimation of transmission costs for new generation." Energy Institute, University of Texas at Austin.

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.

Bird, Lori, et al. 2014. *Wind and solar energy curtailment: Experience and practices in the United States.* Technical report. National Renewable Energy Lab.(NREL), Golden, CO (United States).

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.

Deryugina, Tatyana, et al. 2020. "The long-run dynamics of electricity demand: Evidence from municipal aggregation." *American Economic Journal: Applied Economics* 12 (1): 86–114.

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.

ERCOT. 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.

Gonzales, Luis E, et al. 2022. "The Value of Infrastructure and Market Integration: Evidence from Renewable Expansion in Chile." Working Paper.

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.

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

Joskow, Paul L, and Jean Tirole. 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).

LaRiviere, Jacob, and Xueying Lyu. 2022. "Transmission constraints, intermittent renewables and welfare." *Journal of Environmental Economics and Management.*

Lasher, Warren. 2008. "The development of competitive renewable energy zones in Texas." In *2008 IEEE/PES Transmission and Distribution Conference and Exposition,* 1–4. IEEE.

———. 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.*

National Academies of Sciences, Engineering, and Medicine. 2021. "Accelerating Decarbonization of the US Energy System." Washington, DC: The National Academies Press. https://doi.org/10.17226/25932.

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.

NREL. 2008. "Texas $4.9 Billion Transmission Plan to Tap Wind." Retrieved from: https://www.nrel.gov/news/program/2008/620.html, *National Renewable Energy Laboratory.*

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

Rennert, Kevin, et al. 2022. "Comprehensive Evidence Implies a Higher Social Cost of CO2." *Nature,* 1–3.

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.*

US Interagency Working Group on Social Cost of Carbon. 2021. *Technical Support Document: Social Cost of Carbon, Methane, and Nitrous Oxide: Interim Estimates under Executive Order 13990.* 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 Details of CREZ Expansion Planning

The process of identifying the locations and cost of CREZ began following the enactment of the Texas Senate Bill 20 in 2005. In April 2008, ERCOT submitted a transmission optimization study that delineated four scenarios of transmission expansion (ERCOT 2008). These scenarios were expected to integrate the existing wind capacity of 6.9 GW by the end of 2008 and varying levels of projected wind capacities to be added until 2012. These scenarios differed widely in total cost and amount of wind the resulting transmission infrastructure could accommodate by 2012. Scenario 1A was expected to cost $2.95 billion and accommodated 5.15 GW of additional wind; Scenario 1B, was deemed more scaleable with a cost of $3.78; Scenario 2 was projected to cost $4.95 billion and accommodate 11.5 GW; Scenario 3 would accommodate 17.9 GW at a cost of $6.38 billion; and Scenario 4 would accommodate 17.5 GW wind with a total cost of $5.75 billion. These scenarios were evaluated based on three main objectives in ERCOT's transmission optimization study:

1. All of these scenarios would integrate existing wind capacity of 6.9 GW in West Texas.

2. The overall wind curtailment due to transmission congestion would be no more than 2 percent (curtailment as a share of total wind generation). For each scenario, curtailments on existing and planned wind facilities upto 2012 were considered.

3. ERCOT adopted an incremental approach to transmission planning that would essentially "overlay" the new CREZ lines on the existing grid in West Texas. In other words, the new system would not even be indirectly connected to the existing grid in West Texas. This was done in order to prevent widespread congestion and overloads in the existing low voltage system due to additional wind generation in the West and Panhandle region.

B.3 Transmission congestion and market power

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.[39]

Following example illustrates how inadequate transmission can prevent ERCOT from dispatching the cost effective generating units and incentivize them to exercise market power. Consider 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. Since electricity demand is fairly inelastic in the short-run, high cost generators could exercise market power by charging a price for electricity that is well above their marginal cost of generation. Note that 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.

39. 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.

C Conceptual model of wind project location choice

This section presents a simple conceptual model to build intuition on a wind developer's location choice for its wind project. Wind developer i choose location j to site their wind projects in order to maximize present value of annual profits written as:

$$\pi_{ij} = p_i \mathbb{E}(Q_j) - F_{ij} - OM_{ij} \tag{22}$$

where, p_i is the per MWh price that the wind farm receives, $\mathbb{E}(Q_j)$ is the expected electricity production from the wind farm which is a function of wind resource quality and the number and types of turbines. F_{ij} are the fixed costs and OM_{ij} are the operations and maintenance costs associated with the project.

The location choice is dependent on availability and access to transmission lines K at site j. Access to transmission lines is necessary for the wind farm to be able to deliver its electricity to the grid. Therefore, for two locations with similar wind quality, profits would be higher at the location with better access and availability of transmission lines,

$$\therefore \mathrm{K}_j > \mathrm{K}_{j'} \implies \pi_{ij}(\mathrm{K}_j) > \pi_{ij'}(\mathrm{K}_{j'}) \tag{23}$$

Next, the developer considers how far to locate from the electrical substation corresponding to the grid. [40] To see this, consider the profit function in Equation 22:

$$\pi_{ij} = p_i \mathbb{E}(Q_j) - \underbrace{[C_i + \kappa_j \cdot l]}_{\text{fixed costs}} - OM_{ij} \tag{24}$$

The fixed costs is a combination of two main components. The first is C_i, fixed costs incurred in building the wind project (like purchasing wind turbines), and second is the cost of constructing a spur transmission line, denoted by $\kappa_j \cdot l$. Spur transmission line is

40. Electrical (step-up) substations increase the voltage of electricity generated by power plants in order to make it efficient for transmission using long distance transmission lines. Therefore, these substations typically serve as the point of injection of electricity from the power plants into the grid.

a relatively short transmission line that connects the generator to the bulk transmission grid (Andrade and Baldick 2016). The cost of building spur lines is borne by the developer of the project. The schematic in Figure C1 illustrates the cost allocation of spur lines and bulk transmission lines between developer and end use consumers of electricity in Texas.

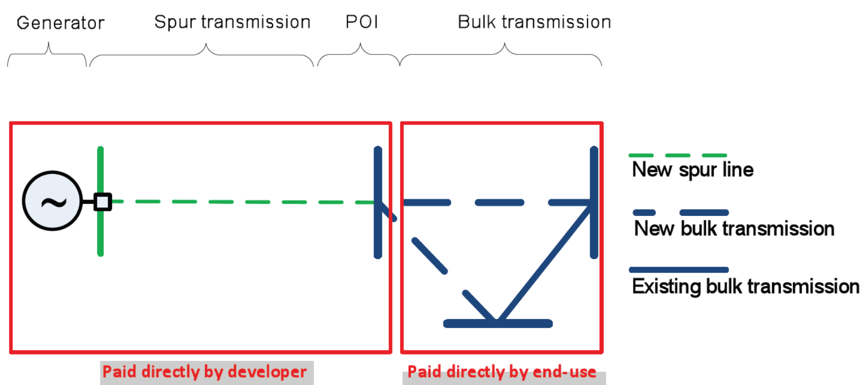


Figure C1: Illustration of transmission cost allocation in Texas for a new generation project. Source: Andrade and Baldick (2016)

The length of a spur line in Equation 24 is denoted by l (> 0) and κ_j is a positive cost multiplier which summarizes the costs associated with building a unit length of spur line (of a specific voltage) at location j. These costs are mainly due to land prices, terrine features, and generation technology (example wind, coal, natural gas). Partially differentiating π_{ij} with respect to length l shows that profits are decreasing in spur line length, i.e.

$$\frac{\partial \pi_{ij}}{\partial l} = -\kappa_j < 0 \tag{25}$$

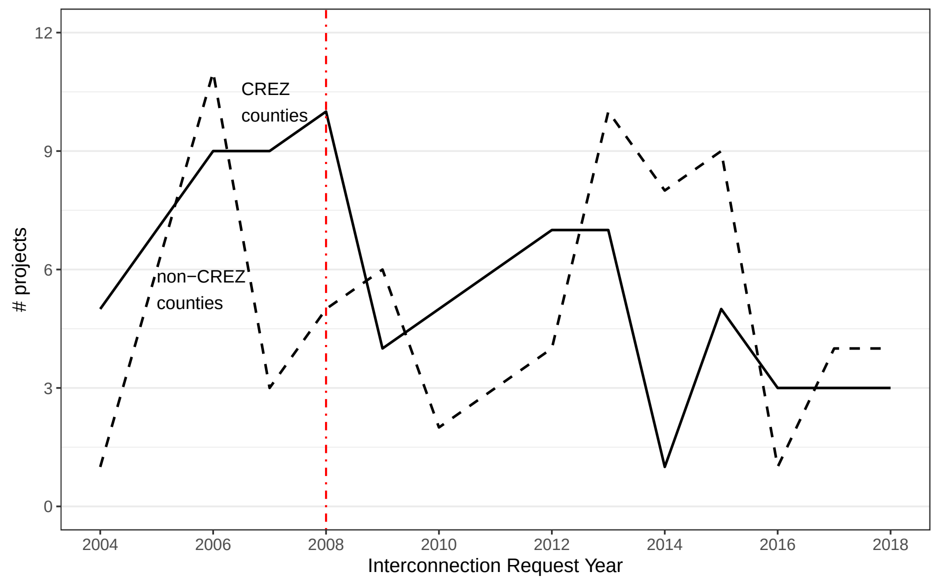
therefore, wind developers have an incentive to locate near the substation associated with the bulk transmission grid in order to maximize profits (or minimize costs). The simplified model shows that wind developers site their project in a region with access and availability to the grid, and then tend to locate near the grid substations to minimize the costs of building the spur transmission line.

D Anticipation Effects

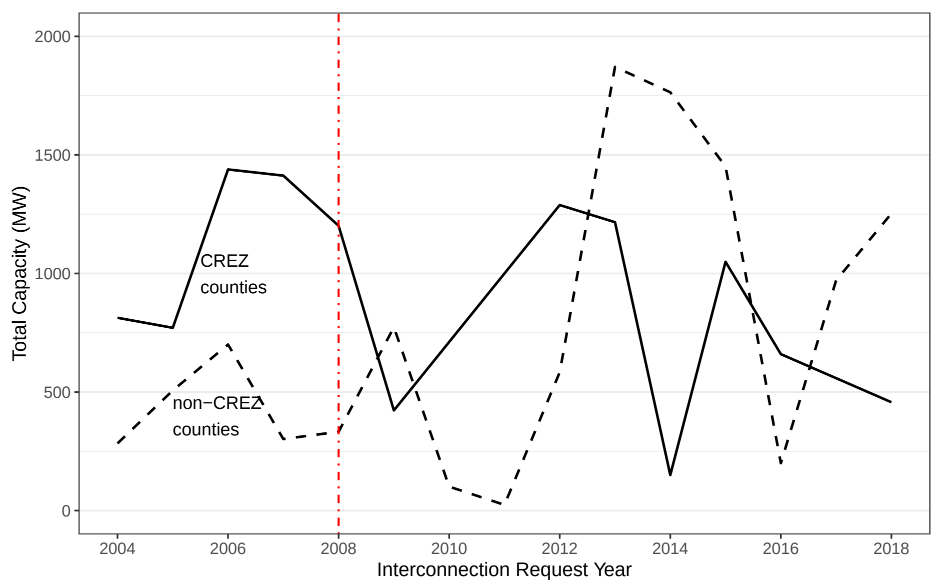
In this section, I examine whether there was an anticipation amongst wind developers to invest in wind projects in the period leading upto the announcement of CREZ transmission expansion in late 2008. Existence of such an anticipation could lead to biased estimates of the impact on CREZ announcement on wind investment in **??**. 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 D1 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 D1a and Figure D1b 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 D1: # 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:

$$y_{it} = \alpha_i + \beta \cdot \mathbb{1}\{year \in [k,\ 2008]\} + \mathbf{X}'\Pi + \epsilon_{it} \tag{26}$$

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 26 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 D1 reports the results of OLS regression of Equation 26 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 D2 reports the results of OLS regression of Equation 26 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 D1, 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 D1: 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 ∈ [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 ∈ [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 ∈ [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 ∈ [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 D2: 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

E Supplementary Figures

E.1 Renewable projects (solar and wind) in the US and county level population density

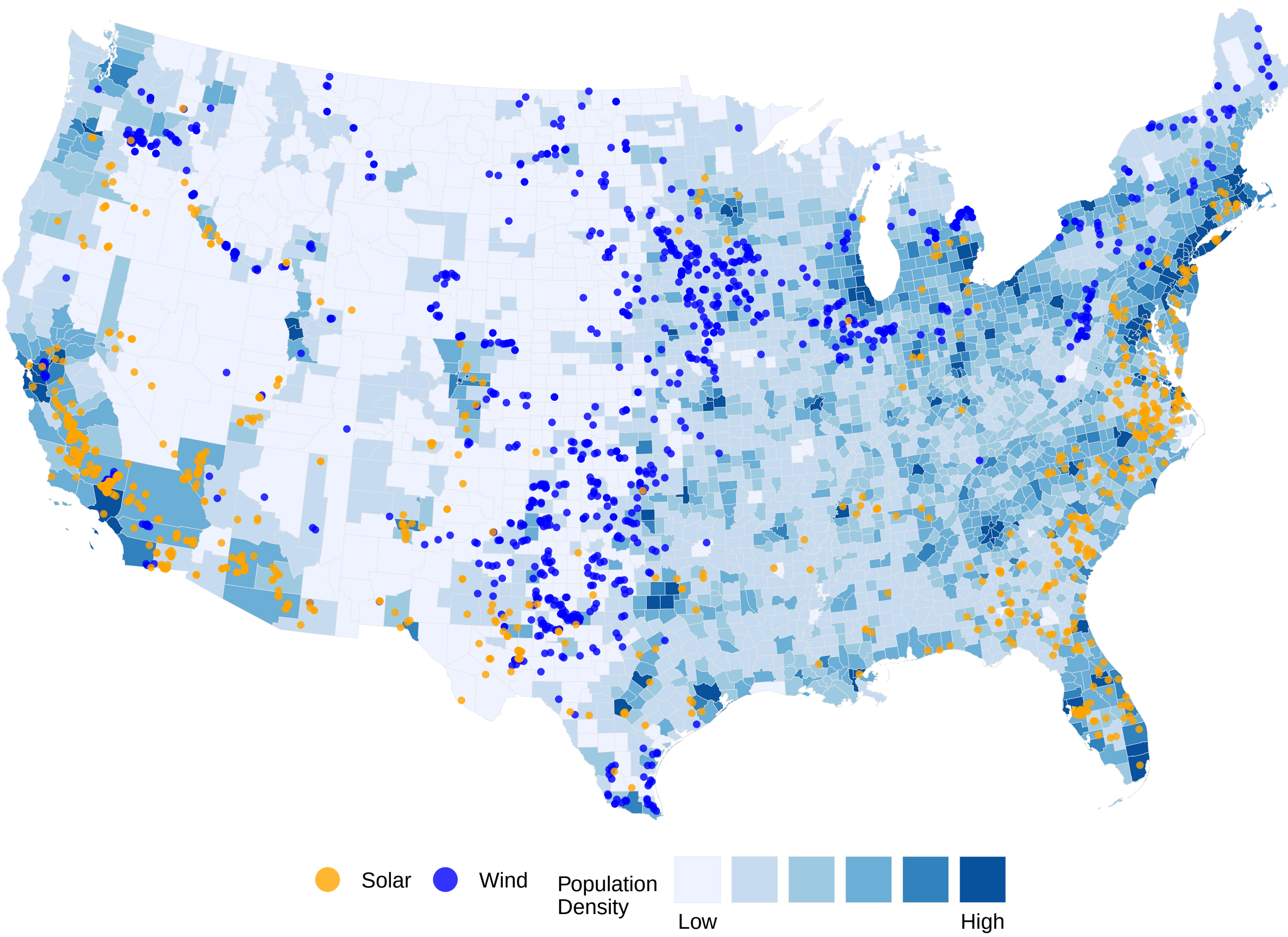


Figure E1: All solar and wind projects with $\geq$ 10 MW of nameplate capacity that started operation post 2001.

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]. This figure shows that most utility scale wind and solar power plants are located far from the demand centers.

E.2 Marginal cost of coal and natural gas generators used in the analysis

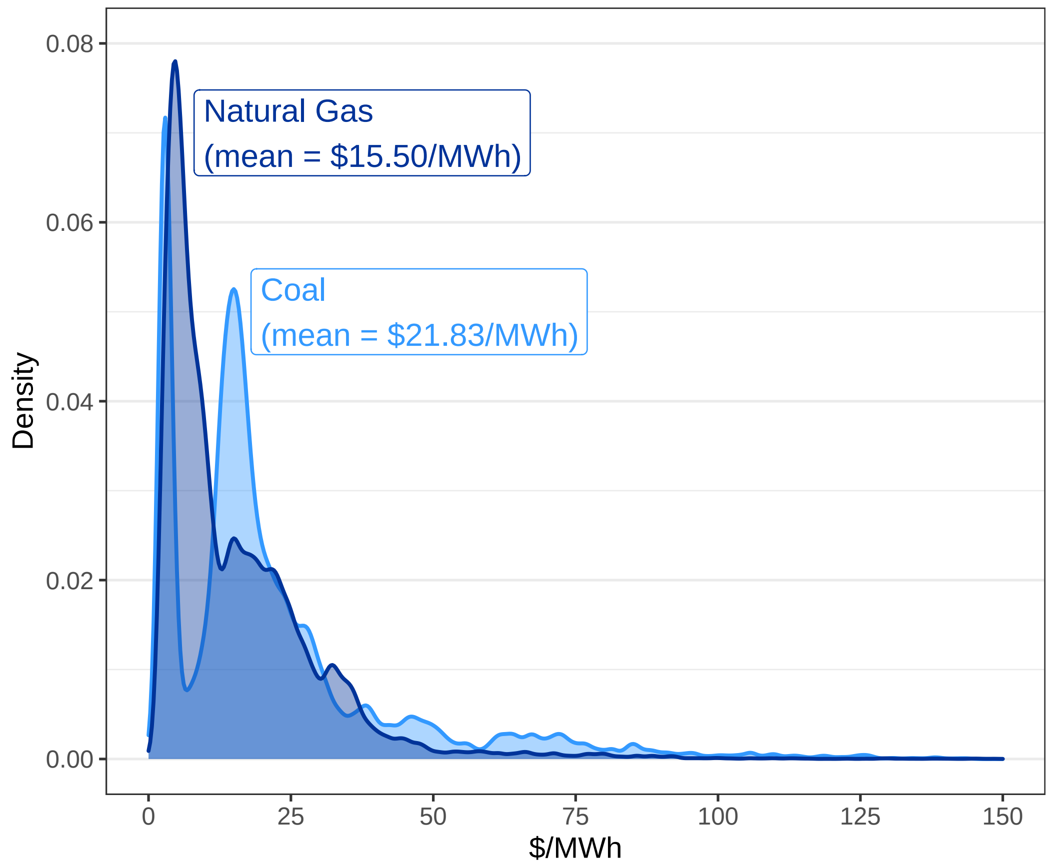
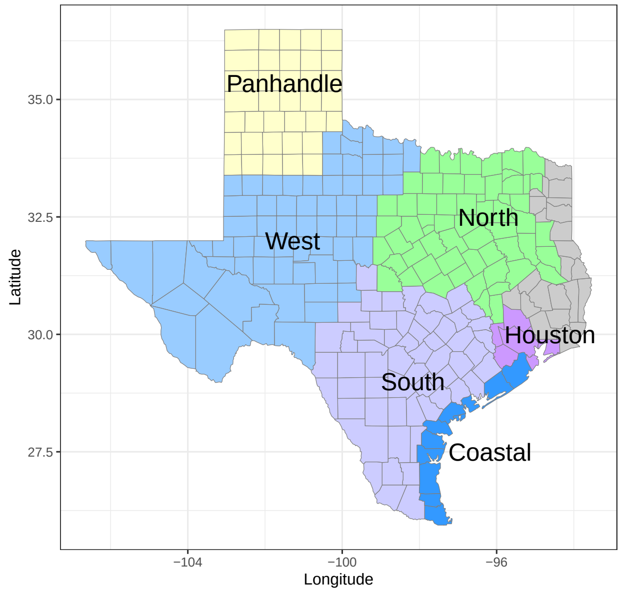
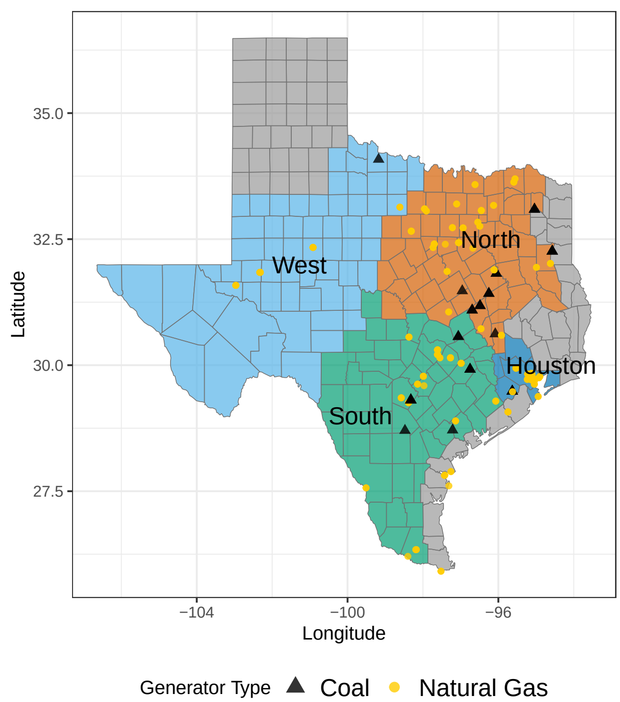


Figure E2: Distribution of marginal costs ($/MWh) of coal and natural gas generators

E.3 ERCOT Zones and marginal coal and natural gas generators



(a) ERCOT Zones Map



(b) ERCOT load zones and the locations of marginal coal and natural gas generators used in the emissions analysis

Figure E3: ERCOT load zones and marginal coal and natural gas generators

E.4 Illustration of theory model setup

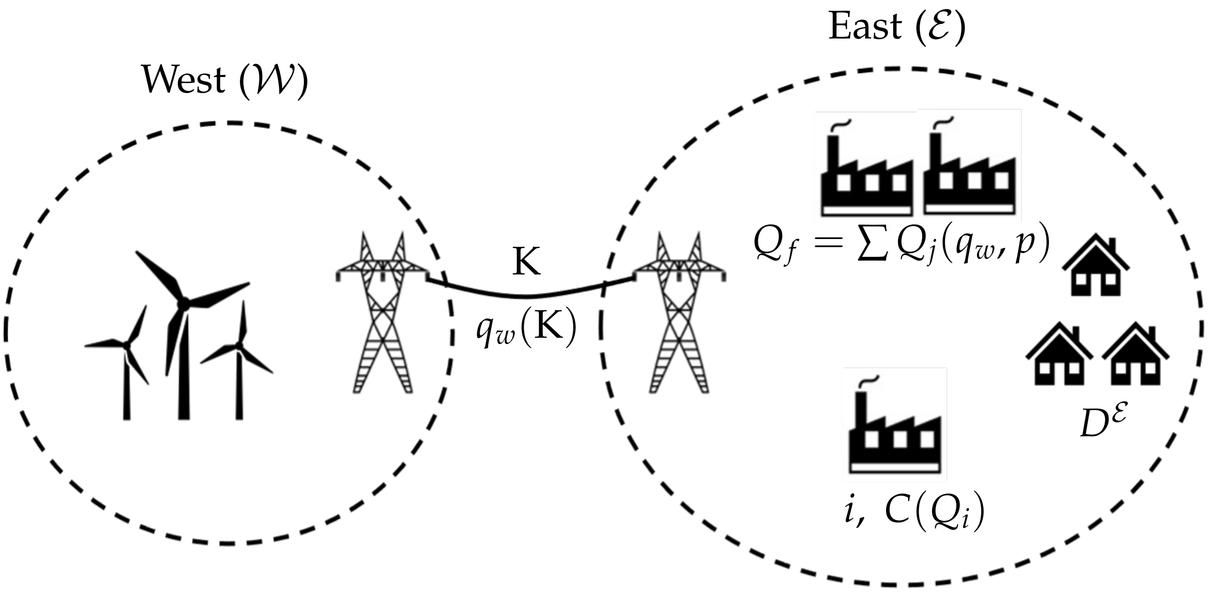


Figure E4: Theory model setup

Note: K denotes the transmission capacity between West ($\mathcal{W}$) and East ($\mathcal{E}$), $q_w(\mathrm{K})$ is the amount of wind generation transported into $\mathcal{E}$. $D^{\mathcal{E}}$ 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.

E.5 Robustness results for the impact of wind generation on markups

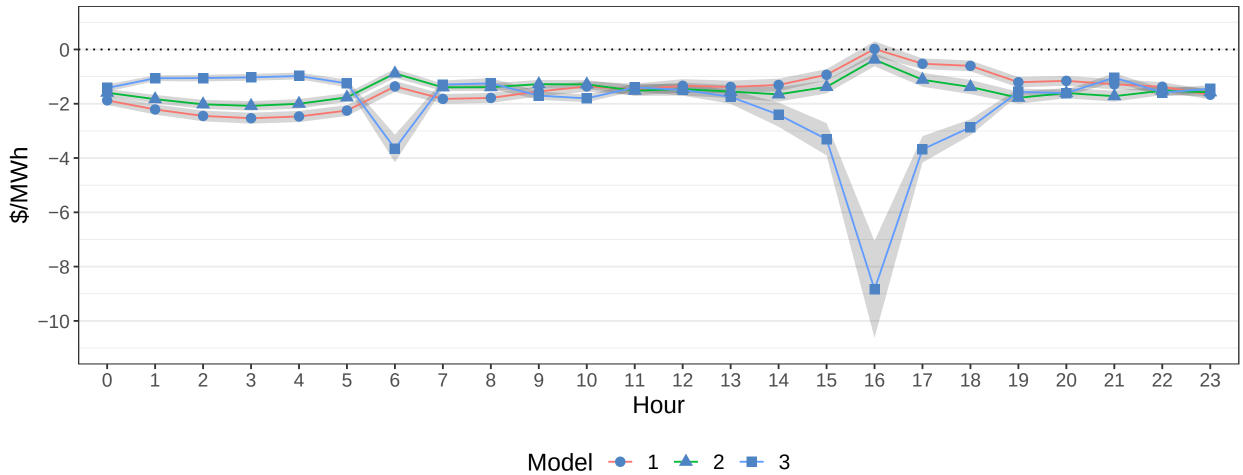


Figure E5: Effect of integration of 1 GWh of wind energy on marginal fossil fuel generator markups ($/MWh)

Sample: Hourly data from August 2011 to December 2014. 95 percent confidence intervals shown as shaded bands. Standard errors clustered at the generator level.

Model Specifications:

	(1)	(2)	(3)
Generator FE	✓	✓	✓
Load and Load^2		✓	✓
Hour × Month × Year FE			✓
Number of FE	284	284	1244
Observations	619,864	619,864	619,864
R^2	0.141	0.152	0.253

E.6 Robustness results for the impact of CREZ expansion on wind generation

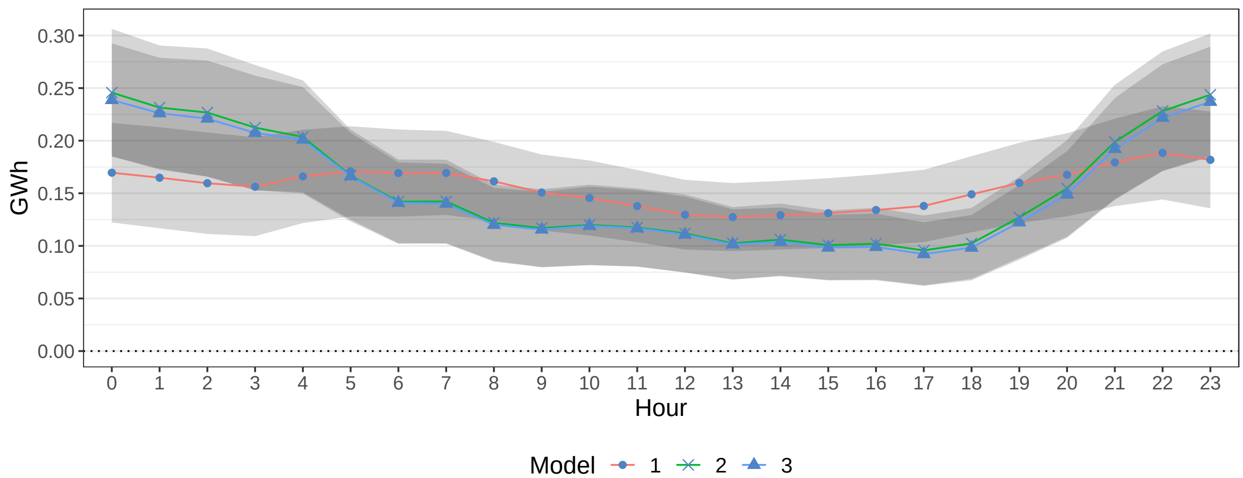


Figure E6: Effect of CREZ completion on integrating wind energy (GWh)
Sample: Hourly data from August 2011 to December 2014. 95 percent confidence intervals shown as shaded bands. Autocorrelation corrected standard errors computed using Newey-West procedure with 7 day lag structure.

Model Specifications:

	(1)	(2)	(3)
Max. predicted generation	✓	✓	✓
Hour FE		✓	✓
Hour × Month FE			✓
Number of FE	0	24	288
Observations	29,205	29,205	29,205
R^2	0.991	0.992	0.992

E.7 Impact of CREZ expansion on realized markups

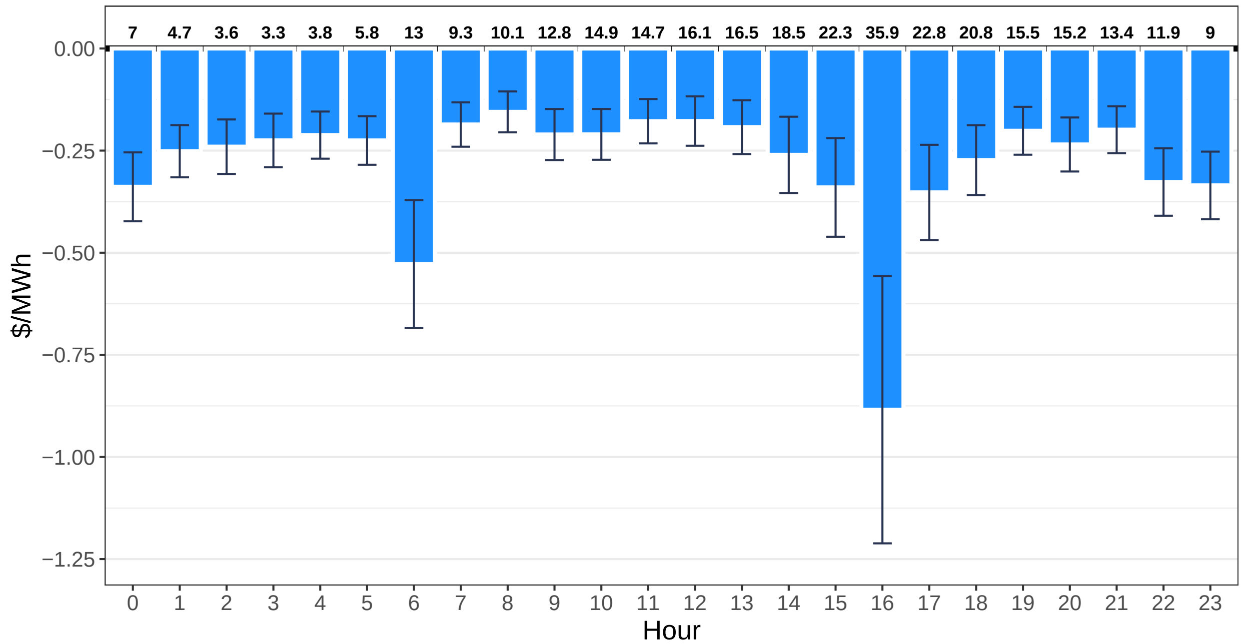
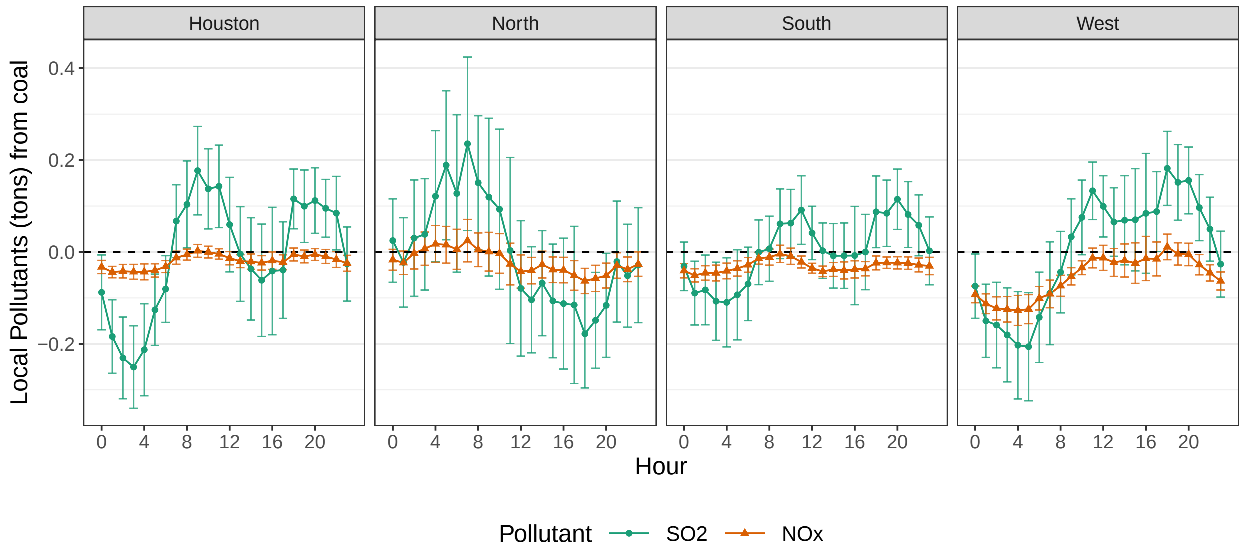
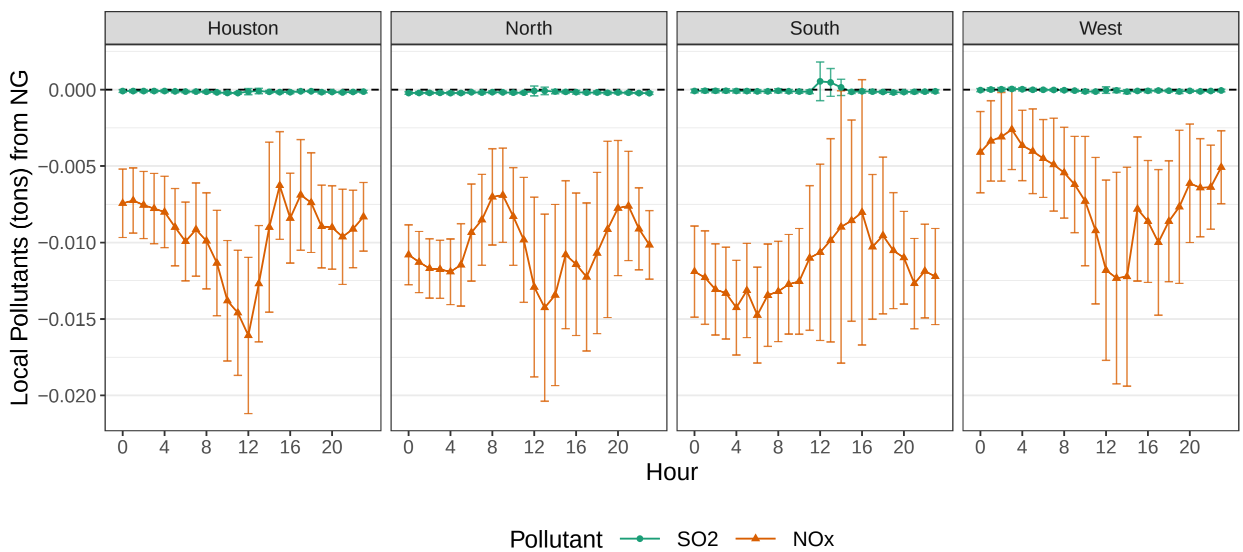


Figure E7: Impact of CREZ completion ($crez_d = 1$) on markups ($/MWh).

E.8 Wind generation and local pollution from marginal generators



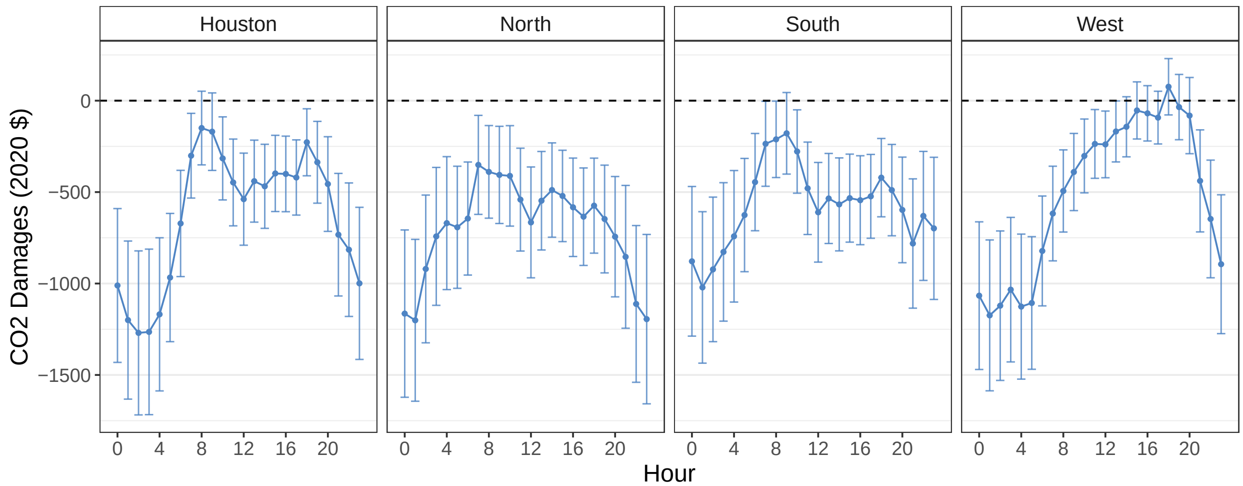
(a) Impact of wind generation on local pollutants (SO_2 and NOx) from coal generators



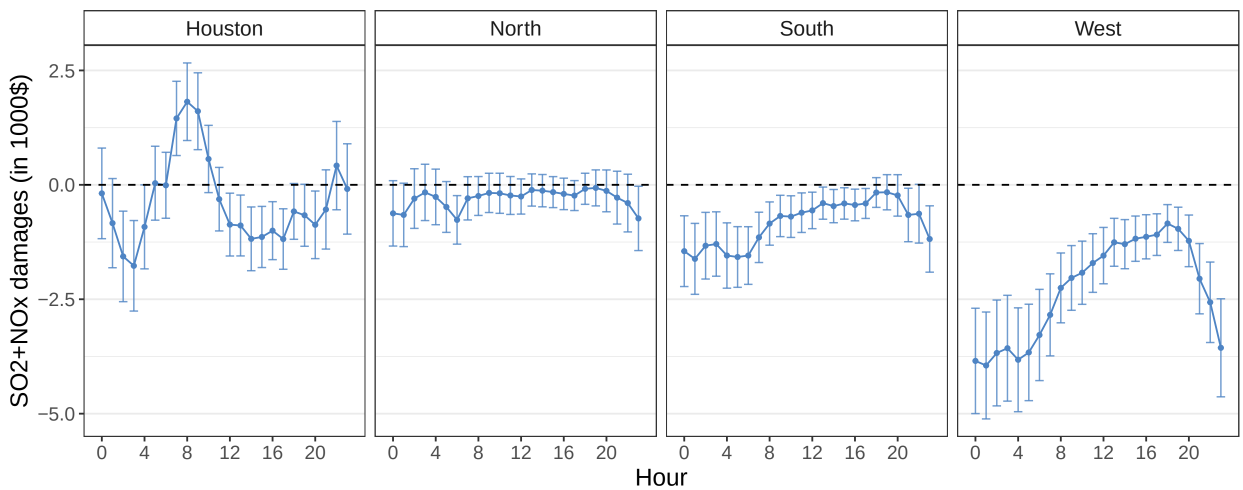
(b) Impact of wind generation on local pollutants (SO_2 and NOx) from natural gas generators

Figure E8: Short-run impact of wind generation on local pollutants (SO_2 and NOx) by generator type

E.9 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 E9: Hourly averages of the marginal damages (2020 $) avoided due to CREZ expansion for each zone over 2011 - 2014.

E.10 Existence of multi-phase wind projects and project extensions

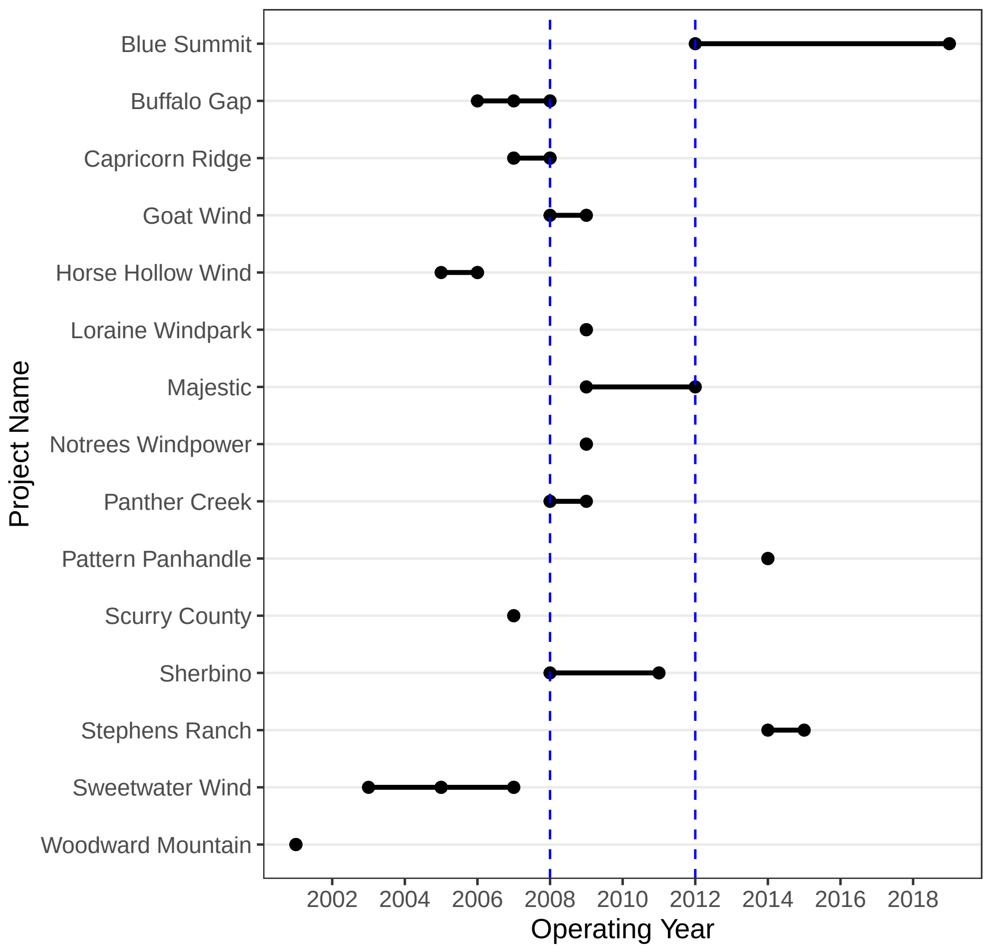


Figure E10: Wind projects with multiple phases and extensions

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).

E.11 Trend in wind curtailment

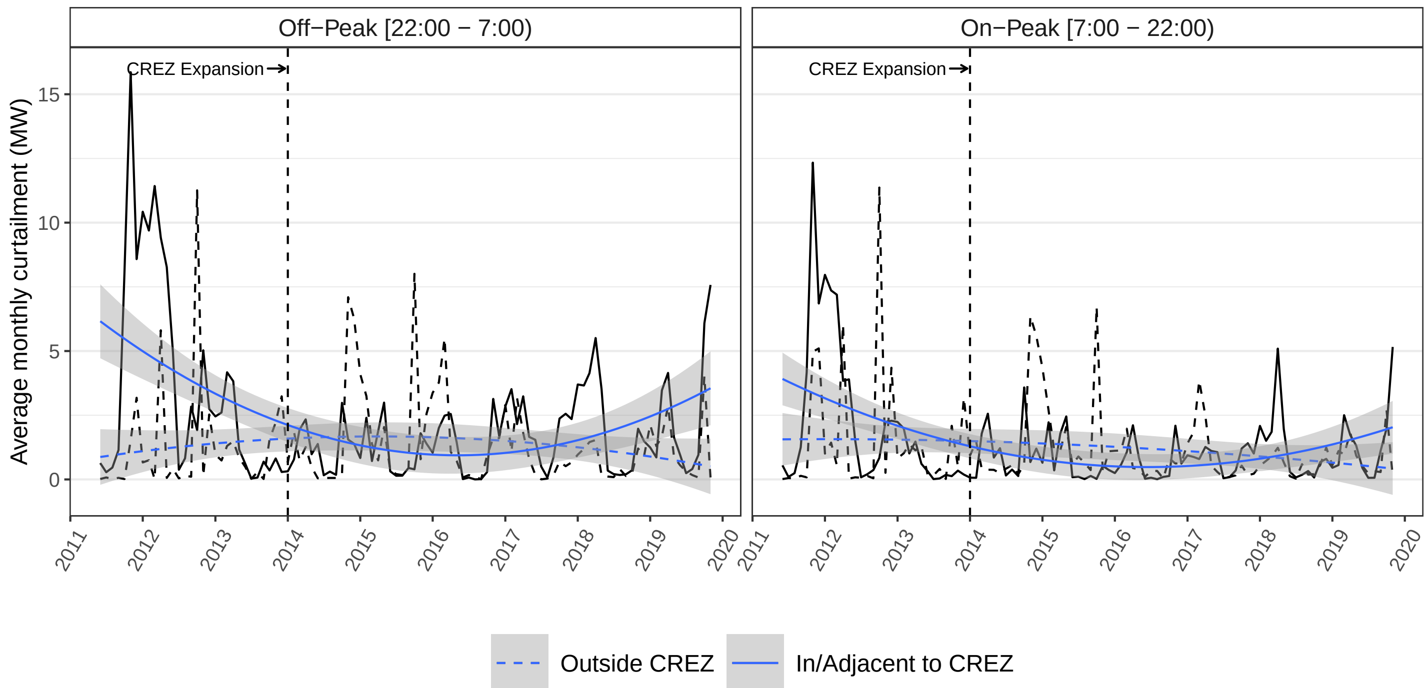


Figure E11: Average monthly wind curtailment (MW) by region and time of use (peak/off-peak)

Note: This figure compares the average monthly wind curtailment from 2011 to 2019 for wind farms in/adjacent to CREZ and those outside this region for off-peak (22:00 - 7:00) and peak (7:00 - 22:00) hours. Dashed vertical line marks the completion of CREZ expansion. Quadratic trend lines and the associated 95 percent confidence intervals are also shown. While the trend of average curtailment outside CREZ regions is almost flat over the years, curtailment in counties near CREZ shows a U shaped pattern. The trends overlap for a few year post CREZ expansion before diverging in 2017. All of these patterns are more pronounced during the off-peak hours when wind flow is high in Texas.

F Supplementary Tables

F.1 Robustness checks for matching on unobservables

F.1.1 Results excluding 'opposing' counties

Table F1: 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	67.29**	36.72***	28.16*
	(25.93)	(12.42)	(15.63)
Mean Dep. Variable	40.807	18.15	29.804
Semi-elasticity (%)	165.9	202.3	94.5
Controls	✓	✓	✓
Group × Trend FE	✓	✓	✓
Matching Weights	✓	✓	✓
Sample	Matched	Matched	Matched
Observations	280	280	280
R^2	0.489	0.505	0.471

Notes: This table reports the result of regressions excluding 'opposing' counties (Kendall, Gillespie, Newton, Kimble, Kerr, Mason, and Schleicher) from the overall sample before using Coarsened Exact Matching to obtain the matched sample. Total number of control counties is 23 and total number of treated counties are 12. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. All specifications include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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

F.1.2 Results excluding 'enthusiastic' counties

Table F2: 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	89.83***	48.53***	37.60*
	(30.60)	(14.78)	(18.74)
Mean Dep. Variable	36.636	16.484	26.761
Semi-elasticity (%)	245.2	294.4	140.5
Controls	✓	✓	✓
Group × Trend FE	✓	✓	✓
Matching Weights	✓	✓	✓
Sample	Matched	Matched	Matched
Observations	312	312	312
R^2	0.498	0.517	0.436

Notes: This table reports the result of regressions excluding 'enthusiastic' counties (Dallam, Sherman, Oldham, Swisher, Lipscomb, Parmer, Lamar, Hall, Deaf Smith) from the overall sample before using Coarsened Exact Matching to obtain the matched sample. Total number of control counties is 26 and total number of treated counties are 13. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. All specifications include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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

F.1.3 Results excluding 'opposing' and 'enthusiastic' counties

Table F3: 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	80.84***	44.47***	32.42*
	(29.04)	(13.43)	(18.84)
Mean Dep. Variable	41.033	18.348	29.000
Semi-elasticity (%)	197.0	242.4	111.8
Controls	✓	✓	✓
Group × Trend FE	✓	✓	✓
Matching Weights	✓	✓	✓
Sample	Matched	Matched	Matched
Observations	256	256	256
R^2	0.517	0.545	0.466

Notes: This table reports the result of regressions excluding 'opposing' (Kendall, Gillespie, Newton, Kimble, Kerr, Mason, and Schleicher) and 'enthusiastic' (Dallam, Sherman, Oldham, Swisher, Lipscomb, Parmer, Lamar, Hall, Deaf Smith) counties from the overall sample before using Coarsened Exact Matching to obtain the matched sample. Total number of control counties is 20 and total number of treated counties are 12. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. All specifications include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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

F.2 Investment spillovers to control counties adjacent to CREZ (treated) counties

Table F4: Regression results with an indicator for control counties adjacent to CREZ

	Dependent variable		
	Total Nameplate Capacity (MW)	Total Turbines	Avg. Capacity of a project (MW)
	(1)	(2)	(3)
CREZ	76.82**	42.49**	36.01*
	(33.45)	(16.08)	(21.11)
Adjacent to CREZ	4.25	3.24	9.20
	(27.49)	(12.99)	(19.47)
Mean Dep. Variable	35.907	16.067	26.951
Semi-elasticity (%)	213.9	264.4	133.6
Controls	✓	✓	✓
Group $\times$ Trend FE	✓	✓	✓
Matching Weights	✓	✓	✓
Observations	344	344	344
R^2	0.467	0.477	0.426

Notes: This table reports the estimate from Equation 19. 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 sited a substation for CREZ lines. 'Adjacent to CREZ' is an indicator specifying whether a control county is adjacent to a treated county. There are 17 adjacent control counties. All specifications include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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: ***$p<0.01$;**$p<0.05$;*$p<0.1$

F.3 Robustness check results with different specifications for full and matching samples

Table F5: Effect of CREZ on total wind capacity (MW)

	Dependent variable: Total Nameplate Capacity (MW)			
	(1)	(2)	(3)	(4)
CREZ	51.14**	43.04*	57.11	73.73**
	(24.31)	(22.60)	(34.43)	(29.45)
Controls		✓		✓
Sample	Full	Full	Matching	Matching
Mean Dep. Variable	33.069	33.069	35.907	35.907
Observations	2,024	2,024	344	344
R^2	0.027	0.221	0.061	0.467

Notes: The dependent variable is total wind capacity (MW) in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. Full 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. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. Specification in Columns (2) and (4) include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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 in Column (4) 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

Table F6: Effect of CREZ on total wind turbines

	Dependent variable: Total Turbines in a County			
	(1)	(2)	(3)	(4)
CREZ	27.49**	23.36**	31.36*	40.13***
	(12.74)	(11.82)	(17.77)	(14.46)
Controls		✓		✓
Sample	Full	Full	Matching	Matching
Mean Dep. Variable	15.928	15.928	16.067	16.067
Observations	2,024	2,024	344	344
R^2	0.033	0.209	0.081	0.476

Notes: The dependent variable is the total number of turbines in a county in year *t*. The independent variable is a binary variable indicating whether a county is CREZ or not. Full 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. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. Specification in Columns (2) and (4) include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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 in Column (4) to allow for time-varying unobserved factors affecting matching groups. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***$p<0.01$;**$p<0.05$;*$p<0.1$

Table F7: Effect of CREZ on size of a wind project

	Dependent variable: Average Capacity (MW) of a project			
	(1)	(2)	(3)	(4)
CREZ	19.64**	10.62	25.51	29.33
	(9.86)	(10.04)	(19.49)	(17.71)
Controls		✓		✓
Sample	Full	Full	Matching	Matching
Mean Dep. Variable	19.990	19.990	16.067	16.067
Observations	2,024	2,024	344	344
R^2	0.014	0.200	0.027	0.425

Notes: The dependent variable is the average capacity (MW) 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. Full 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. The independent variable is a binary variable indicating whether a county sited a substation for CREZ lines. Specification in Columns (2) and (4) include cubic polynomial of time trend and controls for wind quality, land price, terrain ruggedness, county level regulation, and demographics. Wind controls include power curve, capacity factor, and cubic polynomial of wind speed. Land price 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 in Column (4) to allow for time-varying unobserved factors affecting matching groups. Robust Standard Errors clustered at the county level reported in parenthesis. Significance: ***$p<0.01$;**$p<0.05$;*$p< 0.1$