The Employment Impact of Wind and Solar Energy in the US

A thesis submitted by

Chao Fang

in partial fulfillment of the requirements for the degree of

Master of Science

In

Economics

Tufts University

May 2022

Advisor: Professor Gilbert Metcalf

Abstract

This paper uses Quarterly Census of Employment and Wages (QCEW) employment data and Energy Information Administration (EIA) electricity generation data to study the relationship between wind development, solar development, and employment. There is some evidence that wind development has a positive and higher employment impact in both construction and operation phases compared to solar development. The IV approach with wind speed and solar radiation data is used for the potential endogeneity problem. Policy implications can be drawn from this research that different subsidy policies may have different impacts on wind and solar energy. For example, the 1603 grant program, a cash subsidy for investment, may be higher on wind than solar, considering that wind has a higher employment impact in the construction phase. More data needs to be collected, and more research needs to confirm this conjecture.

Acknowledgment

I would like to thank Professor Gilbert Metcalf for the wonderful thesis supervision, Professor Jeffery Zabel for his Econometrics and data training, Professor Ujjayant Chakravorty for his research mentoring in environmental economics, My parents for their warm company, my friend Shengzhuo (Phineas) Yuan for helpful discussion and comments.

Contents

List of Tables:

List of Figures

1. Introduction

Renewable energy such as wind and solar electricity makes us rely less on conventional energy such as fossil fuel and increases social welfare by increasing "green job"—the employment in renewable energy. The renewable energy sector is estimated to have a significant employment impact. van der Zwaan et al. (2013) estimated that in the Middle East, wind and solar power will account for 60% of the total electricity supply in 2050. There are needs for 155,000 direct and 115,000 indirect jobs in the local workforce. While most literature on job creation is based on the input-output analysis (Fragkos & Paroussos, 2018; Ram et al., 2020; Warlick, 2009), there is limited econometrics analysis on the employment impact of renewable energy development. Therefore, there is no clear evidence of which industry creates more jobs and whether the job creation in each sector is economically significant in the general industry.

Understanding the employment impact of renewable energy is essential when evaluating the social impact of renewable subsidy policy. The renewable subsidy policy can help the development of renewable energy, thus increasing social welfare by increasing clean energy and renewable employment. The investment subsidy encourages the investment in renewable energy since it reduces the cost of investment. The output subsidy increases both the investment and long-term operation of the power plants since the output subsidy increases the profit from the renewable investment. It is important for policymakers to understand how much such investments are encouraged and how many green jobs are created from those investments when they decide to renew or introduce the subsidy policies.

Renewable energy such as wind and solar have direct employment impacts such as the operation and maintenance of power plants, and it also has indirect employment impacts such as the manufacture of equipment and component for power plants; construction and installation; and the transmission of electricity (Ram et al., 2020). Therefore, this paper will utilize employment data from the Quarterly Census of Employment and Wages (QCEW) and electricity generations data from the Energy Information Administration (EIA) to explore the employment impact of wind and solar power plants during the construction and operation phases. It will argue that different subsidies should be applied to wind and solar energy to maximize the employment impact. Instrumental variables such as wind speed and solar radiation will be used here to solve the endogeneity problem.

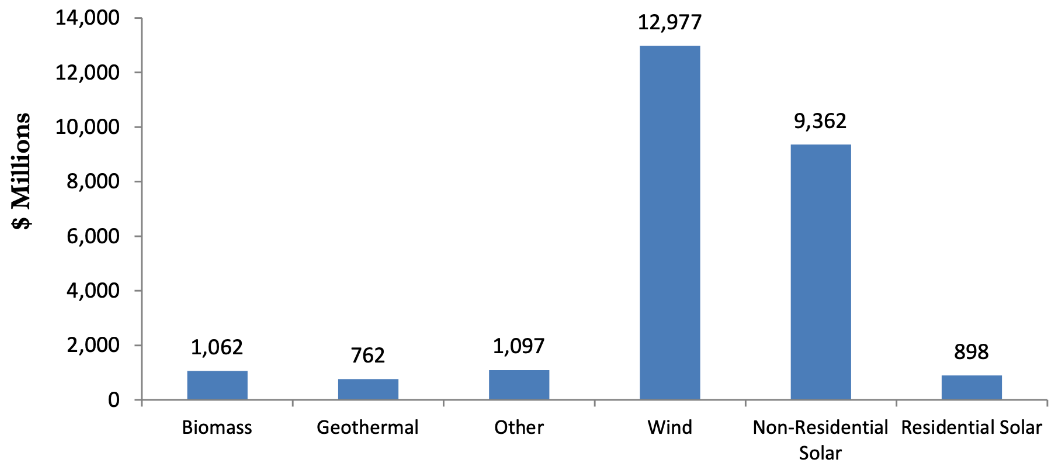
The rest of this paper is organized as follows. The following section will present a background of subsidy policy in renewable energy. A literature review on the employment impact of renewable energy is given in the third section. The fourth section explains the datasets used in this analysis, the finding from the overview of the data, and the strategy for merging them to get the final dataset of the study. The fifth and sixth sections describe the empirical strategy and results. It contains my concerns regarding research strategy and potential solutions, such as using different instrumental variable settings. The last section summarizes the results and discusses potential improvements to this research, and it connects my conclusion in this paper to the current subsidy policy for renewable energy in the U.S.

2. Background on subsidy policies in the U.S.

There are ongoing debates on whether the government should subsidize renewable energy via investment subsidy or output subsidy(Aldy et al., 2018; Alizamir et al., 2021; Johnston, 2019). It is an important topic when considering the employment impact of renewable energy as the subsidy policies have a significant impact on employment in the renewable sectors(Dell'Anna, 2021; Goolsbee, 2003), it is essential to know which policies are more suitable for increasing employment.

There are two forms of subsidies in the renewable electricity generation sector, the cash grant and the non-refundable tax credit. The non-refundable tax credit subsidizes investment for energy sources such as solar (Investment Tax Credit—ITC) and output for other energy sources such as wind (Production Tax Credit—PTC). The 1603 grant, available between 2009 and 2011, subsidizes investment with cash. The purpose of the 1603 payment was that in the several years after the 2008 financial crisis, many firms couldn't use the tax credits because they didn't have sufficient taxable income. This subsidy applies to solar, wind, geothermal, biomass, fuel cells, hydropower, combined heat and power, landfill gas, municipal solid waste, and microturbine projects. Graph 1 from the program report of the 1603 grant indicates that wind and solar are the two energy sources that benefited the most from this program (*1603 Program*, n.d.). Therefore, if a subsidy policy like the 1603 grant policy is introduced, the policy should be designed mainly according to how the policy would affect those two industries.

Figure 1—1603 grant awards by projects type (source: U.S. Department of the Treasury)



Previous research (Aldy et al., 2018; Johnston, 2019) uses the 1603 grant for wind developers to argue the subsidy choice: should the government provide the 1603 grant or the PTC? The 1603 grant, a cash grant for investment subsidy, is compared to the PTC, a non-refundable tax credit for output subsidy. Aldy et al. (2018) studied the choice between subsidizing investment or output to promote socially desirable production and found that wind farms choosing the investment subsidy generated less power per unit of capacity than those choosing the output subsidy. Johnston (2019) found that wind developers significantly discount non-refundable tax credits relative to grants. There is an efficiency loss since more tax revenue is required to reach a given level of wind energy investment. The author only focuses on choosing between the Production Tax Credit (PTC) and the 1603 grant since developers are expected to prefer the grant to the Investment Tax Credit (ITC).

Compared to previous research, the current paper will study the subsidy choice based on the employment impact of the two critical phases of wind and solar projects: the construction and operation of power plants. Investment subsidies such as the ITC and the 1603 grant should have more employment impact on the construction phase since more investment leads to more construction, while the production subsidy such as the PTC should have more job impact on the operation phase since wind farms and solar projects are motivated to operating longer to gain more from the production subsidy. Even in harsh weather conditions, it may also be worth hiring more employees to operate the wind power plant since the production subsidy increases as more electricity is generated.

3. Literature review

Previous research has provided some evidence on the wage impact of development in wind energy. Silva et al. (2013) compared the wage before and after the rapid expansion of wind power development that followed the ex-ante renewal of the PTC. They find that average payrolls for wind power generators grow relative to fossil fuel-based electricity generators after 2006. It is an indication that more investment and production in a wind power plant can increase the wage in wind electricity generation. This research only considers the direct impact of wind investment, which is the employment in the electricity generation sector. However, investment in renewable energy can also have a significant indirect impact on employment, such as jobs in manufacturing equipment, construction and installation of power plants, and transmission of electricity(Ram et al., 2020). This research is also limited since it only looks at wind farms but not other renewable energy. Those limitations bring up the question of how those indirect employment levels change from wind development and how the impact of wind farms is compared with that of solar power plants. Therefore, the current paper will investigate the employment impact of renewable energy development from both the construction and operation phases and compare the different impacts of wind and solar energy.

Previous literature on the employment impact of renewable energy focus on calculating job creation using an economic input-out model rather than using an econometrics analysis. Warlick (2009) classified job creation in wind power generation into three segments: wind turbine manufacturing, wind farm construction and development, and electric power delivery. Fragkos & Paroussos (2018) found that renewable energy sources jobs in the E.U. workforce were primarily created in the construction of solar photovoltaics, the supply and production of advanced biofuels, and the manufacturing and installation of wind turbines. Ram et al. (2020) estimated the job creation in wind power plants, and they calculated the direct and indirect job coefficients from construction and operation using collected questionnaire data.

Researchers also consider the job losses in conventional industries, such as fossil fuel electricity generation. Haerer & Pratson (2015) estimated that between 2008 and 2012, the coal industry in the U.S. lost more than 12% of jobs, while in other sectors such as natural gas, solar, and wind industries, employment increased by nearly 21%. It also happens in other counties such as Australia, Diesendorf (2004) found that the wind power industry in Australia already creates two to three times the number of direct, local job-years per kWh generated than coal power in 2002. Sohrab et al. (2019) estimated the job creation and loss during the global energy transition using the employment factor approach adopted from Rutovitz et al. (2015). This approach considers all categories of jobs created and lost during the transition from fossil fuels to renewable energy, including manufacturing, construction and installation, operation and maintenance, and job loss in fossil fuel and nuclear power plants. They concluded that the job losses in the fossil fuel and nuclear power sectors were outweighed by the job creation in the renewable power generation and storage sectors.

There is some research on the employment impact of the renewable industry using econometrics methods. Hartley et al. (2015) estimated the job-creating performance of wind, shale oil, and gas in Texas at a county level from 2001 to 2011. They did not find a significant impact of wind industry development on local employment or wages. The employment measurement they used is the total employment in all industries since they want to measure total employment effects, including indirect job creation. Compared to this research, the current paper uses more narrow definitions of employment: Power and Communication Line and Related Structures (PCL) construction and Electric Power Generation, Transmission, and Distribution (GTD) defined by the North American Industry Classification System (NAICS); Compared to total employment, those two measurements are more related to the wind and solar industries. Therefore it is more likely to find a significant impact. Another difference is that they used newly installed wind capacity as a sign of wind development. This paper uses both the capacity installed (i.e., constructed) and the capacity operated to study the construction and operating phases' impact separately. Their sample is limited to the power plants in Texas, while the current paper utilizes national-level data.

4. Data

This paper uses four primary datasets:

(1) The Quarterly Census of Employment and Wages (QCEW) data on county employment level;

(2) Electricity generation data from the Energy Information Administration (EIA) 923 Form data set.

(3)The electricity generation operating date and nameplate capacity data from EIA 860 Form dataset.

(4) The National Renewable Energy Laboratory (NREL) data on wind speed and solar radiation.

This paper also uses the county unemployment rate and county population data from the U.S. Bureau of Labor Statistics. All data are monthly data at the county level. The analysis period is limited to 2000-2020 since the EIA data is updated to 2020 at the time of writing.

4.1 QCEW employment data

The Quarterly Census of Employment and Wages (QCEW) provides county-level employment data for different industries, and data is classified using the North American Industry Classification System (NAICS). The NAICS has definitions for electricity generations, such as 221115—wind electric power generation. However, to study the employment impact of the wind and solar power plant investment, job creation during other processes related to power generation should also be included—for example, the employment in transmitting and distributing electric power. Therefore, the current paper uses employment in the NAICS industry—electricity generation, transmission, and distribution to assess the employment impact of power generation. Additionally, to study the employment impact of the investment in wind and solar power plants, this research also looks at the effects of the construction of power plants. The corresponding industry defined by NAICS is 237130—power and communication line and related structures construction.

*Table 1-Summary Statistics of monthly employment and capacity data**

Variables	Mean	Median	Minimum	Maximum	Number of observations	Number of observed counties
Construction employment	250	117	3	4,540	88,860	1,064
Operation	319	103	1	10,753	109,974**	922
Wind Construction Capacity (MW (megawatt))	35.26	0.00	0.00	2,272.70	120,204	477
Solar Construction Capacity (MW)	2.83	0.00	0.00	1,096.80	209,990	830
Other Construction Capacity (MW)	16.76	0.00	0.00	3,464.70	296,352	1,176
Wind Operation Capacity (MW)	75.30	0.00	0.00	2,543.80	117,898	466
Solar Operation Capacity (MW)	7.27	0.00	0.00	2,118.90	206,448	816
Other Operation Capacity (MW)	70.88	2.20	0.00	1,396.60	254,012	1,004

*This is the statistics for the available sample, not the whole country. For example, the statistics for construction capacity are conditional on the counties that have at least one construction capacity data.

**The QCEW data does not include every month from 2000 to 2020 for those counties, it can be seen that more data is recorded for Operation than Construction, even the Operation covers less counties.

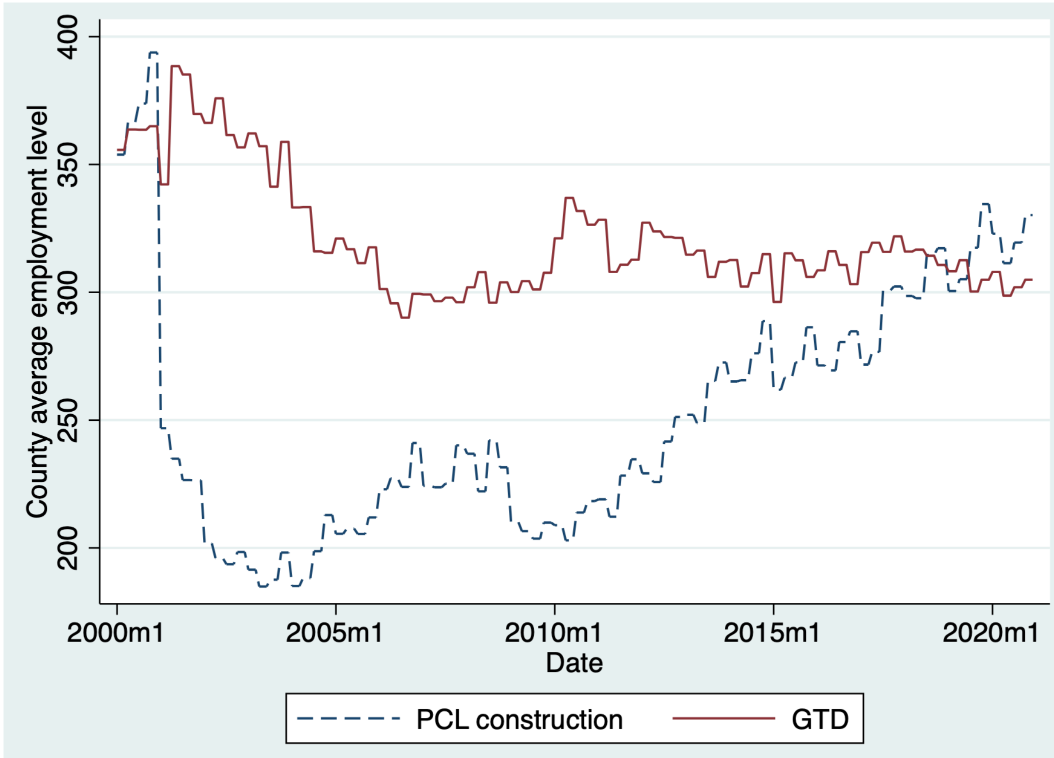
4.1.1 Power and Communication Line and Related Structures Construction

First, I collected the county employment of construction for power plants and communication lines from the QCEW. The NAICS code and industry name is 237130 Power and Communication Line and Related Structures (PCL) Construction; hydroelectric generating facilities are not included here according to its definition. Summary statistics are shown in table 1. The Construction employment level for 1,064 counties is available from the year 2000 to the year 2020. The average county employment in Construction in the sample is shown in Graph 2. From Graph 2 we can see that, the employment level after the construction phase is more stable than the employment level in the construction phase; the construction employment may be more sensitive to the subsidy policy since friendly subsidy policy can attract more investments thus more construction employment is needed. For example, we can see upward trends in Construction employment after the renewal of PTC in 2006 and the introduction of the 1603 grant program in 2009. Power and Communication Line and Related Structures (PCL) Construction employment is referred as construction employment in this paper.

4.1.2 Electric Power Generation, Transmission, and Distribution

Secondly, I collected the employment for electricity generation, transmission and distribution. The NAICS code and industry name are 2211 Electric Power Generation, Transmission, and Distribution (GTD). This is used to assess the employment impact of the power plant operation. Summary statistics are shown in table 1. The Operation level for 922 counties is available from the year 2000 to the year 2020. Electric Power Generation, Transmission, and Distribution (GTD) employment is referred as operation employment in this paper.

Figure 2—Sample average county employment in construction and operation



4.2 EIA power plant data

4.2.1 Construction capacity

County construction capacity is defined as the sum of the nameplate capacity of all power plants during the construction period in one county. The construction period for solar is defined as the twelve months before the operating date. For wind, the construction period is defined to be thirty-six months before the operating date. The definitions for these construction periods are from Ram et al. (2020). The data about nameplate capacity and operating date of the power plant come from the Form EIA-860 data. It contains data for wind, solar, and other electricity generation data, such as fossil fuel electric power generation. The construction capacity is separated for wind and solar and includes the rest of the power generation in the category—other construction capacity. The construction of hydroelectric generating facilities is excluded to be parallel with the definition of Construction employment.

The summary statistics are shown in Table 1 above. Between 2000-2020, 477 counties have wind power generation construction, 830 counties have solar power generation construction, 1,176 counties have other power generation construction.

4.2.2 Operation capacity

The county-level operation capacity is defined as the sum of all operating power plants' nameplate capacity in a county. The plant is defined as operating if the current date is after its operating date, which is when the power plant started operating. This data is from Form EIA-860.

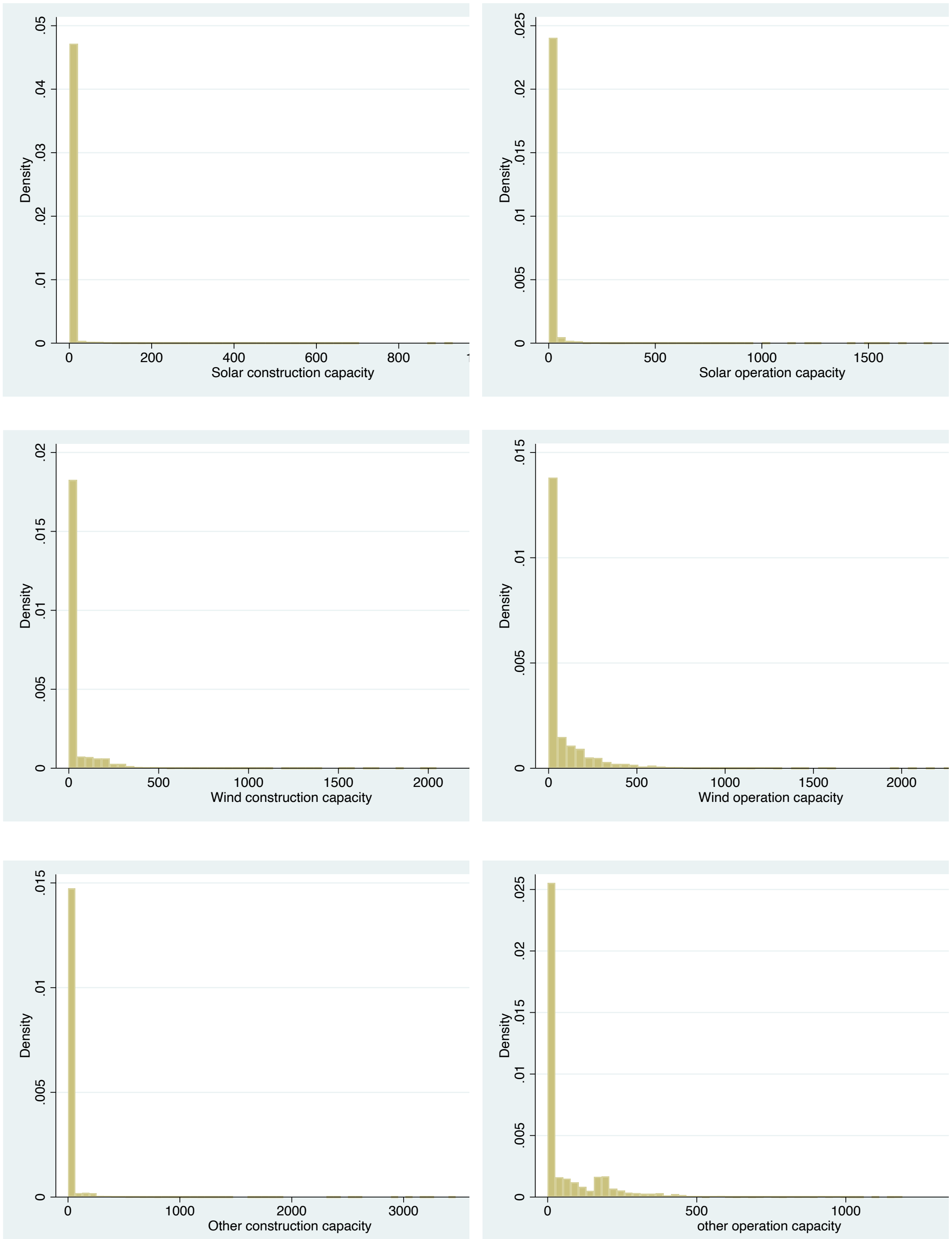
A potential concern of using this definition of operation capacity is that power plants that end their lifetime cannot be in operation. It is not a problem anymore by using the 2020 version of the EIA 860 Form since this survey only includes existing and planned power plants. Another measurement could be the net electricity generation. The net electricity generation is the gross electricity generation minus the internal electricity consumption for a production process. There are some periods when power plants do not produce electricity; thus, this value would be negative. The negative value is hard to be correlated to the operation activity, so it is not appropriate to use this as a measurement of operation to predict Operation. Also, it is assumed that the operation job in wind farms and solar plants is not short-term, so it should not fluctuate monthly with the net generation. The net generation fluctuates with weather conditions, while employment should not as long as the power plant operates. The gross generation could be a better measurement than the net generation. However, this data is only recorded by EIA-923 after 2018, so it is not usable for this research.

Another concern is that the power plants may not be operating in some periods. Thus the definition of operating capacity may not be valid during some periods when the power plant is not operating due to bad weather conditions. Therefore, the operating capacity is over-estimated, so there would be a negative bias in the coefficient estimates of operating capacity. To correct this bias, I replace the operating capacity with zero when the net generation of one power plant is negative. It is assumed that when the power plant has zero gross generation when the net generation is zero, it is based on the data between 2018 to 2020 when the EIA-923 records both gross generation and net generation. It is also assumed that when the gross generation is zero, the power plant is not operating; thus, no workers are hired.

The summary statistics are shown in Table 1 above, noting that these statistics are conditional on the counties that have wind, solar power plants operation or other power plants operation. Between 2000-2020, 466 counties have wind power generation operations, 816 counties have solar power generation operations, 1,004 counties have other power generation operations.

From Graph 3, we can see that both construction and operation capacity dataset is right-skewed and contain a lot of zero values, many counties do not have construction or operation in many periods from 2000 to 2020.

Figure 3—Histograms of construction and operation capacities



4.3 NREL climate condition data

The final dataset comes from NREL National Solar Radiation Database (NSRDB). Its online tool—*PVWatts Version 6*, contains the annual solar radiation value, the capacity factor for solar power plants (calculated based on solar radiation), and the yearly average hourly wind speed for climate stations in the United States. Climate condition data are collected for each city and town in the United States. The county-level data used in this paper is the simple average value for cities and towns in that county.

The mean of wind speed and solar radiation for the full sample; conditional on positive construction and conditional on zero construction are shown in Table 2, the wind speed and solar radiation are higher in those counties that have wind or solar projects than those that do not, so, it is assumed that the investment of solar and wind projects in one county depends on its solar radiation or wind speed.

Table 2—County average wind speed and solar radiation

Variable	Sample mean	Mean conditional on positive Wind/Solar construction	Mean conditional on no Wind/Solar construction
Wind speed (m/s)	1.60	2.16	1.43
Annual solar radiation(kWh/m2/day)	5.17	5.21	5.17

4.4 Final dataset for Construction and Operation analysis

4.4.1 Construction dataset

To conduct the research for both Construction employment and Operation, the datasets described above are merged into two datasets—the construction dataset and the operation dataset. The construction dataset includes Construction employment, solar construction capacity (in MW); wind construction capacity (in MW); other construction capacity (in MW); average wind speed; average solar radiation, the population, and the unemployment rate for 697 counties. The final dataset excludes the counties that have capacity data but not employment data since it is not possible that there is construction going on in one county and there are no workers hired. Therefore, there must be a missing record of the employment data of this county. If employment is zero for this county, there would be bias in the size of the estimated employment impact. The counties that have PCL employment data but not capacity data are included since the Construction employment measurement consists of the industry communication line. If there is no construction capacity for the power plant, the construction of a communication line could also be positive, so there is positive PCL employment data. The final dataset contains 18.7% of observations with no capacity but positive employment. Summary statistics are given in Table 3. We can see that both medians of employment data and the capacity data are smaller than the means, which indicate that the data is highly skewed, so the log value is used in regression. There are a lot of zero value in the capacity data since for some counties, there are many periods when no construction of wind or solar project is going on. The inverse hyperbolic sine transformation is used here so that I can take the log of those zero values.

Table 3—Summary statistics for construction dataset

Variable	Mean	Median	Standard Deviation	Min	Max
Construction employment	240.75	105.00	383.39	3.00	4,540.00
Solar Construction Capacity (MW)	2.86	0.00	25.02	0.00	1,096.80
Wind Construction Capacity (MW)	5.37	0.00	43.66	0.00	2,044.70
Other Construction Capacity (MW)	20.15	0.00	102.90	0.00	3,464.70
County average wind speed(m/s)	1.50	1.52	0.89	0.00	5.26
County average solar radiation(kWh/m2/day)	5.19	5.22	0.44	3.99	6.77
County population	444,470	196,788	824,523	1,801	1.01×10^7
County unemployment rate(percentage)	5.70	5.10	2.52	0.90	36.00

Number of observations: 72,494; Number of counties: 930

4.4.2 Operation dataset

The operation dataset includes Operation; solar operation capacity (in MW); wind operation capacity (in MW); other operation capacity (in MW); average wind speed; average solar radiation, the population, and the unemployment rate for 623 counties. The final dataset excludes the counties that have capacity data but no employment data since if power plants are operating in one county, there should exist workers employed for electric power generation, transmission, or distribution. Hence, the employment data for this county is not zero but not recorded by the QCEW. The counties with Operation data but not capacity data are included since there may not be a power plant operating in one county but there could exist employment for the transmission and distribution of electricity. The final dataset contains 21.7% of observations with no capacity but employment data. Summary statistics are given in Table 4. Similar to the construction dataset, both medians of the employment data and the capacity data are less than the means, so the data is highly skewed, the log of value is used in regressions. The inverse hyperbolic sine transformation is also applied in the zero values of capacity data.

Table 4—Summary statistics for the operation dataset

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Operation	406.95	155.00	828.23	2.00	10,753.00
Solar Operation Capacity (MW)	35.44	0.00	135.63	0.00	2,543.80
Wind Operation Capacity (MW)	9.76	0.00	76.10	0.00	2,118.90
Other Operation Capacity (MW)	105.11	2.30	184.82	0.00	1182.60
County average wind speed(m/s)	1.44	1.43	0.90	0.00	5.26
County average solar radiation(kWh/m2/day)	5.13	5.04	0.48	4.10	6.83
County population	384,351	144,646	705,226	2,089	$1.009*10^7$
County unemployment rate	6.01	5.40	2.76	0.90	36.00

Number of observations: 66,267; Number of counties: 584

5. Empirical Strategy

5.1 Main regressions

My research question is about the difference in the employment impact of subsidies for wind and solar power projects. By answering this question, policy suggestions can be made on which sector should receive more focus on subsidy policy, such as the 1603 grant. To answer this question, I am planning to run separate regressions for Construction employment (i.e., power and communication line and related structures construction) and Operation (i.e., the employment for electricity generation, transmission and distribution):

(1) Construction employment regressions:

$$y_{it} = \beta_0 + \beta_1 WCC_{it} + \beta_2 SCC_{it} + \gamma' X_{it} + \mu_m + \eta_y + w_i + \epsilon_{it} \quad (1)$$

Where: $WCC_{it} = \sum_{j \in i} capacity_j * wind\ construction_{jt}$;

$SCC_{it} = \sum_{j \in i} capacity_j * solar\ construction_{jt}$;

X_{it} is the vector of variables includes:

$\sum_{j \in i} capacity_j * other\ construction_{jt}$; $unemploy\ rate_{iy}$;

$population_{iy}$

Where i indexes county, y indexes year, m indexes 12 months of the year, and t indexes monthly observation. y_{it} is the Construction employment for county i at time t. WCC_{it} and SCC_{it} are the construction capacity for wind or solar power plants in county i at time t, it is the sum of the nameplate capacity of the power plants that is under construction in county i. $capacity_j$ is the nameplate capacity for power plant j. $wind\ construction_{jt}$ and $solar\ construction_{jt}$ are the constructions dummies for wind and solar plant j at time t; $other\ construction_{jt}$ is the construction of other power plants such as fossil fuel electricity generation. $unemploy\ rate_{it}$ is the monthly unemployment rate for county i at year y. $population_{iy}$ is the population of county i at year y. μ_m is the month fixed effect. η_y is the year fixed effect. w_i is the county fixed effect. ϵ_{it} is the error term. γ is a vector of coefficients for variables in X_{it}.

(2) Operation regressions:

$$y_{it} = \beta_0 + \beta_1 WOC_{it} + \beta_2 SOC_{it} + \gamma' X_{it} + \mu_m + \eta_y + w_i + \epsilon_{it} \qquad (3)$$

Where: $WOC_{it} = \sum_{j \in i} capacity_j * Wind\ operation_{jt}$;

$SOC_{it} = \sum_{j \in i} capacity_j * Solar\ Operation_{jt}$;

X_{it} is the vector of variables includes:

$other\ employment_{it}$; $unemploy\ rate_{iy}$; $population_{iy}$

Where i indexes county, y indexes year, m indexes month, and t indexes monthly date. y_{it} is the employment for electricity generation, transmission, and distribution. WOC_{it} and SOC_{it} are the operation capacity for wind or solar power plant j at time t, it is the sum of all operating power plants' nameplate capacity. $capacity_j$ is the nameplate capacity for power plant j. $Wind\ Operation_{jt}$ and $Solar\ Operation_{jt}$ are dummy variables that indicate whether there is any power plant operation for wind and solar power plant j separately, it is a value of 1 if the time t is after operating date of plant j. $other\ employment_{it}$ is the sum of employment in electricity generations other than wind and solar electricity generation. $unemploy\ rate_{iy}$ is the monthly unemployment rate for county i at year y. $population_{iy}$ is the population of county i at year y. μ_m is the month fixed effect. η_y is the year fixed effect. w_i is the county fixed effect. ϵ_{it} is the error term. γ is a vector of coefficients for variables in X_{it}.

5.2 Instrumental variable strategy

The main regression has a potential endogeneity problem: (1) The investment decision could be affected by the county Construction employment since low county-level employment could attract more investment in renewable energy. For example, some states have financial incentives and renewable portfolio standards to support wind development (Black et al., 2014). (2) The investment in renewable energy such as wind and solar electricity generation could potentially "crowd out" the employment in other electricity generation sectors such as fossil fuel electricity generation, or be "crowed out" by those sectors. Since those variables are not available and not included in the main regression, there may be endogeneity. The current paper uses two instruments—wind speed and solar radiation to solve this endogeneity issue.

5.2.1 Instruments and first stage regression

The instrument for wind construction and operating capacity is the wind speed, and the one for solar construction and operating capacity is the solar radiation level. Those two instruments can be used because of the following two assumptions: (1) The investment decisions of the wind farm and solar plant depend on local wind speed and solar radiation level. (2) The instrument wind speed and solar radiation only affect employment through its effect on the endogenous variables.

For the first-stage Construction employment regression, I'm running regressions of construction capacity on wind speed/solar radiation. The regressions for operation capacity are similar.

$$WCC_{it} = \pi_0 + \pi_1 wind\ speed_i + AX_{it} + \eta_y + \mu_m + \epsilon_{it} \quad (3)$$

$$SCC_{it} = \pi_0 + \pi_1 solar\ radiation_i + AX_{it} + \eta_y + \mu_m + \epsilon_{it} \quad (4)$$

Where i indexes county, and t indexes monthly date. μ_m is the month fixed effect. η_y is the year fixed effect. ϵ_{it} is the error term. Since the capacity data contain a lot of zeros, it would be problematic if I take the log directly. Therefore, here I'm using the inverse hyperbolic sine transformation. The transformation for variable z would be $In(z + \sqrt{z^2 + 1})$.

5.2.2 Issues with the first stage analysis

Two issues with the instrument are that:

(1) The instrument is county-specific characteristics. If I run the above equations with a county fixed effect, the instruments—wind speed and solar radiation, which are time-invariant characteristics of counties, would be dropped out and have no prediction power on the endogenous variables.

(2) The capacity variables are time-variant variables, while the instruments are time-invariant. This paper uses time-invariant wind speed and solar radiation data, it is based on the assumption that those climate conditions do not change much

over the years, and the investment decision of wind farms and solar plants should depend on the annual estimate of wind speed and solar radiation instead of monthly climate condition. A potential solution to this issue could be allowing the coefficients of wind speed and solar radiation to vary over time.

Six options for the first stage are used here to explore the potential solution to those issues, the first stage for wind construction capacity is used here as the example, the settings for wind operation capacity and solar construction/operation capacity are the same.

$$WCC_{it} = \alpha_0 + \alpha_{1t}Wind_i + \alpha_{2t}SR_i + \alpha_{3t}Wind_i * SR_i + AX_{it} + \mu_m + \eta_y + w_i + e_{it} \quad (5)$$

Option 1: This is a model with a county fixed effect and time-varying coefficients for wind speed and solar radiation. SR_i is the solar radiation for county i. The time-variant coefficients allow me to apply county fixed effect here, and the instruments would not drop out. This first stage includes both instruments and their interaction.

$$WCC_{it} = \alpha_0 + \alpha_1 Wind_i * X_{it} + \alpha_2 SR_i * X_{it} + \alpha_3 Wind_i * SR_i * X_{it} + AX_{it} + \mu_m + \eta_y + w_i + e_{it} \quad (6)$$

Option 2: This is a model with county fixed effect and let the wind speed and solar radiation to be interacted with X_{it}, here the coefficients is not allowed to be varied across time, but county fixed effect can still be applied since now the new instruments includes interactions with X_{it}, which makes the instruments to be time-variant. I also include the interaction of wind speed, solar radiation and X_{it}.

$$WCC_{it} = \alpha_0 + \alpha_1 Wind_i + \alpha_2 SR_i + \alpha_3 Wind_i * SR_i + AX_{it} + \mu_m + \eta_y + e_{it} \quad (7)$$

Option 3: This is a model without a county fixed effect and time-varying coefficients, and this is based on the assumption that investors' preferences for wind speed and solar radiation do not vary across time. The interaction between two instruments is included. The coefficients for instruments are likely to be biased since the OLS estimators pick up the impact of other county characteristics when the fixed effect is not included. Still, it may not be an issue here in the first stage of 2SLS since only the first stage regression's predicted value is used in the structural equation. A county fixed effect is not used here, which assumes county fixed effect is not helpful since wind speed/solar radiation are county-specific characteristics, and they can have higher prediction power to the construction capacity without county fixed effect.

$$WCC_{it} = \alpha_0 + \alpha_1 Wind_i * X_{it} + \alpha_2 SR_i * X_{it} + \alpha_3 Wind_i * SR_i * X_{it} + AX_{it} + \mu_m + \eta_y + e_{it} \quad (8)$$

Option 4: This is a model without county fixed effect and time-varying coefficients, but the instruments are the interactions of wind speed/solar radiation and X_{it}, this assume county fixed effect is not helpful here since wind speed/solar radiation are county-specific characteristics and they can have higher prediction power to the construction capacity without county fixed effect.

$$WCC_{it} = \alpha_0 + \alpha_1 Wind_i * Year\ Trend_{it} + \alpha_2 SR_i * Year\ Trend_{it} + \alpha_3 Wind_i * SR_i * Year\ Trend_{it} + AX_{it} + \mu_m + \eta_y + w_i + e_{it} \quad (9)$$

Option 5: $Year\ Trend_{it} = year - 1999$, i.e., $Year\ Trend_{it} = 1$ for $year = 2000$; This is a model with a county fixed effect and time-invariant coefficients for wind speed and solar radiation, and the instruments are the interactions of wind speed/solar radiation and $Year\ Trend_{it}$.

$$WCC_{it} = \alpha_0 + \alpha_1 Wind_i * Trend_{it} + \alpha_2 SR_i * Trend_{it} + \alpha_3 Wind_i * SR_i * Trend_{it} + AX_{it} + \mu_m + \eta_y + w_i + e_{it} \quad (9)$$

Option 6: $Trend_{it} = year - 1999$, i.e., $Trend_{it} = 1$ for $year = 2000\ and\ month = 1$; This is a model with a county fixed effect and time-invariant coefficients for wind speed and solar radiation, and the instruments are the interactions of wind speed/solar radiation and $Trend_{it}$.

5.3 Placebo Test

To test the robustness of the result from above models, the current paper uses a different definition of construction and operation period as a placebo: The "Placebo Construction period" for wind energy is defined as four years before the time when actual construction started, the "Placebo Construction period" for solar project is defined as two years before the actual construction started. I add one additional year to how long it actual takes to construct for wind and solar just to extend the "grace period" of placebo. The "Placebo operation period" is defined as the year prior when construction started.

5.4 Event Study

$$y_{it} = \beta_0 + \sum_{l=-5}^{6} \beta_l WOC(l)_{it} + \sum_{m=-5}^{6} \beta_m SOC(m)_{it} + \gamma' X_{it} + \mu_m + \eta_y + w_i + \epsilon_{it} \quad (10)$$

Where: $WOC_{it} = \sum_{j\in i} capacity_j * Wind\ operation_{jt}$;

$SOC_{it} = \sum_{j\in i} capacity_j * Solar\ Operation_{jt}$;

X_{it} is the vector of variables includes:

$other\ employment_{it}$; $unemploy\ rate_{iy}$; $population_{iy}$

$WOC(l)_{it}$ is the lth lead or lag of WOC, $SOC(m)_{it}$ is the mth lead or lag of SOC

To see if there is a pre-trend in operation employment for wind and solar development, I include 5 leads and 6 lags of the operation capacity variables. The event study is only performed for operation employment because of the evidence of pre-trend in the placebo test for operation employment.

6. Empirical Results

6.1 Construction employment regression

Table 5 includes positive results for the Construction employment regression. Only the positive results are analyzed here for policy implication since I believe those negative results are biased, it is based on the assumption that more construction of wind and solar should increase the Construction employment level; the full regression results, including other covariates and negative results, are shown in Table 9 and 10 of the Appendix, reasons for rejecting those negative results are provided in the appendix.

The results in column 1 to column 3 in table 5 are positive. If assuming the result without IV is causal, then it can be concluded that the construction of wind projects has a higher employment impact than solar projects. Still, it is not likely to be casual because of the endogeneity problem. If we assume the result in columns 2 and column 3 of Table 5 to be causal, it can be concluded that wind projects construction may have a higher employment impact.

Table 5—Regression result for Construction regression

	(1)	(2)	(3)
VARIABLES	Log[Construction employment(y)]		
Log(SCC)	0.0048*	0.0813***	0.0714***
	(0.0026)	(0.0277)	(0.0277)
Log(WCC)	0.0088***	0.3899***	0.3778***
	(0.0025)	(0.0684)	(0.0684)
Constant	4.6006***	4.6533***	4.6451***
	(0.0202)	(0.0304)	(0.0304)
Using IV†	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
F-test†† for 1st stage IV (Solar)	N/A	315.83	314.75
F-test for 1st stage IV (Wind)	N/A	49.74	49.40
Observations	72,497	72,497	72,497
R-squared	0.0859	0.0862	0.0861
Number of counties	930	930	930

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
County, year, month fixed effects and other variables X_{it} are included in the second stage.
†County fixed effect is included in the first stage of 2SLS.
†† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

From the result in column 2, it can be interpreted that 1% increase in the solar construction in a county would lead to 0.08% increase in the Construction employment in that county on average and all else equal, it is statistically significant under 1% significance level. 1% increase in the wind construction in a county would lead to 0.39% increase in Construction employment in that county, on average and all else equal. It is statistically significant under 1% significance level. The economics significance is not discussed here since research indicates that when the inverse hyperbolic sine transformation is used, the actual elasticity can be substantially different with the elasticity from using the natural log (Bellemare & Wichman, 2020).

The results in those four columns provide some evidence to the conjecture that the wind construction may have a higher and more significant employment impact than the solar construction. Those results are similar thus provide some robustness to our conclusion. However, it is different with the estimation in Rutovitz et al. (2015), in which it is estimated that utility-scale PV(Photovoltaics) requires 13.00 job-yrs/MW in construction and installation, while wind onshore only requires 3.20 job-yrs/MW in construction and installation, wind offshore requires 8.00 job-yrs/MW in construction and installation. This conclusion is different with our results, which could be the reason that the labor in installation is not included in the analysis of the current paper. Another potential reason is that my results have their drawbacks: there is potential endogeneity problem in column 1 since no instrumental variable is used; column 2 and column 3 use time trend while there is no evidence that there is a trend of the effect of instrumental variables. More analysis using better datasets is needed to confirm my conjecture here.

6.2 Operation regression

Table 6 includes positive results for Operation regression. I am only analyzing the positive result for policy implication here based on the assumption that there is positive relationship between operation activity and Operation. Full regression results including other covariates and negative results are shown in Table 11 and 12 of the Appendix, reasons for rejecting the negative results are provided in the appendix.

The endogeneity problem arises from the assumption that the low Operation area attracts more investment in the electricity industry, especially renewable energy. Therefore, there should be a negative bias of capacity's impact on employment. Column 1 is the original regression without using the IV. After using IV, from the result in columns 2 and column 3, we can see now there are larger impacts. Therefore, it is what I expected that the negative bias is corrected by using the IV.

Table 6—Regression result for operation regression

	(1)	(2)	(3)
VARIABLES	Log[Operation employment(y)]		
Log(SOC)	0.0159***	0.0372***	0.0374***
	(0.0015)	(0.0046)	(0.0046)
Log(WOC)	0.0105***	0.0636***	0.0637***
	(0.0013)	(0.0067)	(0.0067)
Constant	4.9092***	4.8889***	4.8890***
	(0.0146)	(0.0159)	(0.0159)
Using IV†	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
Interaction in 1st stage	N/A	No	No
F-test††for 1st stage IV (Solar)	N/A	2551.07	2553.75
F-test for 1st stage IV (Wind)	N/A	901.48	897.62
R-squared	0.0123	0.0118	0.0118

Standard errors in parentheses; number of observations: 66,267; number of counties: 584
*** p<0.01, ** p<0.05, * p<0.1
County, year, month fixed effects and other variables X_{it} are included in the second stage.
†County fixed effect is included in the first stage of 2SLS.
††Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

From the result in column 2, it can be interpreted that 1% increase in the solar operation of a county would lead to 0.04% increase in Operation in that county on average and all else equal, it is statistically significant under 1% significance level. 1% increase in wind operation of a county would lead to a 0.06% increase in Operation of that county on average and all else equal, it is statistically significant under the 1% significance level. The economics significance is not discussed here since research indicates that when the inverse hyperbolic sine transformation is used, the actual elasticity can be substantially different with the elasticity from using the natural log (Bellemare & Wichman, 2020).

This result provides some evidence for the conjecture that wind power plant operations may have a higher employment impact than the solar project. This conjecture was also made by other literature using job estimation method. According to Rutovitz et al. (2015), wind onshore and offshore averagely requires 0.25 jobs /MW for operation and maintenance, while utility-scale PV(Photovoltaics) requires 0.800 job-year/MW; But utility-scale PVrequires 0.800 job-year/MW for decommissioning while wind onshore and offshore averagely requires 1.855 job-yea1r/MW Therefore, wind operation might have a higher employment impact in the sector "Electric Power Generation, Transmission and Distribution". There is drawback with using the year trend or monthly trend, both yearly trend and monthly trend interaction model (option 5 and 6) for both solar and wind fails the test for putting restriction on time-variant coefficient model, so we do not have strong reason to put the restriction on the time-variant coefficient model so that the wind speed and solar radiation is interacted with trend. More analysis would need to be done to confirm or refute this conjecture.

6.3 Placebo Test Result

6.3.1 Placebo test for construction regression

The results for construction using IV are generally robust. The placebo result indicates that the results in column 2 and column 3 of Table 6 are robust since the results in the placebo test are insignificant. The result for column 1 is not robust, it is probably because of the endogeneity problem.

Table 7—Placebo test for construction regression

	(1)	(2)	(3)
VARIABLES	Log[Construction employment(y)]		
Log(SCC)	0.0036	-0.0986	-0.1056
	(0.0022)	(0.0719)	(0.0712)
Log(WCC)	0.0263***	-0.0849	-0.0886
	(0.0024)	(0.1338)	(0.1328)
Constant	4.6338***	4.5717***	4.5667***
	(0.0199)	(0.0409)	(0.0406)
Using IV†	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
F-test†† for 1st stage IV (Solar)	N/A	174.22	174.94
F-test for 1st stage IV (Wind)	N/A	58.58	58.59
R-squared	0.0855	0.0841	0.0841

Standard errors in parentheses; number of observations: 72,504; number of counties: 930

*** p<0.01, ** p<0.05, * p<0.1

County, year, month fixed effects and other variables X_{it} are included in the second stage.

†County fixed effect is included in the first stage of 2SLS.

†† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

6.3.2 Placebo test for operation regression

Table 8—Placebo test result for operation regression

	(1)	(2)	(3)
VARIABLES	Log[Operation employment(y)]		
Log(SOC)	-0.0113***	-0.0451***	-0.0453***
	(0.0017)	(0.0065)	(0.0065)
Log(WOC)	-0.0109***	-0.0552***	-0.0551***
	(0.0012)	(0.0059)	(0.0059)
Constant	4.9303***	5.0513***	5.0516***
	(0.0156)	(0.0204)	(0.0204)
Using IV†	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
Interaction in 1st stage	N/A	No	No
F-test†† for 1st stage IV (Solar)	N/A	216.71	1584.43
F-test for 1st stage IV (Wind)	N/A	37.16	971.13
R-squared	0.0106	0.0110	0.0110

Standard errors in parentheses; number of observations: 66,267; number of counties: 584

*** p<0.01, ** p<0.05, * p<0.1

County, year, month fixed effects and other variables X_{it} are included in the first and second stage.

†County fixed effect is included in the first stage of 2SLS.

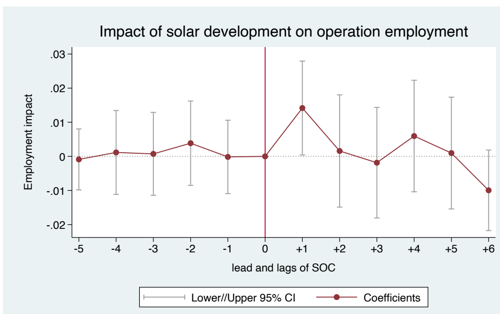
†† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

The placebo result indicates that the results in column 1 to column 3 of Table 7 are all negative and statistically significant under 1% significance level. It indicates that there may be a negative pre-trend in operation employment before the construction of a wind or solar power plant in a county. More analysis about pre-trend is done in the following section to determine the entire impact of wind and solar development.

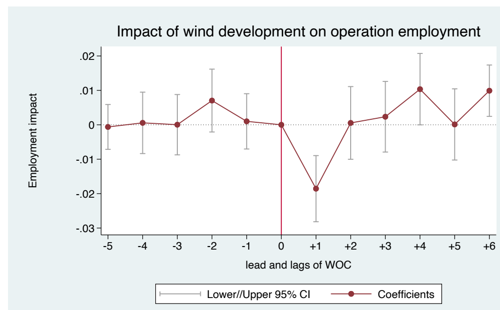
6.3.3 The event study for operation employment

The event study results are shown in figure 4, we can see the period before period 0 are zero(or zero is in the confidence interval), so we can see there is no clear evidence about pre-trend for wind and solar operation.

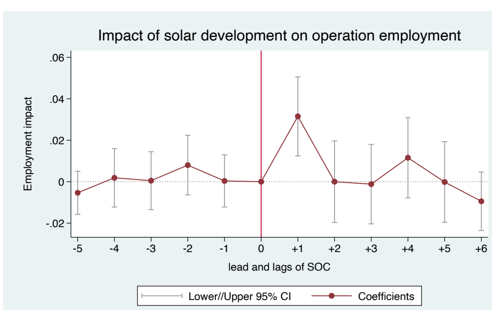
Figure 4 Event study for pre-trend analysis



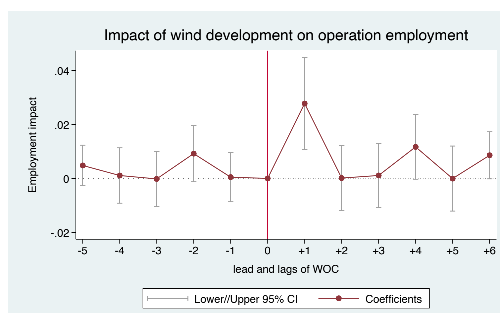
Event study for solar using OLS



Event study for wind using OLS



Event study for solar using year trend IV



Event study for wind using year trend IV

7. Conclusion and improvement

7.1 Conclusion

This paper uses the employment data from QCEW and the electricity generation data from EIA to investigate the job creation in two main phases of wind and solar power projects—the construction phase and the operation phase. Instrumental variables, including wind speed and solar radiation, are used to solve the potential endogeneity problem.

From the result of the Construction regression, it can be conjected that the construction of wind projects may create more employment compared to solar projects construction; A similar conclusion is drawn from the result of Operation regression, where the 2SLS result indicates a larger employment impact from wind operation than from solar operation.

The subsidy policy should consider its employment impact since subsidy can significantly increase wages, thus increasing "green job"—jobs in the renewable energy sector (Dell'Anna, 2021; Goolsbee, 2003); The policymaker for production subsidies such as PTC should pay more attention to its impact on wind energy considering the larger impact of wind farm operation on Operation relative to the solar power plant. The policymaker for investment subsidies such as the ITC and the 1603 grant should consider its employment impact, with a focus on wind energy considering its larger employment impact on the Construction sector compared to solar power.

7.2 Improvement

Improvement on datasets could lead to more rigorous results:

1. Establishment level employment data could help estimate the job creation from wind and solar operations more precisely; Silva et al. (2013) used confidential establishment-level employment data from QCEW to examine the employment impact of production subsidy. If this dataset is available, it can be combined with the EIA-923 Form data to study the relationship between employment level and electricity generation on an establishment-level instead of the county-level analysis in this paper.
2. A large amount of counties' power plant data is recorded by EIA, but QCEW does not record the Construction and Operation data in those counties. Therefore, the sample used in the current analysis is not the population. An improvement in the matching of these two datasets could lead to a more rigorous result.

References

1603 Program: Payments for Specified Energy Property in Lieu of Tax Credits. (n.d.). U.S. Department of the Treasury. Retrieved March 28, 2022, from https://home.treasury.gov/policy-issues/financial-markets-financial-institutions-and-fiscal-service/1603-program-payments-for-specified-energy-property-in-lieu-of-tax-credits

Aldy, J. E., Gerarden, T. D., & Sweeney, R. L. (2018). *Investment versus output subsidies: Implications of alternative incentives for wind energy*. National Bureau of Economic Research.

Alizamir, S., Iravani, F., & Yücel, Ş. (2021). *Investment in Wind Energy: The Role of Subsidies* (SSRN Scholarly Paper ID 3868573). Social Science Research Network. https://doi.org/10.2139/ssrn.3868573

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50–61. https://doi.org/10.1111/obes.12325

Black, G., Holley, D., Solan, D., & Bergloff, M. (2014). Fiscal and economic impacts of state incentives for wind energy development in the Western United States. *Renewable and Sustainable Energy Reviews*, *34*, 136–144. https://doi.org/10.1016/j.rser.2014.03.015

Dell'Anna, F. (2021). Green jobs and energy efficiency as strategies for economic growth and the reduction of environmental impacts. *Energy Policy*, *149*(C).

https://ideas.repec.org/a/eee/enepol/v149y2021ics0301421520307424.html

Diesendorf, M. (2004). Comparison of employment potential of the coal and wind power industries. *International Journal of Environment, Workplace and Employment*, *1*(1), 82–90. https://doi.org/10.1504/IJEWE.2004.005605

Fragkos, P., & Paroussos, L. (2018). Employment creation in EU related to renewables expansion. *Applied Energy*, *230*, 935–945. Scopus. https://doi.org/10.1016/j.apenergy.2018.09.032

Goolsbee, A. (2003). Investment Subsidies and Wages in Capital Goods Industries: To the Workers Go the Spoils? *National Tax Journal*, *56*(1.2), 153–165. https://doi.org/10.17310/ntj.2003.1S.02

Haerer, D., & Pratson, L. (2015). Employment trends in the U.S. Electricity Sector, 2008–2012. *Energy Policy*, *82*, 85–98. https://doi.org/10.1016/j.enpol.2015.03.006

Hartley, P. R., Medlock, K. B., Temzelides, T., & Zhang, X. (2015). Local employment impact from competing energy sources: Shale gas versus wind generation in Texas. *Energy Economics*, *49*, 610–619. https://doi.org/10.1016/j.eneco.2015.02.023

Johnston, S. (2019a). Nonrefundable tax credits versus grants: The impact of subsidy form on the effectiveness of subsidies for renewable energy. *Journal of the Association of Environmental and Resource Economists*, *6*(3), 433–460.

Johnston, S. (2019b). Nonrefundable Tax Credits versus Grants: The Impact of Subsidy Form on the Effectiveness of Subsidies for Renewable Energy. *Journal of the Association of Environmental and Resource Economists*, *6*(3), 433–460. https://doi.org/10.1086/702736

Ram, M., Aghahosseini, A., & Breyer, C. (2020). Job creation during the global energy transition towards 100% renewable power system by 2050. *Technological Forecasting and Social Change*, *151*, 119682. https://doi.org/10.1016/j.techfore.2019.06.008

Rutovitz, J., Dominish, E., & Downes, J. (2015). *Calculating global energy sector jobs: 2015 methodology*.

Silva, D. G. D., McComb, R. P., & Schiller, A. R. (2013). Do production subsidies have a wage incidence in wind power? *Applied Economics*, *45*(28), 3963–3972. https://doi.org/10.1080/00036846.2012.741679

Sohrab, T., Rafiee, M., Karkoodi, S., & Parvin, M. (2019). Estimation of the employment rate of wind power plants in Iran in the horizon of 2050. *International Journal of Ambient Energy*, *0*(0), 1–7. https://doi.org/10.1080/01430750.2019.1653974

Steinberg, D., Porro, G., & Goldberg, M. (2012). *Preliminary Analysis of the Jobs and Economic Impacts of Renewable Energy Projects Supported by the ..Section..1603 Treasury Grant Program* (NREL/TP-6A20-52739). National Renewable Energy Lab. (NREL), Golden, CO (United States). https://doi.org/10.2172/1038342 van der Zwaan, B., Cameron, L., & Kober, T. (2013). Potential for renewable energy jobs in the Middle East. *Energy Policy*, *60*, 296–304. https://doi.org/10.1016/j.enpol.2013.05.014

Warlick, D. (2009). Attractive employment potential in the wind power sector. *Oil and Gas Journal*, *107*(21), 4–7. Scopus.

Appendix:

A: Full second stage results

Table 9—Full positive regression result for Construction regression

VARIABLES	(1)	(2)	(3)
	Log[Construction employment(y)]		
Log(SCC)	0.0048*	0.0813***	0.0714***
	(0.0026)	(0.0277)	(0.0277)
Log(WCC)	0.0088***	0.3899***	0.3778***
	(0.0025)	(0.0684)	(0.0684)
Log(other construction capacity)	0.0159***	0.0177***	0.0175***
	(0.0014)	(0.0015)	(0.0015)
Unemployment rate	-0.0198***	-0.0210***	-0.0209***
	(0.0014)	(0.0014)	(0.0014)
Population	0.0000***	0.0000*	0.0000**
	(0.0000)	(0.0000)	(0.0000)
Constant	4.6006***	4.6533***	4.6451***
	(0.0202)	(0.0304)	(0.0304)
Using IV	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
County fixed effect in 1st stage	N/A	Yes	Yes
Interaction	N/A	No	No
F-test† for 1st stage IV (Solar)	N/A	315.83	314.75
F-test for 1st stage IV (Wind)	N/A	49.74	49.40
Observations	72,497	72,497	72,497
R-squared	0.0859	0.0862	0.0861
Number of counties	930	930	930

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
County, year, month fixed effects and other variables X_{it} are included in the second stage.
† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

Table 10—Full regression result for other construction regressions

	(1)	(2)	(3)	(4)
VARIABLES	Log[Construction employment(y)]			
Log(SCC)	-0.0165	-0.1475***	2.8705***	-0.1693***
	(0.0139)	(0.0133)	(0.1359)	(0.0437)
Log(WCC)	-0.0375**	-0.0727	-11.4216***	-0.0516**
	(0.0177)	(0.0554)	(0.5544)	(0.0229)
Log(other construction capacity)	0.0156***	0.0136***	0.3136***	0.0208***
	(0.0014)	(0.0014)	(0.0148)	(0.0017)
Unemployment rate	-0.0196***	-0.0188***	0.0590***	-0.0114***
	(0.0014)	(0.0014)	(0.0037)	(0.0024)
Population	0.0000***	0.0000***	-0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	4.5836***	4.4700***	4.6210***	4.5316***
	(0.0235)	(0.0229)	(0.0190)	(0.0264)
Using IV	Yes	Yes	Yes	Yes
IV coefficient is time-variant	Yes	No	No	No
County fixed effect in 1st stage	Yes	Yes	No	No
Interaction	No	X_{it}	No	X_{it}
F-test† for 1st stage IV (Solar)	3.31	353.05	478.76	352.10
F-test for 1st stage IV (Wind)	1.94	18.80	523.79	281.02
Observations	72,497	72,497	72,497	72,497
R-squared	0.0858	0.0874	0.0857	0.0861
Number of counties	930	930	930	930

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

County, year, month fixed effects and other variables X_{it} are included in the second stage.

† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

Potential explanation for negative results in construction regression:

Most results using IV in Table A2 are negative, which are not what I expected and are likely to be biased since more construction in power plants requires more labor in the Construction sector. Therefore, the Construction employment should increase as construction capacity in wind and solar projects increases.

There are some potential explanations for those negative results. The result in column 1 has a weak instrument problem so the instrument is not working. Column 2 does not allow the coefficients for instruments to be time-variant, this could lead to the negative result. Then we can see column 3 and 4 do not use county fixed effect, thus the characteristic across counties could lead to those negative results.

Another potential explanation for these negative results could be: according to the assumption and estimation in Ram et al. (2020), there is job loss in conventional power plants such as fossil fuel and nuclear electricity generation during the energy transition from nonrenewable energy to renewable energy. Since the construction employment of conventional power plants is not controlled in this analysis, there could be a negative bias in the estimation of renewable construction, and negative results could come from this bias.

The datasets for the construction regression have two issues that could also lead to this negative result:

1. Construction includes not only power construction but also communication line construction, so I cannot control for variations in communication line construction.
2. The match between Construction employment and construction capacity is not perfect. There are 83.07% of capacity data not matched with employment data, which means many counties have power plant constructions, but QCEW has not recorded their employment level in the Construction sector. Therefore, the sample used in this paper is not the whole sample but the selected sample, the result is likely to be biased since the selected sample may miss some positive effect of wind and solar investment.

Table 11—Full positive regression result for operation regression

	(1)	(2)	(3)
VARIABLES	Log[Operation employment(y)]		
Log(SOC)	0.0159***	0.0372***	0.0374***
	(0.0015)	(0.0046)	(0.0046)
Log(WOC)	0.0105***	0.0636***	0.0637***
	(0.0013)	(0.0067)	(0.0067)
Log(other operation capacity)	-0.0090***	-0.0076***	-0.0076***
	(0.0012)	(0.0012)	(0.0012)
Unemployment rate	-0.0057***	-0.0062***	-0.0062***
	(0.0009)	(0.0009)	(0.0009)
County population	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)
Constant	4.9092***	4.8889***	4.8890***
	(0.0146)	(0.0159)	(0.0159)
Using IV	No	Yes	Yes
IV coefficient is time-variant	N/A	Year trend	Monthly Trend
County fixed effect in 1st stage	N/A	Yes	Yes
Interaction in 1st stage	N/A	No	No
F-test†† for 1st stage IV (Solar)	N/A	2551.07	2553.75
F-test for 1st stage IV (Wind)	N/A	901.48	897.62
Observations	66,267	66,267	66,267
R-squared	0.0123	0.0118	0.0118
Number of counties	584	584	584

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
County, year, month fixed effects and other variables X_{it} are included in the second stage.

††Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

Table 12—Full regression result for other operation regressions

	(1)	(2)	(3)	(4)
VARIABLES	Log[Operation employment(y)]			
Log(SOC)	0.0414***	-0.0261***	-1.4110***	0.1786***
	(0.0041)	(0.0050)	(0.1252)	(0.0148)
Log(WOC)	0.0506***	0.0921***	1.1826***	0.0345***
	(0.0066)	(0.0108)	(0.1080)	(0.0078)
Log(other operation capacity)	-0.0080***	-0.0057***	0.1761***	-0.0077***
	(0.0012)	(0.0013)	(0.0168)	(0.0015)
Unemployment rate	-0.0061***	-0.0062***	0.0378***	-0.0150***
	(0.0009)	(0.0009)	(0.0039)	(0.0012)
County population	0.0000***	0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	4.9248***	4.8069***	4.4218***	4.9615***
	(0.0146)	(0.0184)	(0.0449)	(0.0166)
Using IV	Yes	Yes	Yes	Yes
IV coefficient is time-variant	Monthly variant	No	No	No
County fixed effect in 1st stage	Yes	Yes	No	No
Interaction in 1st stage	No	X_{it}	No	X_{it}
F-test† for 1st stage IV (Solar)	13.14	958.16	1119.67	687.61
F-test for 1st stage IV (Wind)	3.74	152.31	1341.67	614.48
Observations	66,267	66,267	66,267	66,267
R-squared	0.0117	0.0106	0.0113	0.0131
Number of counties	584	584	584	584

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

County, year, month fixed effects and other variables X_{it} are included in the second stage.

† Here I use the rule of thumb that it passes the weak instrument test if F-test is larger than 10.

Potential explanation for negative results in operation regression:

Results for SCC in column 2 and column 3 are negative, which are not what I expected and are likely to be biased since more operation in power plants requires more labor in the operation sector.

There are some potential explanations for those negative results. The settings in column 2 and column 3 do not allow the coefficients for instruments to be time-variant, this could lead to the negative result.

B: First stage results

All first stages results are included here except the regressions that allow coefficients of instruments to be time-variant, there are 240 coefficients for each regression. Those regression are not included here due to page limitation.

Table 13—First stage results for option 2 and option 4 in construction regression

IV option	(1) Option 2	(2) Option 2	(3) Option 4	(4) Option 4
VARIABLES	SCC	WCC	SCC	WCC
Solar Radiation ×	0.0196*	0.0233**	0.0011	0.1693***
Log(other construction capacity)	(0.0100)	(0.0104)	(0.0100)	(0.0116)
Wind Speed ×	0.1142***	0.0088	0.0613**	0.3299***
Log(other construction capacity)	(0.0289)	(0.0299)	(0.0283)	(0.0329)
Wind Speed ×	-0.0240***	-0.0024	-0.0138**	-0.0641***
Solar Radiation ×				
Log(other construction capacity)	(0.0056)	(0.0058)	(0.0055)	(0.0064)
Solar Radiation ×	0.0065	0.0002	0.0349***	-0.0640***
Unemployment Rate	(0.0064)	(0.0066)	(0.0029)	(0.0034)
Wind Speed ×	0.0110	-0.0651***	-0.0014	-0.2899***
Unemployment Rate	(0.0181)	(0.0188)	(0.0087)	(0.0101)

Wind Speed × Solar Radiation × Unemployment Rate	-0.0013	0.0112***	0.0002	0.0600***
	(0.0036)	(0.0037)	(0.0017)	(0.0020)
Solar Radiation × Population	0.0000***	-0.0000***	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Wind Speed × Population	0.0000***	-0.0000***	-0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Wind Speed × Solar Radiation × Population	-0.0000***	0.0000***	0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log(other construction capacity)	-0.0873*	-0.1186**	0.0370	-0.8501***
	(0.0511)	(0.0529)	(0.0509)	(0.0593)
Unemployment Rate	-0.0396	0.0114	-0.1381***	0.3076***
	(0.0327)	(0.0338)	(0.0150)	(0.0175)
Population	-0.0000***	0.0000***	-0.0000*	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.3438***	0.0889***	-0.3207***	-0.0482
	(0.0331)	(0.0343)	(0.0253)	(0.0294)
Observations	72,497	72,497	72,497	72,497
R-squared	0.1542	0.0147	0.1791	0.0476
Number of counties	930	930	N/A	N/A

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 14—First stage results for option 3 in construction regression

IV option	(1) Option 3	(2) Option 3
VARIABLES	SCC	WCC
Solar Radiation	0.3142***	-0.1637***
	(0.0161)	(0.0187)
Wind Speed	0.0725	-0.9929***
	(0.0448)	(0.0519)
Wind Speed × Solar Radiation	-0.0235***	0.2144***
	(0.0088)	(0.0102)
Log(other construction capacity)	0.0253***	0.0324***
	(0.0020)	(0.0024)
Unemployment Rate	0.0439***	0.0179***
	(0.0017)	(0.0019)
Population	0.0000***	-0.0000***
	(0.0000)	(0.0000)
Constant	-1.9193***	0.5649***
	(0.0843)	(0.0976)
Observations	72,497	72,497
R-squared	0.1599	0.0353

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 15—First stage results for option 5 and option 6 in construction regression

IV option:	(1) Option 5	(2) Option 5	(3) Option 6	(4) Option 6
VARIABLES	SCC	WCC	SCC	WCC
Solar Radiation ×Trend	0.0573***	-0.0313***	0.0048***	-0.0026***
	(0.0026)	(0.0027)	(0.0002)	(0.0002)
Wind Speed ×Trend	0.0593***	-0.0819***	0.0049***	-0.0068***
	(0.0074)	(0.0075)	(0.0006)	(0.0006)
Wind Speed × Trend	-0.0127***	0.0166***	-0.0010***	0.0014***
Solar Radiation	(0.0015)	(0.0015)	(0.0001)	(0.0001)
Log(other construction capacity)	-0.0116***	-0.0022	-0.0116***	-0.0022
	(0.0020)	(0.0021)	(0.0020)	(0.0021)
Unemployment Rate	0.0058***	0.0016	0.0056***	0.0016
	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Population	0.0000***	0.0000	0.0000***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.9685***	0.1938***	-0.7048***	0.0516*
	(0.0323)	(0.0330)	(0.0302)	(0.0308)
Observations	72,497	72,497	72,497	72,497
R-squared	0.1282	0.0144	0.1282	0.0144
Number of counties	930	930	930	930

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 16—First stage results for option 2 to option 4 in operation regression

IV option	(1) Option 2	(2) Option 2	(3) Option 4	(4) Option 4
VARIABLES	SOC	WOC	SOC	WOC
Solar Radiation ×	0.0247*	-0.0225	0.2044***	0.0938***
Log(other construction	(0.0149)	(0.0175)	(0.0084)	(0.0143)
capacity)				
Wind Speed ×	-0.2922***	-0.5826***	0.0945***	-0.1995***
Log(other construction	(0.0446)	(0.0523)	(0.0244)	(0.0414)
capacity)				
Wind Speed ×	0.0533***	0.1171***	-0.0253***	0.0266***
Solar Radiation ×				
Log(other construction	(0.0089)	(0.0105)	(0.0048)	(0.0082)
capacity)				
Solar Radiation ×	-0.0822***	0.0180**	0.0223***	-0.1956***
Unemployment Rate	(0.0069)	(0.0081)	(0.0041)	(0.0070)
Wind Speed ×	-0.0962***	0.0950***	-0.0159	-0.5271***
Unemployment Rate	(0.0197)	(0.0231)	(0.0117)	(0.0198)
Wind Speed ×	0.0219***	-0.0200***	0.0022	0.1185***
Solar Radiation ×				
Unemployment Rate	(0.0039)	(0.0046)	(0.0023)	(0.0040)
Solar Radiation ×	0.0000***	0.0000***	0.0000***	0.0000***
Population				
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Wind Speed × Population	-0.0000**	-0.0000	0.0000***	0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Wind Speed × Solar Radiation × Population	-0.0000**	0.0000	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Log(other construction capacity)	-0.1062	0.0432	-0.9839***	-0.5020***
	(0.0727)	(0.0854)	(0.0416)	(0.0706)
Unemployment Rate	0.3986***	-0.0735*	-0.0580***	0.8779***
	(0.0344)	(0.0404)	(0.0207)	(0.0351)
Population	-0.0000***	-0.0000***	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-0.6436***	1.1644***	-0.4075***	0.3774***
	(0.0407)	(0.0478)	(0.0383)	(0.0651)
Observations	66,267	66,267	66,267	66,267
R-squared	0.3875	0.1773	0.3273	0.1590
Number of counties	584	584	N/A	N/A

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 17—First stage results for option 3 in operation regression

IV option	(1) Option 3	(2) Option 3
VARIABLES	SOC	WOC
Solar Radiation	0.7488***	-0.7234***
	(0.0225)	(0.0377)
Wind Speed	0.5137***	-2.8882***
	(0.0628)	(0.1052)
Wind Speed × Solar Radiation	-0.1338***	0.6594***
	(0.0126)	(0.0211)
Log(other construction capacity)	0.0154***	-0.1010***
	(0.0017)	(0.0029)
Unemployment Rate	0.0640***	0.0369***
	(0.0021)	(0.0035)
Population	0.0000***	-0.0000***
	(0.0000)	(0.0000)
Constant	-4.0107***	3.0716***
	(0.1155)	(0.1934)
Observations	66,267	66,267
R-squared	0.2948	0.1423

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 18—First stage results for option 5 and option 6 in operation regression

IV option:	(1) Option 5	(2) Option 5	(3) Option 6	(4) Option 6
VARIABLES	SOC	WOC	SOC	WOC
Solar Radiation ×Trend	0.1509***	-0.0648***	0.0125***	-0.0054***
	(0.0030)	(0.0034)	(0.0002)	(0.0003)
Wind Speed ×Trend	0.0822***	-0.2737***	0.0067***	-0.0228***
	(0.0089)	(0.0102)	(0.0007)	(0.0009)
Wind Speed × Trend	-0.0239***	0.0606***	-0.0020***	0.0050***
Solar Radiation	(0.0018)	(0.0021)	(0.0001)	(0.0002)
Log(other construction capacity)	0.0065**	-0.0282***	0.0065**	-0.0282***
	(0.0030)	(0.0035)	(0.0030)	(0.0035)
Unemployment Rate	0.0023	0.0035	0.0016	0.0032
	(0.0022)	(0.0026)	(0.0022)	(0.0026)
Population	0.0000***	-0.0000***	0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	-1.3457***	1.2448***	-0.6891***	0.9973***
	(0.0397)	(0.0459)	(0.0366)	(0.0423)
Observations	66,267	66,267	66,267	66,267
R-squared	0.3793	0.1934	0.3794	0.1932
Number of counties	584	584	584	584

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$