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

Gaurav Doshi*

JOB MARKET PAPER

Work in progress. Please do not cite or distribute.

This version: July 9, 2021

Click here for most recent version

Abstract

Transitioning to an electricity grid that fully realizes the benefits from renewable energy will require substantial investment in transmission lines. This paper examines the short- and long-run effects of large scale grid expansion projects aimed at enhancing the integration of wind energy in the US. I focus on the rollout of a large scale transmission expansion project in Texas for my empirical analysis. Short-run analysis shows that transmission expansion led to a 2-2.5% decline in markups during the peak demand hours and a 7% decline during the off peak hours. Transmission expansion also prevented about $51 million worth of annual damages (2020 $) from marginal emissions in the short-run. In the long-run, transmission expansion on an average led to 62 - 72 MW higher wind capacity, 30 - 40 more turbines, and 32 MW bigger wind projects in counties that received investment in transmission infrastructure. Given a growing need for investment in grid expansion in the US, this paper provides evidence of significant market impacts in the short- and the long-run in response to grid expansion.

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

1 Introduction

The US aims to achieve a carbon-free power sector by 2035 (The White House 2021). Most wind and solar generators in the US produce electricity in geographic locations far from where electricity is consumed. Transitioning to an electricity grid that fully realizes the benefits from renewable energy will require substantial investment in transmission lines. This issue has been covered widely in both energy related and popular news outlets, pointing out the imminent necessity to build transmission lines in order to dramatically cut carbon emissions and achieve ambitions energy goals (New York Times 2016; Temple 2017; Wolff 2021). However, empirical literature in economics on the market and non-market impacts of large scale changes in transmission capacity has been limited.

In this paper, I study how large scale transmission expansion projects aimed at integrating wind energy affect firm entry, pricing behavior, and emissions. I first study the short-run effect of transmission expansion on fossil fuel generator markups and keeping generation capacity fixed. I also analyze the impact on emissions of global and local pollutants (SO_2 and NOx) by marginal generators in the short-run. I then study the long-run investment in wind energy in response to transmission expansion. To answer these questions, I focus on the rollout of a large scale transmission expansion project called Competitive Renewable Energy Zones (CREZ) in Texas. Costing about $7 billion, the CREZ transmission expansion was aimed at integrating the wind resources in West Texas with the major demand centers in other parts of the state. The construction of this project started in 2011 with more than 80 percent of the rollout happening in 2013.

There are several interesting features of the Texas electricity market that make it an excellent context for my research question. First, the roll-out of CREZ over 2011 - 2013 provides an ideal setting to study effects of large scale transmission projects. Second, there has been a gradual increase in the generating capacity of wind over the years. In 2011, wind represented about 13 percent of the total electricity generating capacity which grew to over 25 percent by the end of 2020 (ERCOT 2012, 2021). Finally, the Texas electricity interconnection is isolated from other electricity markets in the US, which avoids concerns about spillover effects due to imports or exports of electricity.

For the short-run analysis, I build a model of optimal bidding for a fossil fuel generator to understand how transmission line expansion affects fossil fuel generator markups. I write this model in the context of a uniform auction wherein the generator participates by bidding the price and quantity of electricity (as a function of price) that it is willing to supply in a particular period. The generator maximizes its profit over the residual demand curve to find the optimal price of electricity to supply. I specifically focus on the case of the marginal generator whose optimal bid determines the wholesale price. Therefore, the markups set by the marginal generator(s) is the realized markup prevailing in the wholesale electricity market. Solving the generator's optimization problem allows me to derive the optimal markup rule as a function of its residual demand curve.

The model yields several predictions on how large scale transmission expansion affects the marginal generator's markups. More generally, the ability of the generator to set markups depends on two factors. First, the extent to which wind based power displaces production from competing fossil fuel generators. Second, the degree to which wind generation affects the competition amongst fossil fuel generators in the market. The relative magnitudes and directions of these two effects determine whether the generator sets higher or lower markups.

Based on the theoretical model, I construct a two-step estimator to examine how transmission expansion affects markups set by a fossil fuel generator. In the first step, I estimate the effect of transmission expansion on hourly wind generation, followed by the effect of wind generation on hourly markups set by fossil fuel generators. I find that CREZ expansion led to an average decline in markups by about 2-2.5 percent during the peak demand hours and about 7 percent during off peak hours. However, the magnitude of decline in markups due to an additional GWh of wind is strongest at peak hours. Therefore, the potential of transmission capacity to affect markups depends on how much wind capacity is integrated into the grid.

The short-run analysis on the emissions from marginal generators finds a decline in emissions to the order of $51 million (2020 $) annually with the majority of the reduction coming from local pollutants (SO_2 and NOx). I find significant evidence of spatial heterogeneity in the damages avoided due to both carbon emissions and local pollutants. While the the decline in carbon emissions is spatially distributed across the four zones in Texas, the decline in local pollutants comes primarily from the West. Further, the estimates show an increase in emissions due to ramping up of marginal coal generators as a result of intermittency of wind during early hours of the day. This is especially pronounced in Houston for both carbon emissions and local pollutants.

Next, I analyze whether locations that receive investment in transmission infrastructure also see higher levels of wind investment in the long run. I implement an instrumental variable and a matching strategy to answer this question. I use the designation of certain regions in Texas into five Renewable Energy (RE) Zones in 2007 as an instrument for a county to ultimately receive transmission investment. The instrument is plausibly exogenous to long term wind investment because wind developers are unlikely to respond to such a broad regional characterization and the primary goal of identifying RE Zones was to narrow the choice set of probable transmission locations. The IV estimates suggest that CREZ expansion on an average led to about 62 MW higher wind capacity and 30 more turbines in treated counties over 2012 - 2019.

I use a matching strategy to alleviate the concerns of the support problem in comparing counties that are inherently different in the OLS and IV regressions. I use Coarsened Exact Matching to construct a set of control counties that are identical to the treated counties on a wide set of observable dimensions. The treatment effect for the matched sample suggests that the treated counties had 72 MW higher wind capacity, 40 more turbines, and about 33 MW bigger wind projects than the control counties. Taken together, these results on wind investment provide robust evidence that counties that received investment in electricity transmission also saw substantially higher investment in wind generation.

Data. I use a variety of publicly available data sources for the empirical analysis in this paper. I assemble a hourly level dataset on markups set by marginal fossil fuel generators using data from EIA From 860, Electricity Reliability Council of Texas (ERCOT), and the Continuous Emissions Monitoring System (CEMS) from EPA. I use data on SO_2 and NOx allowance prices and fossil fuel (coal and natural gas) prices from S&P Global Market Intelligence. To get daily data on progress of the CREZ project from 2011 through

2014, I use transmission planning reports provided by ERCOT. For the long-run analysis, I use data on wind projects from EIA Form 860 and wind resource data from NREL.

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

Second, I contribute to the growing literature focusing on the value of transmission infrastructure in mitigating market power in wholesale electricity markets. Theoretical studies in this area employ Cournot models and simulations to show how expansion in transmission capacity leads to higher competition and mitigates the effects of market power (Borenstein, Bushnell, and Stoft 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 effects of transmission constraints in exacerbating the effects of market power (Woerman 2019; Ryan 2021).

My theoretical model is most closely related to Ryan (2021) who derives the optimal bidding condition for a fossil fuel generator and uses structural estimation to simulate the effects of transmission expansion in the Indian electricity market. I add to this discussion by deriving conditions in my theoretical model under which renewable electricity generation due to transmission expansion could incentivize fossil fuel generators to set lower markups. Utilizing high frequency generator level data from Texas electricity market and detailed data on CREZ, I provide hourly estimates on changes in markups due to transmission expansion.

Third, this paper contributes to the nascent literature looking at the link between transmission and renewable energy in electricity markets. Recent papers study how transmission expansion in Texas enhanced the environmental value of wind measured by emissions avoided (Fell, Kaffine, and Novan 2021) and led to significant welfare gains by reducing wholesale price dispersion across different regions (LaRiviere and Lu 2020). I complement this literature by showing how intermittency of wind during the day leads to higher emissions from marginal generator(s) in certain regions. This could eat away some of the short-run gains due to transmission expansion. Utilizing the spatial data on transmission expansion, I analyze the long term investment in wind energy in response to transmission expansion, a novel contribution to the literature.

Finally, the analysis in this paper provides a framework to study short- and long-run impacts of transmission expansion in the power sector. The feedback effects from the long-run analysis are relevant from a policy standpoint due to the growing importance of upgrading and expanding the electrical grid so as to integrate renewable sources. Changes in markups in the short-run can also have long-run consequences in the sense that higher cost firms might retire due to lower markups over time. This could perhaps lead to lower competition amongst the existing generators thereby pushing the wholesale prices higher in the long-run.

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 the variables along with descriptive statistics in Section 3. The theoretical model for the short-run, empirical strategy, and results are presented in Section 4. The long-run analysis of CREZ expansion and wind project investment is presented in Section 6. Section 7 concludes.

2 Institutional Details

2.1 The Texas electricity market

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

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

The ERCOT interconnection is comprised of six zones within Texas - Panhandle, West, North, South, Houston, and Coastal. Figure 1 provides a sense of geographic distribution of counties in these zones. Wind rich Panhandle and West Zone is where most of the wind generation of Texas comes from. North and the South Zones host most of the fossil fuel generating firms along with major demand centers like Dallas-Fort Worth Area, San Antonio, and Austin. Houston forms the fourth zone which is a major demand center in itself. These zones are connected by a network of over 46,500 miles of transmission lines that carried about 74,820 MW of electricity at a record peak demand on August 12, 2019 (ERCOT 2021).[1]

2.2 Competitive Renewable Energy Zones

An interesting aspect of the Texas electricity market is the increasing capacity of wind based power generation. Since ERCOT schedules lowest cost generation to dispatch first, wind based generators are always scheduled to dispatch first. Thus conditional on wind flow, wind based generating units are always scheduled to meet the demand. Fossil fuel

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

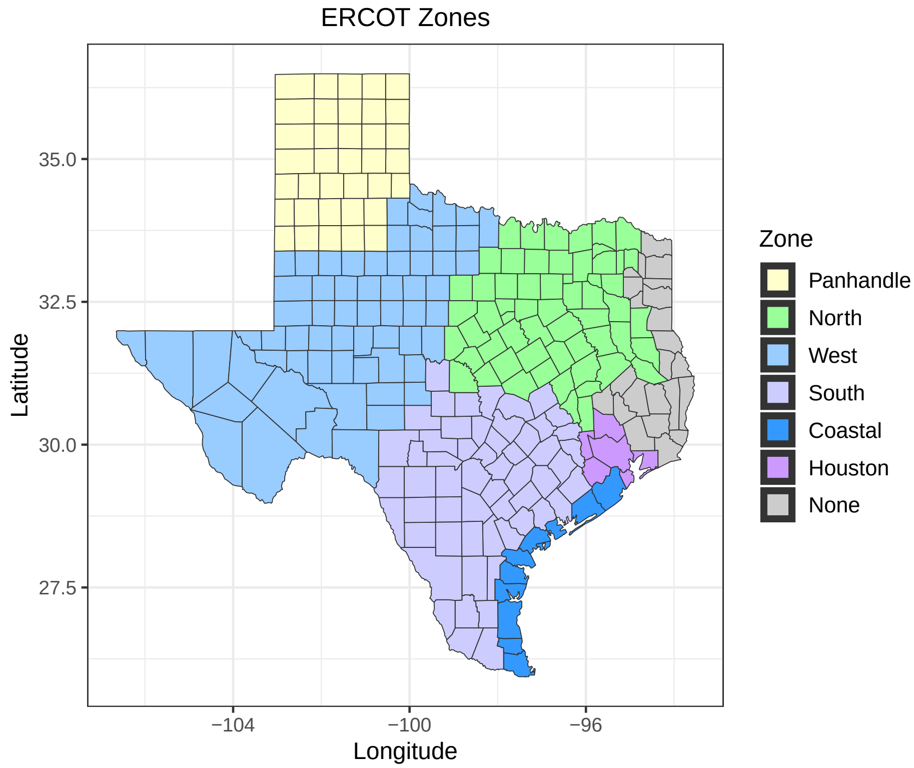


Figure 1: ERCOT Zones Map

generators on the other hand are dispatched to meet the remaining demand as well as to meet any sudden surge in demand at Peak hours.[2]

Inadequate transmission capacity between West and other Zones in Texas could lead to transmission congestion in a way that prevents trade of electricity from West to demand centers in other parts of the state. Recall that most of the wind based electricity generation in Texas comes from the West Zone. Presence of transmission constraints could cause ERCOT to schedule electricity from local generating units that are typically fossil fuel fired fired generators. This not only leads to CO_2 emissions that could've been offset by clean wind based electricity but also incentivize local fossil fuel generators to charge a markup over their marginal cost of production.

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

I examine this phenomenon under the backdrop of a recent transmission expansion project in Texas, Competitive Renewable Energy Zones (CREZ). CREZ was a large scale transmission expansion project aimed at integrating electricity generation from wind farms located in West to the major demand centers in North, South, and Houston Zones. The project, commissioned in 2008 by the Public Utilities Commission of Texas was aimed to accommodate over 18.5 GW of electric power by building about 3,600 circuit miles of 345 kV electricity transmission lines. However, the transmission lines are open access meaning that the use is not limited to only wind generators (Billo 2017).

Transmission lines were built over a period of 2011 through 2013 with a total cost of approximately $6.8 billion. All of the CREZ based facilities were brought in service by December 2013 and has been credited in reducing transmission congestion and wind curtailment (Lasher 2014). CREZ transmission expansion provides an excellent opportunity to study how large scale changes in grid affects pricing behavior of fossil fuel generators as well as investment in wind energy and exit of fossil fuel generators in the long-run.

3 Data and Variables

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

3.1 Electricity Generation and Demand

I use hourly electricity generation data at the generator level for fossil fuel generators and data on system wide electricity generation from wind from ERCOT. The electricity generation data is used to construct marginal costs and data on wind generation is a variable of interest in the empirical specifications. I also use hourly data on system electricity demand from ERCOT.

3.2 Markups

One of the main outcomes of interest for the short-run analysis is the generator markups. Markups are defined as $p - c$ where p is the Locational Marginal Price (LMP) and c is the marginal cost. I use publicly data available from ERCOT to identify the price setting (marginal) generators and the corresponding LMP at each hour of the sample. The other component of markup is the marginal cost of generation. As is common in the literature, I assume constant marginal cost of generation. The two major components of marginal cost are fuel costs and emissions permit costs based on emissions regulations for SO_2 and NOx. I calculate the marginal cost of each generator as the sum of these two components.

To compute fuel costs, I use weekly price data for coal and natural gas. For coal, I use Powder River Basin spot prices from EIA. For natural gas, I use Henry Hub Natural Gas prices from Quandl. I calculate fuel costs by multiplying fuel price and the heat rate (HR_i) of the generator.[3] To compute the emissions permit costs, I use daily data on NOx and SO_2 allowances from S&P Global Market Intelligence. Using hourly emissions data from EPA's Continuous Emissions Monitoring system (CEMS), I calculate the emissions rate (ER_i) for SO_2 and NOx by taking the ratio of emissions to net generation.[4] The generator's emission permit cost is thus the product of the permit price and emissions rate for each emission type. Thus, the marginal cost c_{it} of generator i in period t is:

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

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

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

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

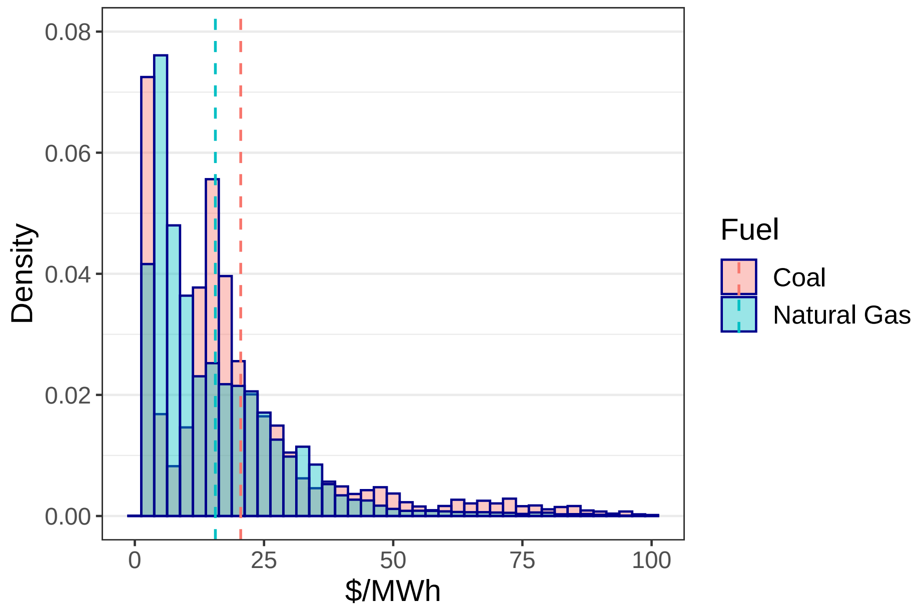


Figure 2: Distribution of marginal costs of Coal and Natural gas generators

Table 1 reports descriptive statistics of generator markups and nameplate capacity by fuel types of generators in the sample. About 70 percent of the observations in the sample are natural gas fired generators whereas coal generators are the remaining 30 percent. From Panel A. in Table 1, the average realized markups set by natural gas generators is about $17.16/MWh which is about three times the average markups of coal generators. This follows from the overall generation pattern in the Texas electricity market wherein coal generators tend to be marginal during the night and early hours of the day whereas natural gas generators operate at the margin during the peak demand hours. From Panel C. the coal generators have an average nameplate capacity of 603 MW whereas natural gas generators are 192 MW.

Generator markups exhibit quite a lot of hourly variation which is not apparent from Table 1. Further, the level of hourly variation also changes over time. To show these patterns I plot the average hourly markups of marginal generators from 2011 to 2014 in Figure 3b. From Figure 3b, we see that on an average markups were about $50/MWh during the peak hour of 16:00 in 2011 and 2013. However, we can see a dramatic drop in markups in 2014 across all peak hours (13:00 - 17:00), perhaps most significant at 16:00 . Average markups are similar over the years for off-peaks hours (0:00 - 12:00 and 18:00 - 23:00), suggesting that CREZ expansion had the strongest effect for peak demand hours.

Table 1: Descriptive statistics of generator markups and nameplate capacity by fuel type

A. Realized Markups ($/MWh)						
Fuel	Mean	Median	SD	Min	Max	Freq.(%)
Coal	5.73	7.26	30.77	-98.78	4597.90	31.87
Natural Gas	17.16	15.17	60.42	-87.32	4899.21	68.13
B. Nameplate Capacity (MW)						
Fuel	Mean	Median	SD	Min	Max	Freq.(%)
Coal	603.18	615	199.73	174.6	1008	31.87
Natural Gas	191.86	178.2	89.78	25	765	68.13

Note: This table presents descriptive statistics for the observations for the markup analysis. Each observation corresponds to the marginal generator(s) at a particular hour. Freq.(%) in each panel refers to the frequency (in percentage) a marginal generator is of specific fuel type in the sample.

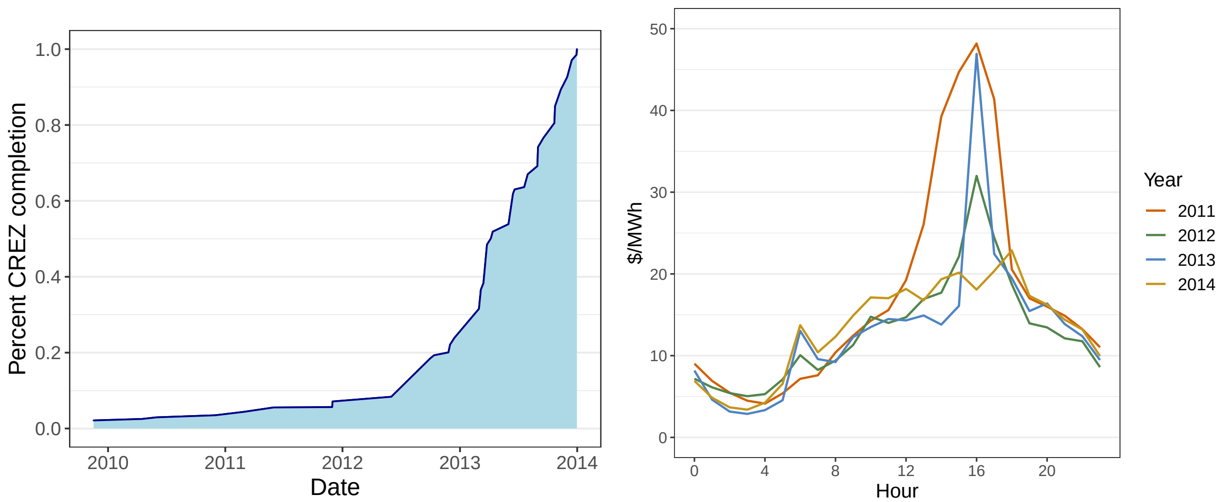
3.3 CREZ Transmission Expansion

Another key variable in the analysis of generator markups 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 for ease of interpretation. As shown in Figure 3a even though the CREZ project was started in 2010, over 80 percent of the project was completed in 2013.

4 Short-run: Impact of CREZ Expansion on Markups

4.1 Real-time electricity market

For this analysis, I focus on the real-time electricity market which sets the expectation for prices in the day-ahead and forward markets (Potomac Economics 2019). The main purpose of a real-time market is to match supply with demand while operating the transmission system within established limits. Real-time operations involve participation from various market participants like generators, retailers, transmission service



(a) Daily progress of CREZ expansion.

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

Figure 3: Daily CREZ progress and generator markups

providers, and distributors. ERCOT serves as the regulatory body that manages the efficient operation of the real-time 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. As noted in (Woerman 2019), 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.

4.2 Transmission Constraints and Market Power

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

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

Presence of transmission constraints prevents ERCOT from dispatching the cost effective generating units to meet the demand at a particular location. To see this, consider a simple example. Say there are two regions A and B. Region A consists of low cost generators that can provide up to 100 MW of electricity and region B consists of high cost generators that can also provide 100 MW of electricity. However, Region A and B are connected by a transmission line that can carry only 50 MW of electricity. Suppose at some time t there is a demand for 80 MW of electricity in region B by households.

ERCOT as the planner, would like to dispatch all of the 80 MW from low cost generators in Region A. However, due to the transmission limit it can only dispatch 50 MW. At this point, the transmission constraint between A and B is said to be binding or there is transmission congestion between A and B. To meet the remaining demand, ERCOT has to dispatch 30 MW of electricity from high cost generators located in region B. Thus, presence of transmission constraint leads to inefficiency in the market clearing process.

A natural question is, how does presence of transmission constraints translate to generators firms exercising market power? Recall that generators submit monotonically increasing offer curves that determines the price of electricity they would charge for a certain quantity. A firm could anticipate demand and transmission constraints and hence submit a bid that is composed of the marginal cost of supplying electricity and a markup term.[5] As I show in the theoretical model in Section 4, this markup term is dependent on the shape of the residual demand curve that the generator faces in the market which in turn is a function of transmission lines.

4.3 A Model of Optimal Fossil Fuel Markups

The theoretical model in this section aims to understand the effect of transmission expansion on 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. Thematically my model is based on Ryan (2021), however, I differ from it in terms of my treatment of transmission constraints and transmission expansion. In what follows, I present the optimal markup rule for a fossil fuel generator and provide intuition on how it is affected by the expansion of electricity transmission lines which enhance electricity generation from wind.

4.3.1 Model Setup

Consider two geographically distinct regions $\mathcal{W}$ and $\mathcal{S}$. Region $\mathcal{W}$ is a wind rich region comprising of wind farms to produce electricity. Region $\mathcal{S}$, on the other hand is comprised of several fossil fuel fired electricity generators that serve a large demand center.

5. In ERCOT, generators do 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.

Presence of electricity transmission capacity (K) enables trade of electricity generated from wind in region $\mathcal{W}$ to demand centers in region $\mathcal{S}$.

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

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

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

The last term in Equation 2 is the payoff from the forward position that is resolved in the real time market. Market demand in $\mathcal{S}$ is denoted by D^S and is assumed to be inelastic. Generator i faces a downward sloping residual demand curve $D_i^r(p, q_w; \mathrm{K})$ which is comprised of three elements: Market demand D^S; electricity generated from wind imported from $\mathcal{W}$, q_w; and the total electricity generated from competitor fossil fuel generators, $Q_f(q_w, p) = \sum_{j \neq i, j \in \mathcal{S}} Q_j(q_w, p)$. I express Q_f as a function of q_w because the dispatch of a fossil fuel generator depends on the amount of electricity generated by wind.[6] Recall that wind-based electricity generation incurs zero marginal cost and is always scheduled to dispatch first. $Q_f(q_w, p)$ is strictly increasing in p and strictly decreasing in q_w.[7] Mathematically, D_i^r can be written as,

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

7. The interpretation of these assumptions is as follows:

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

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

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

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

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

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

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

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

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

The denominator which is the slope of residual demand curve determines the ability of the generator to set markups. An flatter residual demand curve implies that the generator has a lower potential to set markups and thus submits more elastic offer curves. Similarly, a steeper residual demand curve implies that the generator has greater potential to set markups and therefore submits less elastic offer curves.

4.3.2 Predictions from the model

As shown in Equation (5), the extent of marginal generator i's realized markup ($p - c_i$) depends on the shape of the residual demand curve. In order to characterize the effect of change in transmission lines on markups, I perform comparative statics exercise by partially differentiating Equation (5) with respect to K,

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

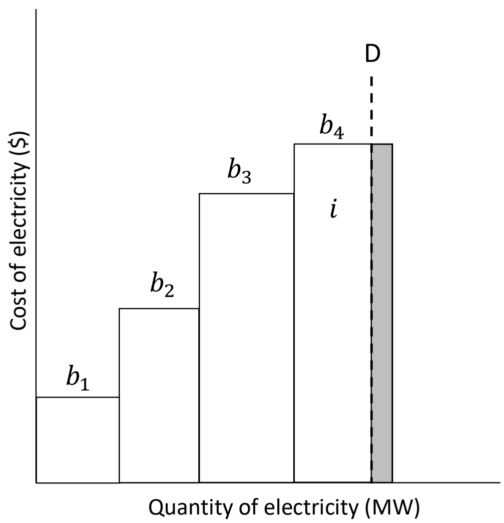
I express Equation (6) as a percentage change in markups by multiplying both sides by the inverse of Equation (5). Algebraic simplification allows me to split the resulting expression into two terms that measure the effect of transmission line expansion on markups. I call these terms Δ*Production* and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

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

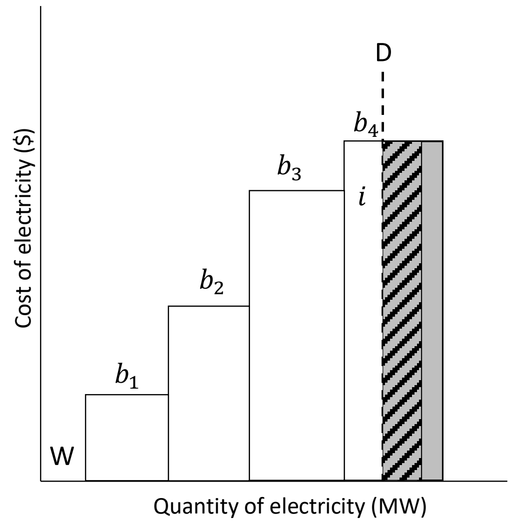
ΔProduction. This term measures the percentage change in net-production by generator i in the real-time market due to change in transmission capacity K.

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

The first term in Equation (8) measures the extent to which production decision of generator i is affected by wind energy. Consider a hypothetical electricity dispatch curve shown in Figure 4a. The supply side assumes four fossil fuel generators indexed by their offer/bid price $b_j (j = 4)$ of electricity they submit. The dispatch curve is a step function comprised of generators arranged in an increasing order of the offer price. The dotted vertical line (D) is the demand of electricity and is assumed to be fixed. Generators are dispatched in the increasing order of the offer price until the demand is met. The generator(s) with the highest offer price that is dispatched is the marginal generator and it determines the wholesale price of electricity.[8] In the scenario below, generator i submits the highest offer price b_4 and is thus the marginal generator.



(a) Hypothetical dispatch curve



(b) Change in the dispatch curve due to wind generation and inward shift in generator i's residual demand.

Figure 4: Hypothetical electricity dispatch curves and the effect of wind generation

Consider the scenario in Figure 4b wherein wind (W) displaces electricity generated from i shown as the grey hashed area. Mathematically, 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 displaying electricity generated by the marginal generator i. This is reflected as an inward shift of i's residual demand curve which in turn reduces i's potential to set higher markups.

The second term in Equation (8) measures the change in import of electricity into $\mathcal{S}$ from wind farms in $\mathcal{W}$ due to a marginal increase in transmission capacity. Transmission expansion would enable higher imports of electricity generated from wind into $\mathcal{S}$, therefore $\partial q_w / \partial \mathrm{K} \geq 0$.

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

Δ**Elasticity.** This term measures the marginal effect of transmission capacity on the slope of marginal generator i's residual demand curve. To understand the direction of this term, I take the derivative of the slope of the generator i's residual demand curve with respect to K. The slope of i's residual demand curve depends only on the production decisions of its competitors as D^S and $q_w(\mathrm{K})$ are invariant to changes in p. Therefore,

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

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

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

The first term on the right hand side in Equation (10) measures the change in the slope of i's competitor fossil fuel generators' supply curve due to an increase in wind energy. The sign of this term is ambiguous because integration of electricity from wind would lead to a rightward shift of the overall dispatch curve leading to generators operating at lower points of their offer curves. However, in the longer-run we could observe high marginal cost or inefficient generators operating less often or even exit the market. This could lead to drop in competition amongst generators incentivizing them to submit steeper offer curves. Therefore, this term could be weakly positive or negative. This translates to Equation (10) being weakly negative or positive meaning a more inelastic or elastic residual demand curve respectively. Multiplying the inverse of the slope of i's residual demand curve in Equation (10) gives us the expression for ΔElasticity in Equation (7).

Substituting the expressions for ΔProduction from Equation (8) and ΔElasticity from Equation (10) in Equation (7):

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

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

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

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

4.4 Empirical Strategy

The objective of this analysis is to examine the impact of transmission expansion effect on fossil fuel generator markups. Equation 12 describes the relationship between transmission lines and markups set by marginal generators. I construct an estimator of the marginal effect of transmission lines on generator markups that is motivated from the theoretical model:

$$\frac{\partial(p-c_i)}{\partial \mathrm{K}} = \underbrace{\frac{\partial(p-c_i)}{\partial q_w}}_{\alpha_h} \times \underbrace{\frac{\partial q_w}{\partial \mathrm{K}}}_{\beta_h} \tag{13}$$

$(p-c_i)$ is the markup of generator i, K measures transmission expansion, and q_w is the electricity generated from wind. I run the following regressions to estimate the terms on the right hand of Equation 13:

$$y_{it} = \alpha_h \cdot w_t + f(\mathrm{D}_t|\lambda) + \kappa_i + \delta_{hmy} + \epsilon_{it} \tag{14}$$

$$w_t = \beta_h \cdot crez_d + \gamma \cdot \mathrm{H}_t + \eta_{hm} + \omega_t \tag{15}$$

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 LMP and c is the marginal cost of generator i at period t. w_t is wind generation (GWh) in hour t, and $crez_d$ is the percentage completion of CREZ transmission project at day d of the sample. The parameters of interest are α_h

which measures the impact of wind generation on markups, and β_h which measures the impact of transmission expansion on wind generation.

I use a wide variety of controls to account for potential confounding factors in Equation 14 and Equation 15. I use a quadratic function of system wide electricity demand D_t in Equation 14 to account for variation in markups driven by spikes in electricity demand.[9] In Equation 15, I use the High System Limit (H_t) of electricity generation from wind at hour t to control for the maximum energy production possible from wind at period t.[10] H_t is a useful control because it not only incorporates the generating capacity of the wind based resourced but it also takes into account the real time meteorological conditions that could affect the amount of power generated through wind farms.

As shown in Figure 5 the actual electricity generated from wind w_t closely tracks H_t for each hour from 2011 to 2014. The difference between the two curves for a specific year arises due to the lack of transmission infrastructure needed to transport the power to demand centers. However, with the CREZ expansion in 2013 we see the gap between the high system limit of wind generation and actual wind generation decreasing with the lowest difference observed across all hours of 2014.

To control for unobserved determinants of markups that could be correlated wind generation I use a battery of fixed effects. In Equation 14, I use generator fixed effects (κ_i) to control for any generator specific heterogeneity in markups. I use hour by month by year fixed effects (δ_{hmy}) to control for seasonality exhibited by electricity market in Texas. This seasonality arises due to varying demand for electricity at different hours of the day over the months in a year. For example, demand for electricity in Texas tends to be higher during the afternoon than at early morning or night. Similarly, demand is typically higher during the summer than the winter months. Similarly, I use hour-by-month fixed effects (η_{hm}) in Equation 15 to control for seasonality in wind generation. ϵ_{it} and ω_t are the random error terms in Equation 14 and Equation 15 respectively.

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

10. ERCOT defines HSL for a generation resource as the limit established by the QSE, continuously updated in Real-Time, that describes the maximum sustained energy production capability of the resource.

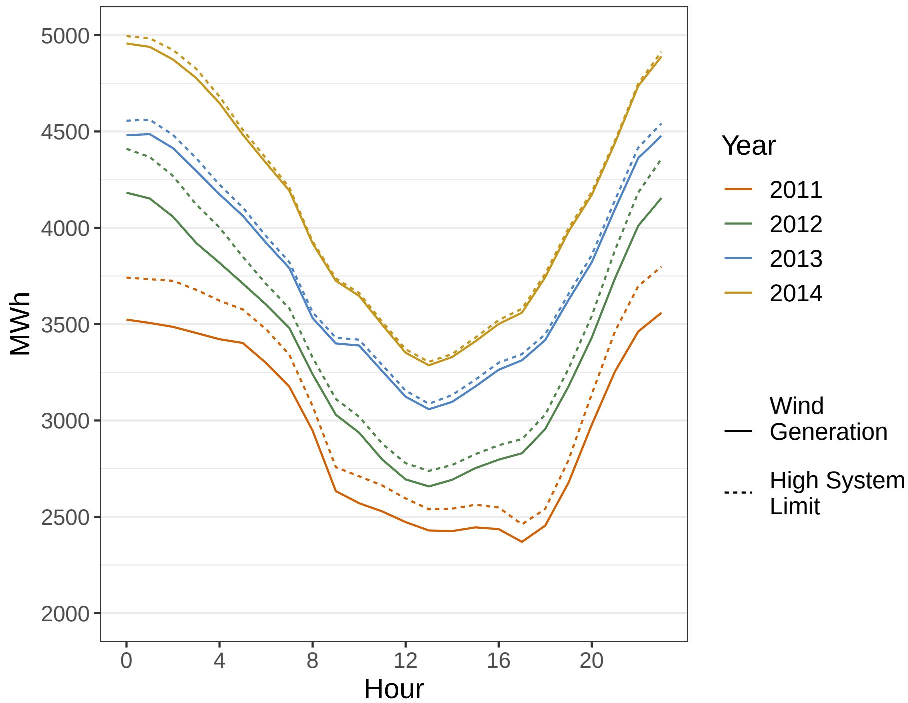


Figure 5: Hourly averages of actual Wind Generation (w_t) and High System Limit of wind generation (H_t) from 2011 - 2014.

The identifying variation in Equation 14 comes from the within generator variation in markups caused by changes in wind generation across hours h within a month m in a given year y. For example, α_{16} is the identified from deviations in markups from generator specific averages across all 16:00 hours within a month, in a given year. A similar argument holds for the identifying variation in Equation 15. Under the identifying assumption that control variables and fixed effects account for the confounding factors, α_h captures the unbiased effect of wind generation on generator markups and β_h is the unbiased effect of CREZ expansion on wind generation. Therefore, the hourly estimator of Equation 13 can be written as $\theta_h = \alpha_h \times \beta_h$. Standard errors in Equation 14 are clustered at the generator level. I use Newey West auto-correlation corrected standard errors with a seven day lag structure for estimates in Equation 15.

4.5 Results

I first present the results of the effect of CREZ expansion $(crez_d = 1)$ on wind generation in Figure 6 followed by the effect of wind generation on generator markups in Figure 7. The plot of $\hat{\beta}_h$ shows the differential impact of CREZ completion $(crez_d = 1)$ on integrating wind generation at hour h. The effect is strongest during the off-peak hours.[11] with the coefficient estimate being the highest at 23:00 and 0:00, about 0.22 GWh. The impact of CREZ is least during the peak demand hours with an average of approximately 0.10 GWh between 15:00 - 18:00 hours.

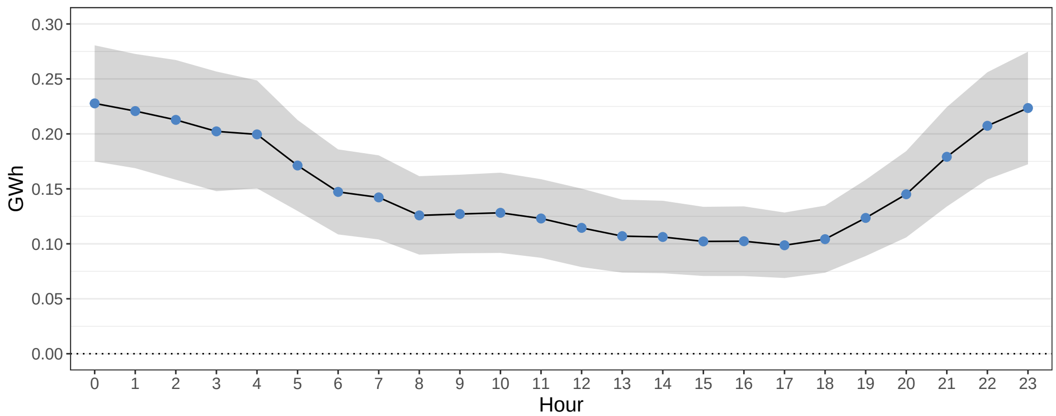


Figure 6: Coefficient estimates $(\hat{\beta}_h)$ for the effect of CREZ expansion on hourly wind generation (GWh) and the 95 percent confidence intervals.

The hourly pattern of the coefficient estimates $(\hat{\beta}_h)$ for the impact of CREZ expansion on wind generation closely follows the hourly wind flow pattern in Texas where the wind flow is strongest during the off-peak hours in the evening compared to the peak hours during the day. This reflects that availability of excess transmission capacity is instrumental in integrating higher levels of wind generation in the Texas electricity market.

Figure 7 shows the coefficient estimates $(\hat{\alpha}_h)$ of the magnitude of the decrease in fossil fuel markups ($/MWh) due to 1 GWh addition of wind energy to the grid. We see that the drop in markups is strongest in magnitude at hour 16:00, about $9/MWh for 1 GWh of electricity generated from wind. The coefficient estimates are smallest for the off-peak hours. Due to low demand and high electricity generation from wind during the off-peak hours, fossil fuel generators typically operate on a smaller net-demand curve as compared to the net demand during the peak hours.[12] This lowers their incentives

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

12. Recall that the net-demand or residual demand curve for fossil fuel generator is defined as $D_i^r(p, q_w; \mathrm{K}) = D^S - q_w - Q_f(q_w, p)$ where D^S is the system demand, q_w is the power generated from wind, and $Q_f(q_w, p)$ is the total electricity generated by other fossil fuel generator.

to set high markups during the off-peak than during the peak times. The impact of an additional GWh of wind is therefore higher on the net demand curve during the peak hours than during the off-peak. This is reflected as the larger magnitude of coefficient estimates for the peak hours than for the off-peak hours.

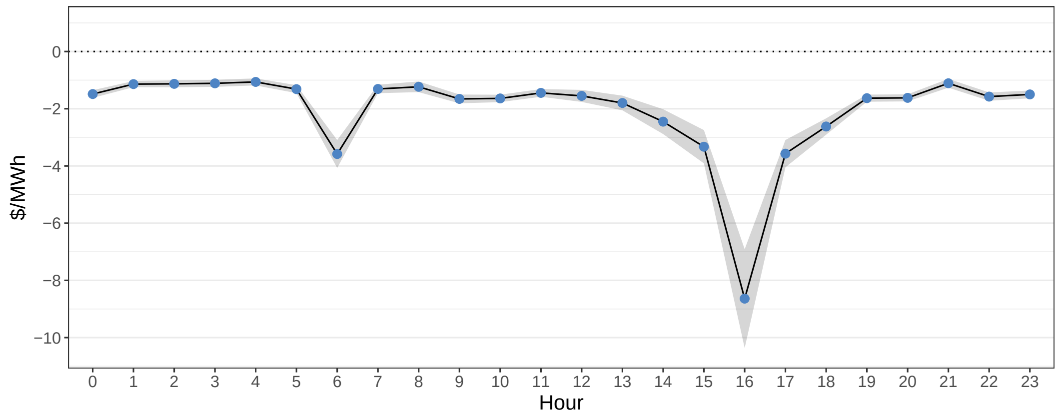


Figure 7: Coefficient estimates ($\hat{\alpha}_h$) for the effect of addition of 1 GWh of wind energy on generator markups ($/MWh) and the 95 percent confidence intervals.

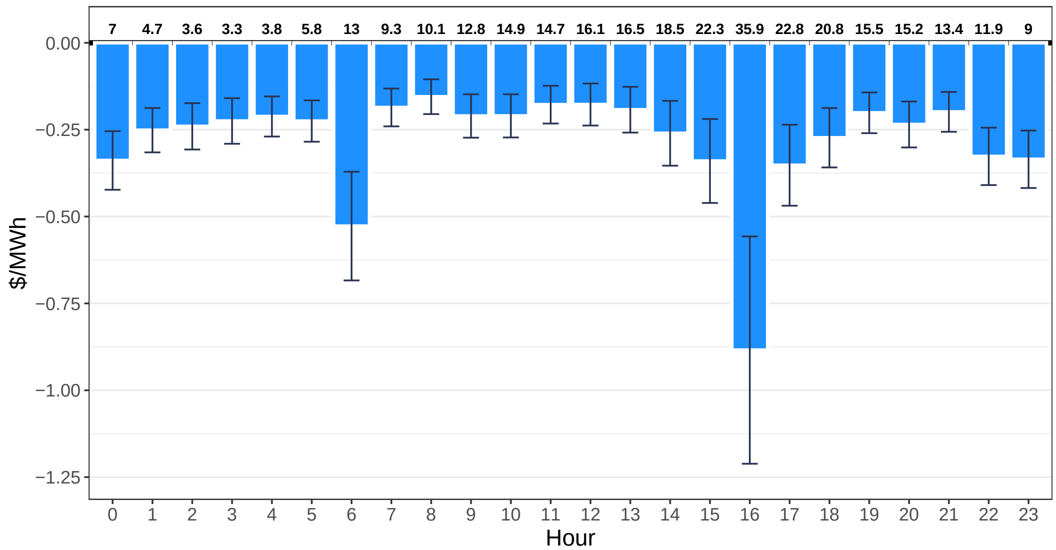
I combine the estimates in Figure 6 and Figure 7 to obtain the impact of CREZ on markups ($\hat{\theta}_h$) for each hour in Figure 8a. Results show that the drop in markups due to CREZ is strongest at 16:00, about $0.88/MWh. This is followed by 6:00 (about $0.53/MWh) which stands out from other hours in the earlier part of the day possibly because it marks the end of the off-peak hours of weekday in ERCOT. Compared to total wind power integrated at hour 0:00, CREZ only led to the integration of about half that amount at 16:00 ($\sim$ 0.10 GWh), but the subsequent drop in markups at 16:00 is about 2.5 times that at 0:00.

To get a better sense of the magnitudes in Figure 8a, I present the semi-elasticity of markups in response to CREZ expansion (i.e. $crez_i = 1$) in Figure 8b.[13] We see a clear distinction between the semi-elasticity of markups between off-peak v.s. peak hours. The magnitude of the semi-elasticity is highest for hours before 7:00 with the maximum decrease of 6.8 percent at hour 3:00. However, the percentage drop in markups for all peak hours (7:00 to 22:00) is less than 3 percent mainly because of the prevalence of

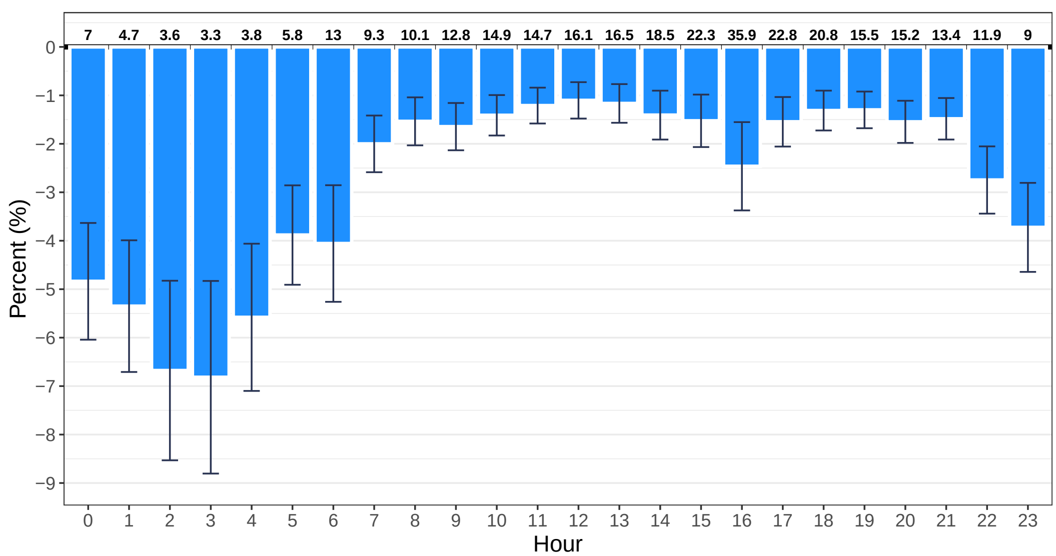
13. I calculate semi-elasticity values by dividing the estimates in Figure 8a with the average markup at each hour.

high average markups during these hours. This highlights that the greater transmission capacity which integrates the existing and growing wind generating capacity has a significant effect in lowering markups especially during the peak hours.

An important caveat is that the patterns observed in these results are within generator changes in the markups as I use generator fixed effects in Equation 14. A simplified version of this effect is shown in Figure 4 in the theoretical model. In Figure 4b we see an inward shift in generator i's residual demand curve due to electricity generated from wind. It does not include changes in markups due to generators operating at different points of their supply curve in response to competition from other generators. As shown in the theoretical model, firm exit could incentivize generators to submit steeper offer curves potentially leading to higher markups. I plan to analyze this channel in future work.



(a) Effect of CREZ completion ($crez_d = 1$) on markups ($/MWh) and the associated 95 percent confidence intervals (average markups for the sample are mentioned above the x axis).



(b) Impact of CREZ completion ($crez_d = 1$) on the semi-elasticity of markups and the associated 95 percent confidence intervals (average markups for the sample are mentioned above the x axis).

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

5 Short-run: Impact of CREZ Expansion on Marginal Emissions

How did wind generation due to CREZ expansion affect emissions from the marginal generators in the short-run? Addition of wind generation to the grid would shift the dispatch curve of electricity rightwards. Further, the intermittent nature of wind generation is likely to affect which generator operates at the margin and therefore the marginal emissions. In this section I examine how additional wind generation due to CREZ expansion affected the emissions from the fossil fuel generator(s) operating at the margin.

Looking especially the emissions from marginal generators is useful because of the variation in the types of marginal generators over the course of a day. For example, coal fired generators typically operate at the margin during the night whereas natural gas generators are usually the marginal generators during the day since they are quicker to ramp up or down to meet any sudden changes in demand. This is evident from Figure 9 wherein marginal CO_2 emissions (tons) per MWh of electricity produced are highest during the night and in the early morning.

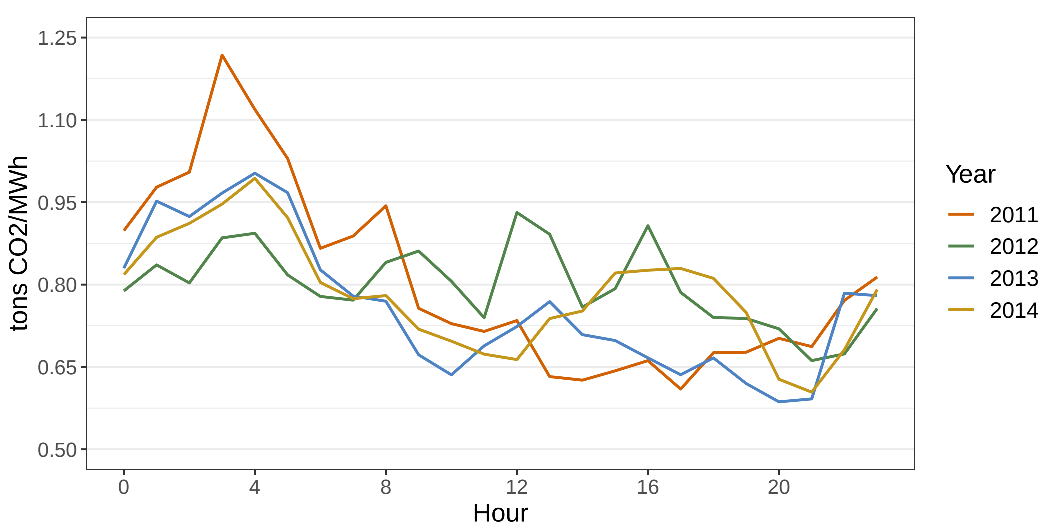


Figure 9: Hourly average of CO_2 emissions (tons) per MWh of electricity generated from marginal fossil fuel generators from 2011 - 2014.

The addition of wind based electricity to the grid in the night could therefore displace high polluting coal generators from the margin whereas addition of wind energy in the day would potentially displace some natural gas generators. This could lead to higher proportional drop in emissions during the night as compared to the day. To isolate the impact of additional wind energy in the grid on marginal carbon emissions, I estimate the following specification:

$$E_{zt} = \rho_{zh} \cdot w_t + f(D_{zt}|\lambda) + \alpha_{zy} + \delta_{hmy} + \epsilon_{zt} \tag{16}$$

where, E_{zt} is the zonal level marginal emissions (i.e. total zonal emissions from generators at the margin) and w_t is the wind generation at hour t of the sample. The parameter of interest is ρ_{zh} which measures the average impact of an additional GWh of system wide wind generation on the marginal emissions in zone z at hour h.

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

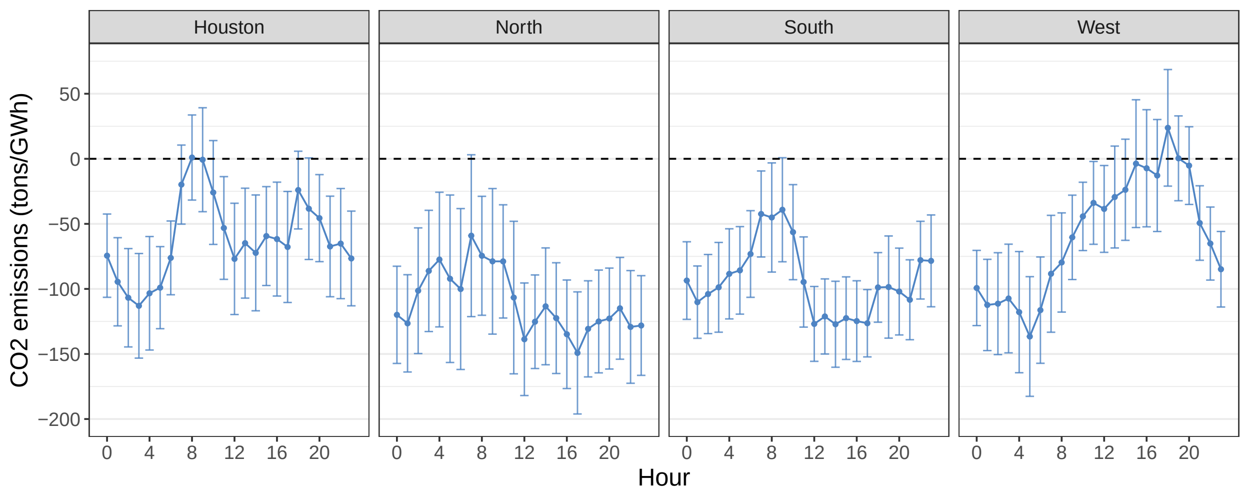
5.1 Results

5.1.1 Impact on marginal carbon emissions

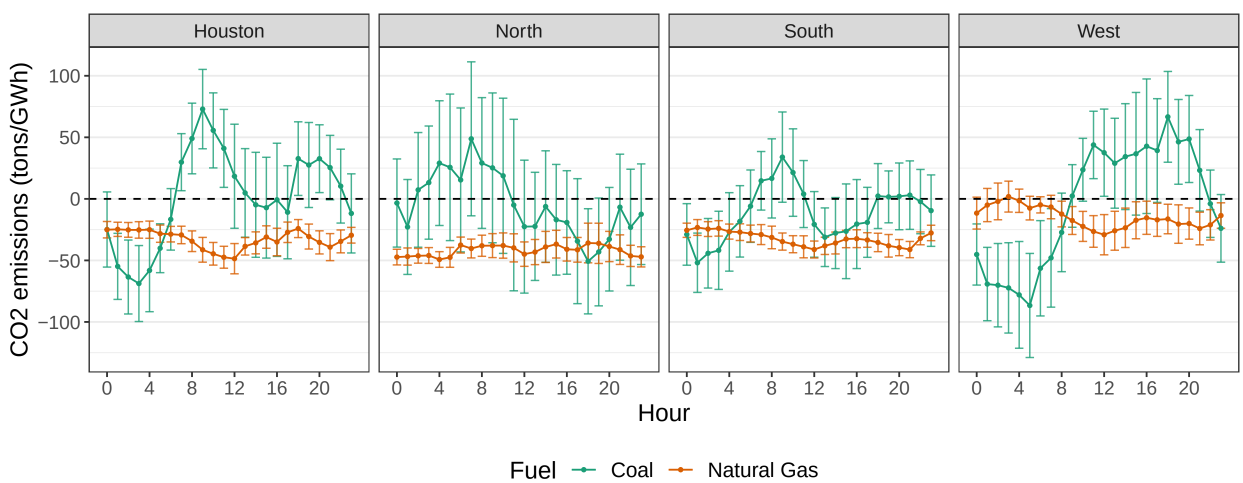
Figure 10a shows the average changes in the marginal CO_2 emissions (tons) during each hour of the day across each Zone in response to an additional GWh increase in wind energy. There is a clear decrease in marginal carbon emissions across all the zones throughout the day. However, the hourly pattern of emissions reduction is quite different across the zones. In the North and the South zones the decline is highest in magnitude between the hours 12:00 - 20:00. This pattern is flipped for the West zone wherein the decline is strongest in the night and statistically insignificant for the hours between 12:00

- 20:00. The emission pattern for Houston shares some similarities with the West zone in that the drop in marginal carbon emissions is highest in the night.

The patterns observed in Figure 10a could be the result of the heterogeneity in the types of marginal generators across the zones at different hours of the day. To explore this, I estimate Equation 16 separately for the sample of marginal emissions from coal and natural gas. The results of the coefficient estimates for impact of wind generation on marginal emissions is shown in Figure 10.



(a) Impact of wind generation on marginal CO_2 emissions



(b) Impact of wind generation on CO_2 emissions disaggregated by Fuel type

Figure 10: Impact of an additional GWh of wind generation on tons of marginal CO_2 emissions by Zone and Fuel type over 2011-2014.

Two key insights emerge from Figure 10b. First, the hourly distribution of estimates for coal is very similar to the hourly distribution of the overall emissions in Figure 10a, suggesting that the pattern in Figure 10a is mainly driven by emissions from coal generators. Second, the drop in emissions from marginal natural gas generators is mostly stable and negative over the hours across all the four zones. This seems to shift the emissions estimates from coal downwards when aggregated by fuel type giving rise to the pattern in Figure 10.

Interestingly, we see a significant decline in coal emissions in Houston, South, and West during the night and early hours of the day. This pattern is reversed for North zone but the estimates are statistically insignificant. This provides empirical evidence that addition of wind energy during the night has a significant impact on reducing emissions from coal generators at the margin. However, we can notice a statistically significant increase in carbon emissions from coal during the day in Houston and West zone.

Zooming in on the marginal generators within West and Houston, we observe that these estimates are driven by the only coal power plants in those zones. In Houston, the coal emissions are due to W.A. Parish Coal Plant (four generators with total capacity of 2.7 GW) whereas in West the emissions are due to Oklaunion Power Plant (single generator with 720 MW capacity).[14] The increase in carbon emissions in the day suggests that during the periods with low wind generation post 8:00, availability of higher transmission capacity tends to promote power from coal generators located near the population centers to meet the demand. This is in line with Fell, Kaffine, and Novan (2021) who find that environmental value of wind in Texas is negligible in periods with low wind generation.

5.1.2 Impact on marginal local pollution (SO_2 and NOx)

I use total damages (in 2020 $) due to SO_2 and NOx emissions from marginal generators as the dependent variable in Equation 16 to estimate the impact of hourly wind generation on damages from local pollutants. Since damages from local pollutants vary across space, I use estimates of county-specific marginal damages due to emissions from SO_2 and NOx from Holland et al. (2016) to calculate the value of damages due to emissions

14. Figure B1 in Appendix B shows the location of various coal and natural gas fired marginal generators in the sample from 2011 to 2014 along with ERCOT Zones and CREZ counties.

from each generator. I aggregate these damages at the Zonal level for the estimation. Figure 11 shows the coefficient estimates associated with each hour of wind generation (GWh) and Zone.

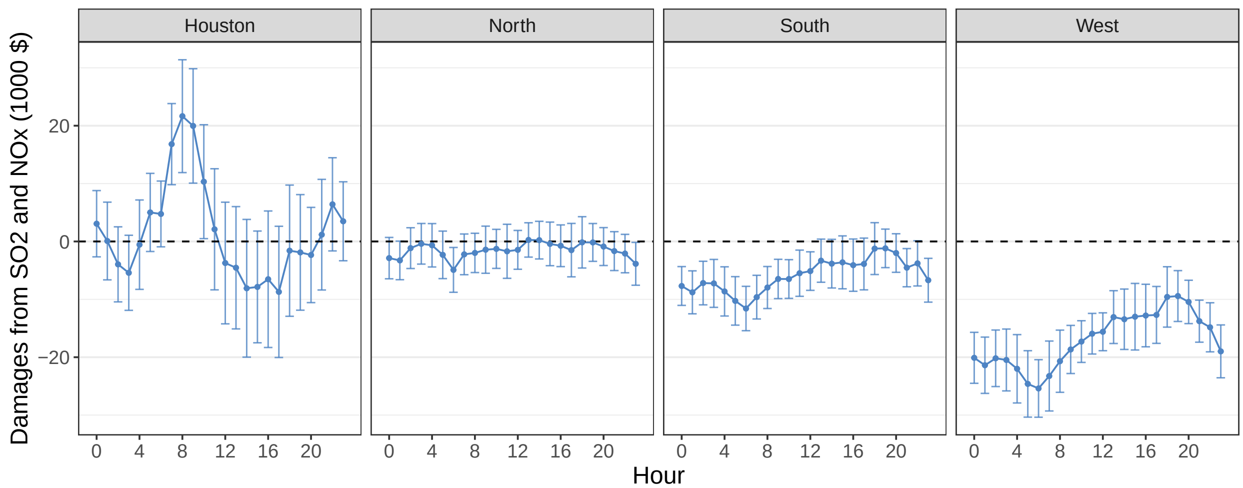


Figure 11: Impact of an additional GWh of wind generation on damages (2020 $) due to local pollutants (SO_2 and NOx) by Zone.

Figure 11 shows evidence of significant heterogeneity in damages avoided from local pollutants across zones due to additional wind generation. For South and West zones, additional wind leads to decline in damages from SO_2 and NOx across all hours, whereas the effect is statistically insignificant for the North zone. In case of Houston, we see a significant rise in local emissions during the early hours of the day. This is similar to the rise in carbon emissions observed from the coal generator W.A. Parish in Houston. It is likely to be driven by ramping up of the coal generator during the day when the wind generation becomes low. These ramping effects are shown to undercut the emissions reductions from wind especially when operating at low levers of maximum generation (Lew et al. 2012).

5.1.3 Damages avoided from marginal emissions

To get a better sense of the magnitudes of hourly marginal emissions, I translate them into average reduction in damages ($). I combine the hourly ρ_{zh} estimates from Equation 16 with the estimates on average wind generation added to the grid due to CREZ expansion (β_h) from Figure 6. I estimate damages as:

$$D_z(\$) = \sum_{h=0}^{24} \tau \times \beta_h \times \rho_{zh} \tag{17}$$

where, D_z is the zonal daily average of damage (in 2020 \$) due to marginal carbon emissions in zone z. I assume social cost of carbon, τ as \$44 per ton of CO_2 emissions (US Interagency Working Group on Social Cost of Carbon 2014), β_h is the hourly average wind generation added due to CREZ in the short-run estimated in Equation 15, and ρ_{zh} is hourly average of marginal carbon emissions avoided due to an additional GWh of wind energy. For local pollution, I simply multiply the coefficient estimates in Figure 11 with β_h and aggregate over the hours to get the value of average daily damage avoided.

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

	Damages Avoided (2020 \$)			
Zone	CO2	SO2 + NOx	Total	Percent (%)
Houston	10,733	-6,939	3,794	3
North	17,448	6,218	23,666	17
South	14,765	22,757	37,522	27
West	10,957	64,515	75,472	54
Total	53,903	86,551	140,454	100

Notes: This table reports the daily average of damages from marginal CO_2 and local pollutants avoided due to additional wind integrated from CREZ expansion for each Zone.

Table 2 reports the average damages avoided (2020 \$) per day from marginal emissions due to CREZ expansion for each zone.[15] We notice that CREZ expansion led to a decline in damages from marginal carbon emissions across all the zones in the short-run. The total damages avoided is about \$54,000 per day with the North and the South zone contributing about three fifths of the total. The distribution of damage reduction across

15. The coefficient estimates of hourly averages of damages avoided for each zone due to CREZ are presented in Figure B2 in Appendix B. The pattern for carbon emissions and local pollution is similar to Figure 10a and Figure 11 respectively.

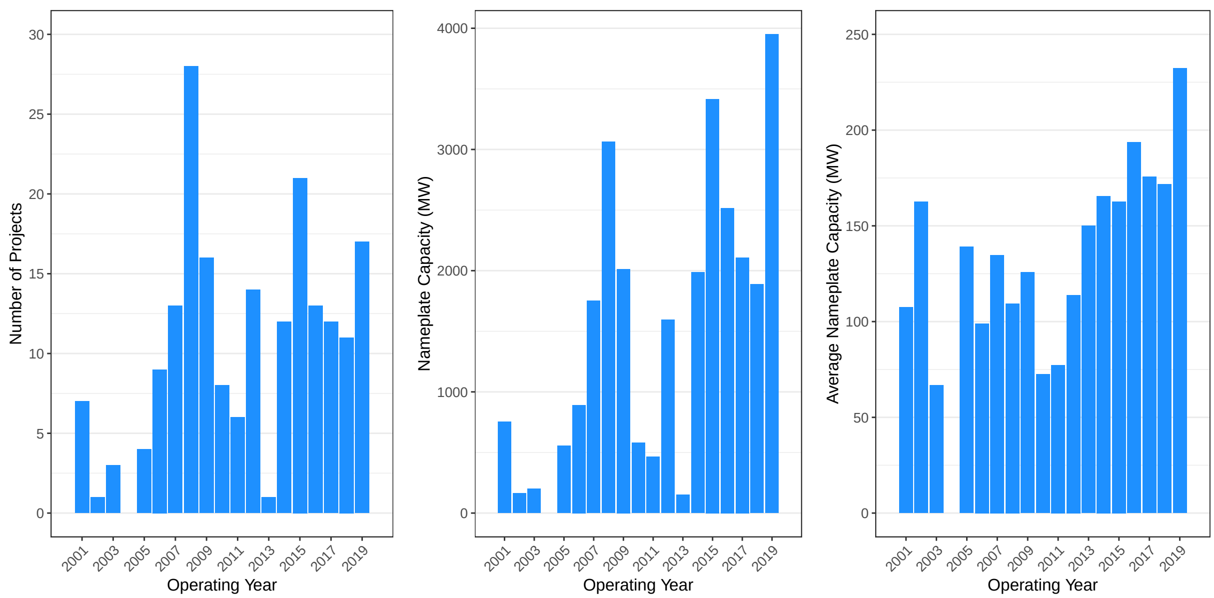
the zones highlights the wide spatial scope of reduction in carbon damages accrued from the integration of wind due to availability of transmission capacity.

The story is a bit different for local pollutants. While the total daily damages avoided is about $87,000, it is mostly concentrated in the West zone followed by South and North. The negative sign for Houston indicates an increase in damages instead of a decline. This is due to the increase in emissions from W.A. Parish Coal plant in Houston during the day. The total daily damages avoided from carbon, SO_2 and NOx emissions is $140,454 which translates to about $51 million annually.

The related literature in economics has generally focused on how reshuffling of generation due to lower congestion and additional wind affects emissions in the power sector (LaRiviere and Lu 2020; Fell, Kaffine, and Novan 2021). This subsumes changes in emissions due to highly polluting units operating at the margin at certain periods of the day. The analysis in this section attempts to identify this aspect by focusing on the emissions from marginal generators in response to additional wind due to transmission expansion. As shown above, I find that while there is significant spatial and temporal heterogeneity in emissions from marginal generators. While there is an overall decline in emissions, we do observe increase in emissions especially during the periods with low wind generation leading to ramping up of some of the coal generators.

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

In this section I examine whether the counties announced to site CREZ related substations and wind collection points saw higher levels of wind investment. I compile a dataset of all the wind projects in Texas post 2001 using the Wind projects supplement of EIA Form 860. This dataset provides detailed information about a wind project including its location and year of operation. Figure 12 shows the total number of wind projects, total capacity (MW) of projects, and average capacity (MW) per project over 2001 - 2019 in Texas. We see a decline in total projects and total capacity in 2013 possibly due to the expiration of Production Tax Credit (PTC) in late 2012. However, both total capacity and number of projects increased sharply post 2013 with an increasing trend in average capacity per project.



(a) Total wind projects (b) Total capacity (MW) (c) Average capacity (MW) per project

Figure 12: Statistics of wind projects in Texas over 2001 - 2019.

The specific technical details of the transmission expansion - the cost breakdown, expected completion dates, and various transmission service providers responsible for the expansion was released in October 2010 in CREZ Progress Report (RS&H 2010). Thus, I refer to October 2010 as the "announcement date" as it provides the first most accurate information of transmission siting in the CREZ transmission expansion.

The 3,600 miles of 345 kV electricity transmission lines were connected between existing and new substations throughout the Panhandle, West, and North zones of Texas.[16] In the data I only see the counties where these substations were located and thus I refer to these counties as 'CREZ counties'.[17]

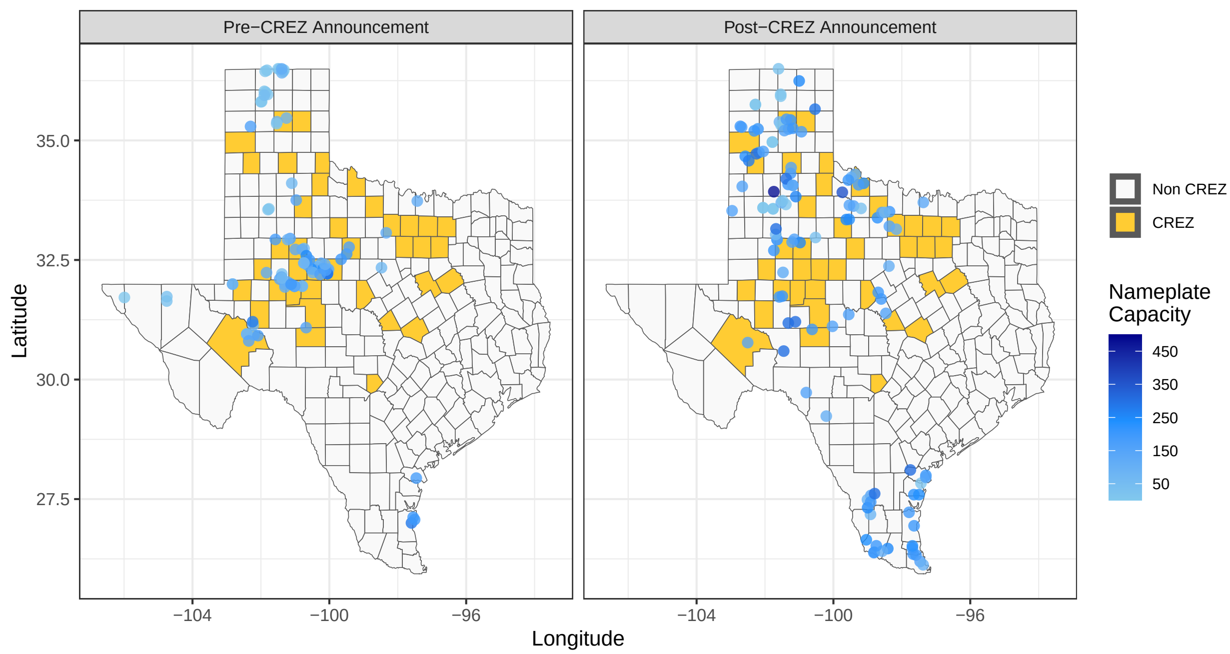


Figure 13: Location and Nameplate Capacity (MW) of wind projects Pre and Post CREZ announcement in Oct 2010

Note: This figure shows wind generators in Texas Pre-CREZ Announcement (2001-2010) and Post-CREZ Announcement (2011-2019) samples.

From Figure 13 we can see a cluster of wind projects located within and near CREZ counties post 2010. However, there is also a cluster of wind farms in coastal Texas. This is because of a superior wind quality in this region which could be profitable for wind

16. Electrical or Transmission substations typically serve as the terminal points for high voltage transmission lines as well as serve as the hub for transmission lines carrying electricity from nearby power generation plants.

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

developers. To parse out whether CREZ counties saw higher levels of wind investment than other counties after accounting for wind quality and other confounding factors, I estimate the following specification:

$$y_{it} = \alpha + \beta \cdot crez_i + \mathbf{X}'\Pi + \epsilon_{it} \tag{18}$$

where, y_{it} is the outcome of interest. I use total wind capacity in county i in year t, average wind capacity of the project (total nameplate capacity/total number of projects in the county), and total number of turbines in county i in year t as the dependent variables for this analysis. The variable $crez_i$ is a binary variable that specifies whether county i is a CREZ county. Since the location details for the CREZ project were initially announced in late 2008 followed by the more accurate information in October 2010, the analysis is restricted to annual county level observations from 2012 through 2019. This also accounts for the fact that wind project planning and development typically takes a few years.

I use a battery of control variables and fixed effects summarized by vector $\mathbf{X}$ in Equation (18). I use power curve, capacity factor, and cubic polynomial of average wind speed to flexibly control for wind resource quality of a county. All of these variables come from NREL's Wind Integration National Dataset (WIND) toolkit (Draxl et al. 2015). I use average yearly wind project cost data from Lawrence Berkeley's Wind Technologies Report, land price data, and median land acreage compiled by the Real Estate Center at Texas A&M University to control for project costs.

To control for demographic factors that could influence CREZ transmission siting and location choice for wind projects I use median household income in 2007 and average population over 2007 to 2010. I use average farm size in a county to control for variation in wind investment due to disseminates from wind turbines in counties with smaller farms (Winikoff and Parker 2019). This data comes from the USDA Census of Agriculture. The rationale behind these variables is that counties with higher household incomes, population, or average farm size could have a higher bargaining power in influencing the regulator's decision to site transmission infrastructure and wind farms.[18]

Cities and counties often enact regulations for wind projects that are sited in their jurisdiction. These regulations are commonly known as setbacks or wind ordinances. They usually specify limits on the size of wind turbines, height of turbines, noise, maximum capacity, etc. Therefore, presence of a wind ordinance could affect investment in wind capacity in a county and could also be correlated with siting of transmission infrastructure. I include a indicator variable specifying whether the county (or a city in the county) has a wind ordinance.[19] I use the publicly available dataset on wind ordinances from WINDExchange for this variable.

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

6.1 Instrumental Variable Strategy

A key concern in Equation (18) is the endogeneity of $crez_i$ due to the selection of specific counties for CREZ expansion. This could lead to biased OLS estimate of β. To alleviate this concern I use an instrumental variables strategy with the designation of counties into five Renewable Energy (RE) Zones in 2007 (z_{i2007}) as the instrument. The IV specification can be written as:

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

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

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

18. 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.

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

6.1.1 Identification

In 2007, the Public Utility Commission of Texas (PUCT or Commission) in its Interim Order on Reconsideration in Docket 33672 designated five zones as Competitive Renewable Energy Zones (PUCT 2007). The five RE Zones - Panhandle A, Panhandle B, Central West, Central, and McCamey are depicted in Figure 14a. These zones were selected based on the existing and potential wind generation capacity, and wind developer interest in these areas. Thus, one of the objectives of the CREZ project was to integrate the wind capacity in these zones to the power grid.

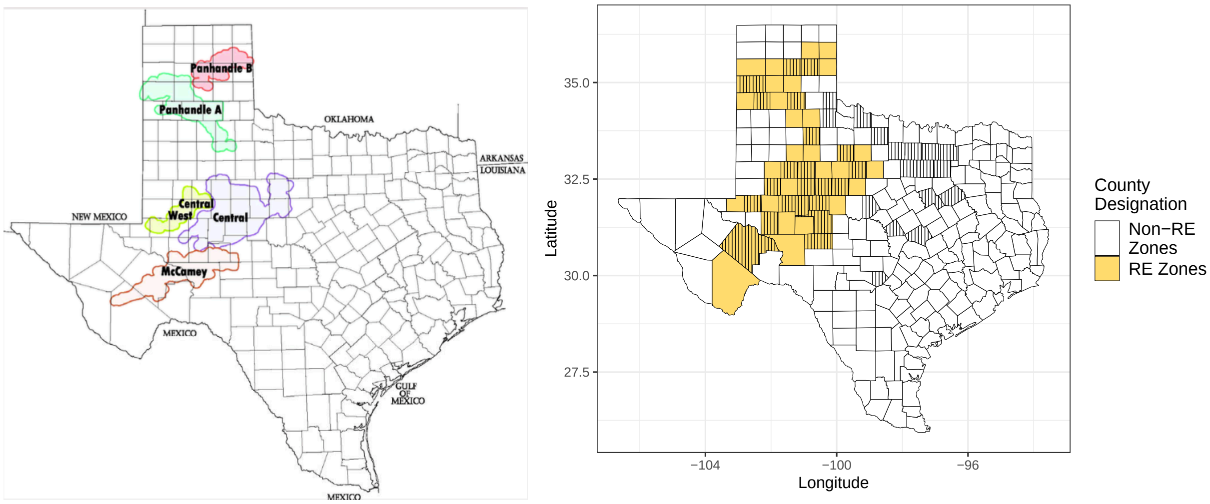
From Figure 14a, I identify the counties that were part of the five RE Zones and denote them as z_{i2007}. These counties are highlighted in yellow in Figure 14b. In April 2008, ERCOT released the CREZ Transmission Optimization Study which analyzed various scenarios for the placement of CREZ transmission lines, substations, and wind collection points within the five RE Zones(ERCOT 2008). Ultimately, scenario 2 was selected by the PUCT as the optimal scenario for CREZ Project. Over the next few years, exact locations of transmission infrastructure was determined based on cost considerations, inputs from county legislatures and wind developers. The hatched counties in Figure 14b are the counties that sited CREZ infrastructure - substations and wind collection points.

As evident from the Figure 14b, there is a significant overlap between the counties that were part of the RE Zones (z_{i2007}) and the CREZ counties $(crez_i)$. Therefore, using z_{i2007} as an instrument for $crez_i$ validates the relevance condition:

$$\mathbb{E}[z_{i2007}, crez_i|\mathbf{X}] \neq 0 \tag{19}$$

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

$$\mathbb{E}[z_{i2007}, \eta_{it}|\mathbf{X}] = 0 \tag{20}$$



(a) RE Zones in 2007

(b) Overlap between counties in RE Zones and county level siting of transmission (hatched)

Figure 14: Designation of counties in Renewable Energy (RE) Zones in 2007 and CREZ counties

Equation 20 refers to the exogeneity of the instrument to the dependent variable y_{it}. In other words, it implies that conditional on $\mathbf{X}$, investment in wind energy in a county from 2012-2019 is only driven by its designation to one of the five RE Zones through its effect on the likelihood of that county being selected as a CREZ county.

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

6.1.2 Results

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

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

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

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

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

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

6.1.3 Robustness Checks and Caveats

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

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

Another concern with the instrumental variable strategy is that the counties within the RE Zones were inherently different than the other counties. Table C2 in Appendix shows that counties within RE Zones were statistically different than the other counties on almost all of the observable characteristics pre-transmission expansion announcement. I attempt to account for these factors by including a rich array of control variables that condition on these inherent differences in the counties. However, the concern of a lack of common support amongst the treated and control counties in the 2SLS regressions still remains. Thus, I further address these concerns by implementing a matching estimator which constructs a control group of identical counties based on observable characteristics.

6.2 Matching Strategy

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

$$\mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|\mathbf{X}, crez_i = 0) \tag{21}$$

where, ϵ_{it} is the unobserved component of dependent variable of interest (y_{it}) - wind capacity (MW), total turbines, and average project size.[20] Under the assumption that the unobserved component (ν_i) of a county that affects the treatment status is time invariant, using county fixed effect would eliminate the selection bias. However, since the treatment variable is assigned at the county level and at the beginning of the sample, I cannot include county fixed effects.

Instead, I assume that ν_i can be approximated using some flexible function of observable county characteristics $\mathbf{Z}$, i.e. $\nu_i = f(\mathbf{Z})$. Therefore, validating CIA involves comparing counties with exactly the same combination of characteristics, such that $\mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 1) = \mathbb{E}(\epsilon_{it}|f(\mathbf{Z}), \mathbf{X}, crez_i = 0)$. This would provide an estimate of the unbiased effect of the treatment. However, 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, King, and Porro (2012) to obtain the set of counties comparable on observable dimensions which includes both continuous and discrete variables. I divide the sample of counties across CREZ (treated) and non-CREZ (control) groups and then match the counties across the two groups based on observable characteristics using CEM. I use a wide variety of pre-treatment observable covariates to account for factors that could have influenced CREZ siting as well as investment in wind energy post 2012. These factors include historical wind capacity,

20. For simplicity, we can think of the linear specification in Equation 18 as the data generating process:

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

wind resource quality, average land price and acreage, and demographic characteristics to account for citizen bargaining power(Billo 2017; Cohn and Jankovska 2020).[21]

Table 4 provides the balance table of these observable characteristics for pre- and post-matched samples. As evident, CEM provides a well balanced group of treated and control counties that look identical on all observable dimensions. From Table 4 we see that counties that do not lie in the common support of observable characteristics used in matching are discarded from the sample. Therefore, the control group comprises of 30 counties and the treated group comprises of 13 counties. Figure 15 shows the map of treated and control counties. Most of the control counties (light yellow) are adjacent to the treated counties (dark yellow).

For the regression analysis on the counties obtained by matching, I use the same set of control variables as in Equation 18. The key assumption is that conditional on $\mathbf{X}$, there are no unobservables that affect the outcome variable and treatment status $(crez_i = 1)$. Therefore, conditional on $\mathbf{X}$, the treatment assignment is as good as random for the given set of counties.

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

Table 4: Balance Table of key observables for Pre- and Post-Matching Sample

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

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

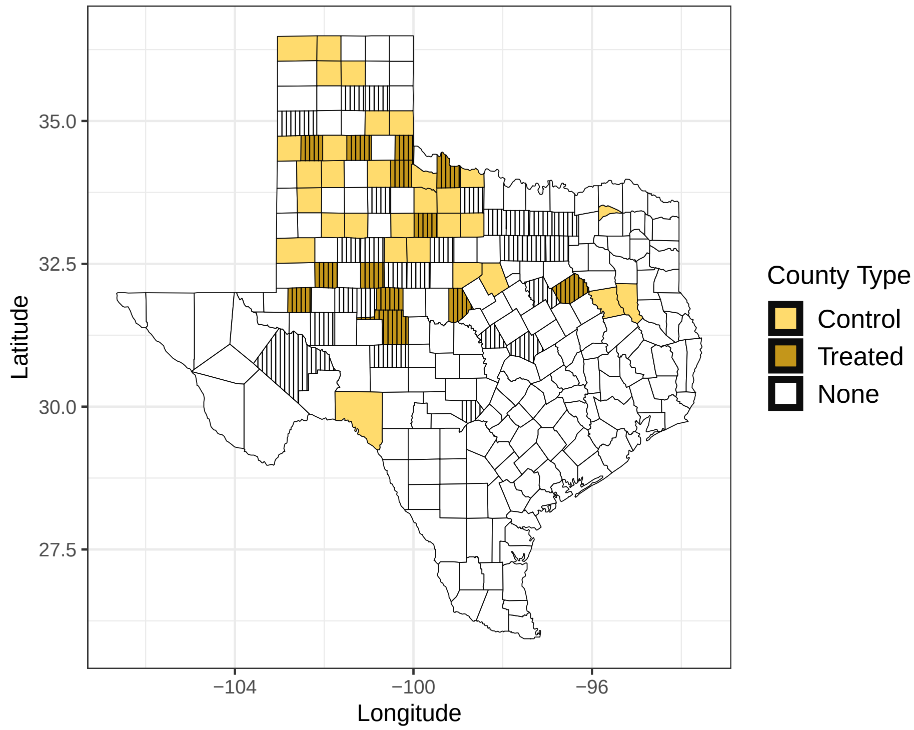


Figure 15: Treated and Control counties obtained using Coarsened Exact Matching (CEM). Hatched counties depict the original CREZ counties.

6.2.1 Results

Table 5 reports the regression results of the baseline specification with total nameplate capacity (MW), total turbines, and average project capacity (MW) as the dependent variables respectively. All the specifications include the same set of control variables used in the IV strategy. I include the interaction of matched group fixed effects with the time trend to allow for unobserved factors that could affect specific groups and vary over time. The sample is a balanced panel of 13 treated and 30 control counties obtained using CEM.

Column (1) in Table 5 shows that the counties that received CREZ expansion on an average saw approximately 73 MW higher wind capacity than the control counties. This result is about 10 MW higher than the IV estimate and about 30 MW higher than the OLS estimate highlighting the presence of counties outside the common support in both of those cases. In a similar vein, Column (2) shows that treated counties on an average had about 40 more turbines than the control counties. This is 10 turbines more than the IV estimate in Table C6. Similar to the earlier results on average project size, coefficient estimate in Column (3) lends weak evidence to the hypothesis that wind projects in CREZ counties were bigger than the projects in non-CREZ counties. The estimate implies that on an average wind projects in CREZ counties were about 33 MW bigger than projects in non-CREZ counties.

6.2.2 Robustness Checks

I conduct a series of robustness checks to see how the coefficient estimate of CREZ changes with different combinations of control variables and fixed effects. Specifically, I show how the results change when only using group fixed effects with and without CEM weights. The results for all the three dependent variables are similar to the estimates of baseline specification reported in Table 5. Appendix C presents the tables for these robustness checks.

Table 5: Effect of CREZ expansion on wind investment - matching results

	Dependent variable		
	Total Nameplate Capacity (MW)	Total Turbines	Avg. Capacity of a project (MW)
	(1)	(2)	(3)
CREZ	72.640***	39.419***	32.756*
	(26.499)	(13.075)	(19.093)
Time Trend	✓	✓	✓
Wind Controls	✓	✓	✓
Project Cost Controls	✓	✓	✓
Regulatory Controls	✓	✓	✓
Demographic Controls	✓	✓	✓
FE	Group × Trend	Group × Trend	Group × Trend
Sample	Matched	Matched	Matched
CEM Weights	✓	✓	✓
Mean Dep. Variable	35.907	16.067	26.951
Observations	344	344	344
R^2	0.467	0.476	0.426

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

7 Conclusions

Efforts to combat climate change in the US have largely focused on building generating capacity using solar and wind energy. A major factor in ensuring that power generated through these sources is fully utilized is substantial investment in transmission lines that could carry the electricity to demand centers. Availability of electricity transmission is also needed to mitigate the intermittency problem with these renewable sources. The grid failure in Texas due to the winter storms in February 2021 has also renewed calls for grid expansion so as to integrate various electricity markets in the US (Plumer 2021; Campbell 2021).

Using CREZ project in Texas as the case study, this paper provides evidence of the short- and long-run impacts of large scale grid expansion. I find that CREZ had a significant effect in integrating wind energy in the short-run, ranging from 0.10-0.22 GWh. This in turn had moderate effect on lowering markups set by fossil fuel generators. On an average, transmission expansion led to a decline of about 2-7 percent in markups. The effect of additional wind added to the grid on fossil fuel markups is strongest at the peak demand hours. This highlights the potential of electricity transmission to aid technologies like energy storage in lowering markups at these hours.

The short-run analysis on the emissions from marginal generators finds a decline in emissions to the order of $51 million annually with the majority of the reduction coming from local pollutants (SO_2 and NOx). I find that the decline in carbon emissions is spatially distributed across the four zones in Texas whereas the decline in local pollutants is primarily from the West zone. Further, there is an increase in emissions due to the intermittency of wind during the early hours of the day. This is likely due to ramping effects from coal generators operating at the margin during some of these hours. This is especially pronounced in Houston for both carbon and local pollutants.

I use an instrumental variables and matching strategy to analyze the long-run impact of CREZ expansion on wind investment. The IV estimates suggest than on an average CREZ led to about 62 MW higher wind capacity in treated counties. I use coarsened exact matching to account for the support problem and obtain a comparable group of control and treated counties. OLS regressions on the matched sample suggests that CREZ counties saw a 72 MW increase in wind capacity over 2012-2019. Further, I find weak evidence that wind projects in treated counties were about 32 MW bigger than wind projects in control counties.

These findings open up several avenues for future research. A key issue with renewable sources is that they are intermittent in nature. Research on the short-run impacts of the availability of electricity transmission for deploying energy stored from renewable sources can be informative. Another aspect is the inefficiency in investment in wind energy, especially due to over-investment in certain regions leading to congestion and curtailment of electricity from wind farms. Atal (2020) studies this issue and finds evidence of inefficiency in the timing and location of wind investments in the US. In another paper, I extend this idea to study the inefficiency due to localized over-investment in wind energy due to excess transmission.

References

Atal, Raimundo. 2020. "Too close and too soon? Spatial and dynamic inefficiency in wind energy investment." Working Paper.

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

Birge, John R, Ali Hortaçsu, Ignacia Mercadal, and J Michael Pavlin. 2018. "Limits to Arbitrage in Electricity Markets: A case study of MISO." *Energy Economics* 75:518–533.

Borenstein, Severin, James B Bushnell, and Frank A Wolak. 2002. "Measuring market inefficiencies in California's restructured wholesale electricity market." *American Economic Review* 92 (5): 1376–1405.

Borenstein, Severin, James Bushnell, Christopher R Knittel, and Catherine Wolfram. 2008. "Inefficiencies and market power in financial arbitrage: a study of California's electricity markets." *The Journal of Industrial Economics* 56 (2): 347–378.

Borenstein, Severin, James Bushnell, and Steven Stoft. 2000. "The competitive effects of transmission capacity in a deregulated electricity industry." *The Rand Journal of Economics* 31 (2): 294.

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

Campbell, Andrew. 2021. *The More We Get Together, The Happier We'll Be.* Energy Institute Blog, UC Berkeley. `https://energyathaas.wordpress.com/2021/05/03/the-more-we-get-together-the-happier-well-be-2/`.

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.

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

Draxl, Caroline, Andrew Clifton, Bri-Mathias Hodge, and Jim McCaa. 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/.

ERCOT. 2008. *CREZ Transmission Optimization Study.* Technical report. Electric Reliability Council of Texas, TX (United States).

———. 2012. "Fact Sheet." Electric Reliability Council of Texas. Retrieved from: http://www.ercot.com/content/news/presentations/2012/ERCOT_Quick_Facts_Nov_%202012.pdf.

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

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

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

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates. 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, Gary King, and Giuseppe Porro. 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.

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

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

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

Lew, Debra, G Brinkman, N Kumar, P Besuner, D Agan, and S Lefton. 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.

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`.

Plumer, Brad. 2021. "A Glimpse of America's Future: Climate Change Means Trouble for Power Grid." Retrieved from: `https://www.nytimes.com/2021/02/16/climate/texas-power-grid-failures.html`, *New York Times.*

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

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

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

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

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

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

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

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

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

Wolff, Eric. 2021. "Down to the wire: Biden's green goals face a power grid reckoning." Retrieved from: https://www.politico.com/news/2021/04/08/biden-green-goals-power-grid-480446, *Politico.*

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.

B Supplementary Figures

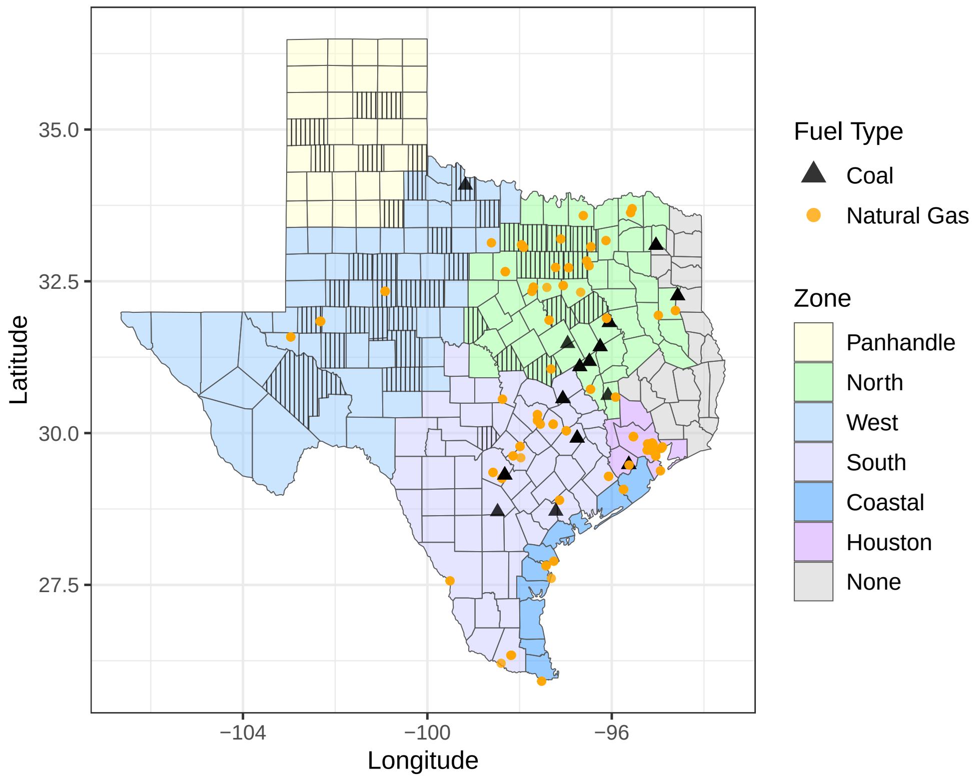
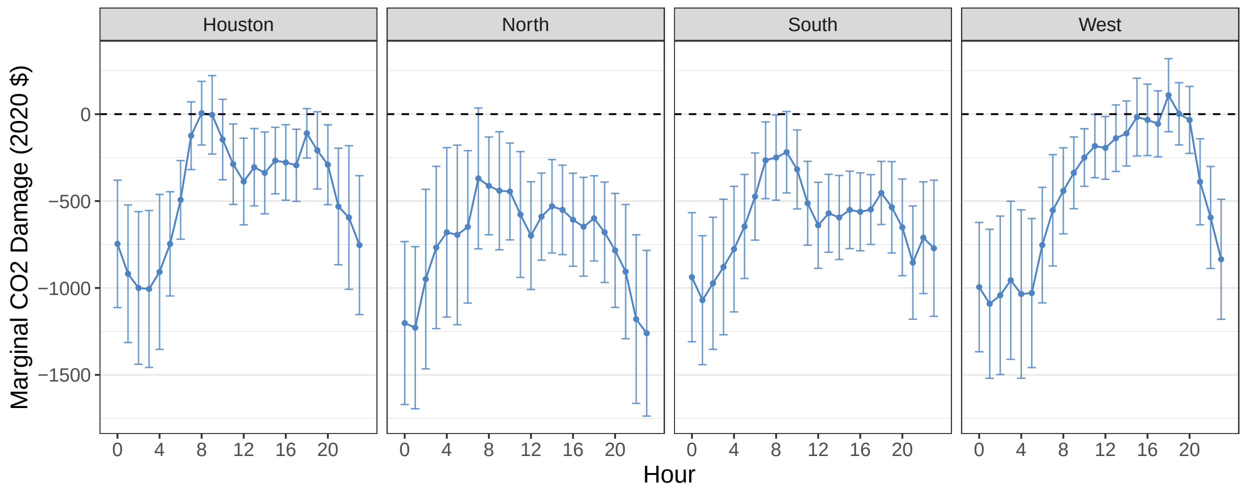
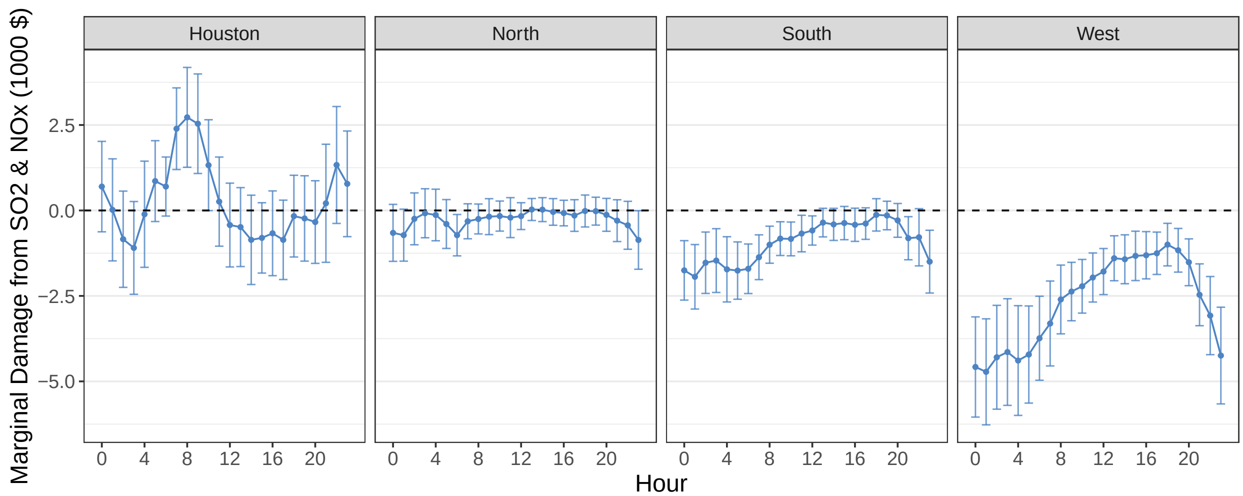


Figure B1: Coal and natural gas fired marginal generators in the sample from 2011 - 2014. Hatched counties denote the counties that received CREZ transmission expansion.



(a) Damages due to global pollution (CO_2)



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

Figure B2: Hourly averages of the marginal damages (2020 $) avoided due to CREZ expansion for each zone over 2011 - 2014.

C Supplementary Tables

C.1 Descriptive Statistics

Table C1: Descriptive statistics of key variables used in the short-run analysis of the effect of CREZ on fossil fuel generator markups

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

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

Table C2: Difference in means test of key observable characteristics of counties within and outside CREZ Zones

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

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

C.2 Robustness check results for markup analysis

Table C3: Effect of integration of 1 GWh of wind energy on fossil fuel generator markups ($/MWh)

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

This table reports the coefficient estimates of Equation 14. The dependent variable is the markup ($/MWh) set by the marginal generator i at hour t. The coefficient of interest are the interaction of hourly wind generation (GWh) and indicator of the hour. All regressions use hourly data from August 2011 to December 2014. Standard Errors clustered by generator reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table C4: Effect of CREZ completion on hourly wind generation (GWh)

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

This table reports the coefficient estimates of Equation 15. The dependent variable is total wind generation (GWh) at hour *t*. The coefficient of interest are the interaction of CREZ progress (*crez*) and indicator of the hour. All regressions use hourly data from August 2011 to December 2014. Newey-West Autocorrelation corrected standard errors with a 7 day lag structures reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

C.3 Robustness check results for OLS and IV regressions

Table C5: Effect of CREZ expansion on total wind capacity (MW) - IV results

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

Notes: Columns (1) and (2) show OLS estimation results for Equation 18 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is total nameplate capacity (MW) of wind projects in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table C6: Effect of CREZ expansion on total turbines in a county - IV results

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

Notes: Columns (1) and (2) show OLS estimation results for Equation 18 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is the total wind turbines in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

Table C7: Effect of CREZ expansion on average capacity of a wind project - IV results

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

Notes: Columns (1) and (2) show OLS estimation results for Equation 18 and Columns (3) and (4) show 2SLS estimation results for Equation IV2. The dependent variable is the total wind turbines in a county in year t. The independent variable is a binary variable indicating whether a county is CREZ or not. The IV for 2SLS regressions is whether a county was designated as one of the five RE Zones in 2007. Sample is a balanced panel of all Texas counties from 2012-2019. Time Trend is a cubic polynomial of linear time trend variable, Wind Controls include power curve, capacity factor, and cubic polynomial of wind speed (m/s). Project Cost Controls include average wind project cost, real land price, and median land acreage. Regulatory Controls include FE for PTC and wind ordinance in a county. Demographic controls include average farm size (acres) in 2007, median household income in 2007, and average population over 2007 to 2010. Robust Standard Errors reported in parenthesis. Significance: ***p<0.01;**p<0.05;*p< 0.1

C.4 Robustness check results for matching regressions

Table C8: Effect of CREZ on total wind capacity (MW)

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

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

Table C9: Effect of CREZ on Total Turbines in a County

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

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

Table C10: Effect of CREZ on Average Capacity (MW) of a wind project in a County

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

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