Inefficiency in Californian Solar Markets

Ellen Osborn

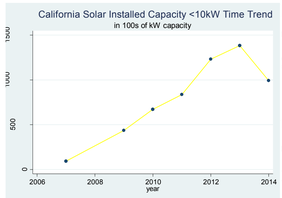
5/10/2016

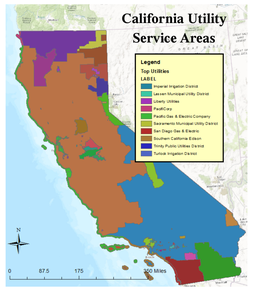
GIS102 - Advanced GIS

Prof. Srinivasan

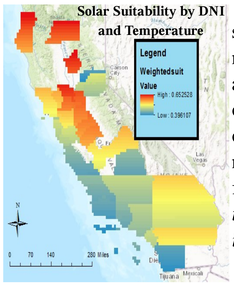
Introduction

In the summer of 2015, the regulatory addition of the EPA's ambitious Clean Power Plan pushed energy policy reform to the forefront of many state regulators' priorities. The EPA assigned each state a unique goal to reduce a portion of emissions contributing to climate change from their energy sectors, and policymakers are scrambling to develop plans to meet these goals. Some have sought examples of successful programs in other states, for whom California might seem to be a leader. Increasing industry incumbency paired with quickly dropping installation costs have helped solar companies to become legitimate competitors in the state's energy generation markets.





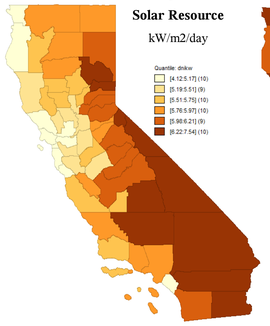
Solar photovoltaic installations rose quickly from the early California Solar Initiative in 2007 to the present. With only about 400MW of generating capacity installed at the beginning of the program, costing on average over $15/watt, solar was not a generally viable option as an energy source. However, by August 2015, California had installed another 3,000MW across the state, with new installations averaging a third of the original cost.

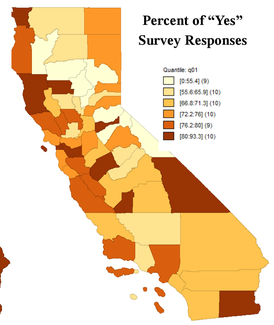


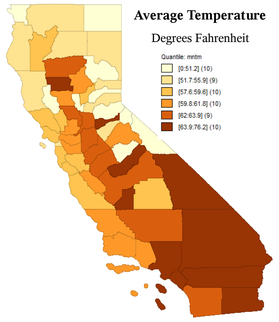
Solar photovoltaic technology relies on solar energy radiated to the surface of the earth, and panel efficiency has remained relatively stagnant over the course of the program, depending primarily on the irradiance and average annual temperature of the installation location. Efficiency optimization of energy policy is of interest because of the politically unrewarding nature of relatively fair energy policy to date, such as market based priced-in values on the negative externalities of fossil fuels. The question this project intends to answer is, *what factors drive the installed solar capacity of counties in California, and what policy inefficiencies could be rectified?*

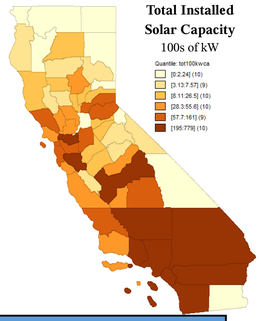
Data

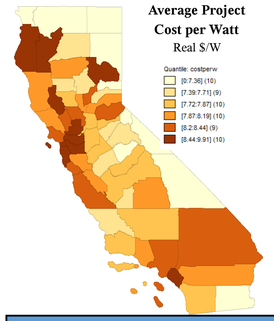
This project uses a variety of data sources geocoded and compiled in a yearly panel from 2009 to 2014 to determine relationships between the measured variables described below. The five main components of the data are the financial and technological data of systems all determined at the county level: county survey question results, county characteristics, average solar resource available, and average temperature. These are used to estimate the effectiveness of solar project installation policy in California. All cumulative data averaged over the period of 2009-2014 is included in the maps below to show the spatial distribution of each variable, although all regression estimates studied are done on a separated annual basis.











Methods

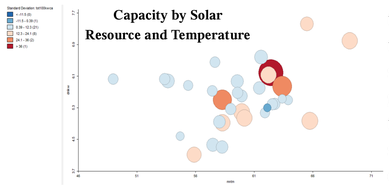
The primary hypothesis of this project is that there is inefficiency in the renewable energy policy structure in Californian solar markets, and that this inefficiency will be demonstrated through a significant positive effect of voting data over solar resource and temperature data. It predicts that the coefficient of the survey question variable would be positively and strongly correlated with county installed capacity, greater than that of solar resource (which should still be positive), and temperature (which should be negative), controlling for other project and county characteristics.

Rated panel capacity (Wp)	Temperate (° C)	Temperature Coefficient	Effective panel capacity (Wp)	Change in Wp
250	20	-0.45%	255.59	102.24%
250	25	-0.45%	250.00	100.00%
250	35	-0.45%	238.83	95.53%
250	45	-0.45%	227.65	91.06%

Less obvious than solar resource, however, is that temperature should in fact decrease as solar installations increase. While it is important to note that annual aggregations do not capture seasonal variation effectively, county average annual temperatures strongly affect solar panel efficiency. The chart to the left shows the estimated decreases in efficiency expected for temperature variation off of the nameplate ratings of solar panels. Large increases in average annual temperature have very significant effects on the efficiency of electricity production, and should therefore affect installation decisions.

Other variables such as economic and persona; characteristics of households to have varied effects. Median income (in thousands of real dollars, medinc1000), total county population and number of housing units are some that demonstrate these. The estimated spatial model is:

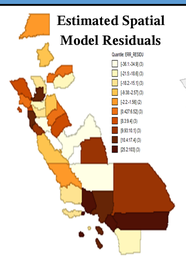
$$\text{totkwcap}_i = \beta_0 + \beta_1 \text{q01}_i + \beta_2 \text{dnikw}_i + \beta_3 \text{mntm}_i + \beta_4 \text{medinc1000}_i + \beta_5 \text{costperw}_i + \beta_6 \text{tot\_housin}_i + \beta_7 \text{tot10kpop}_i + u_i$$

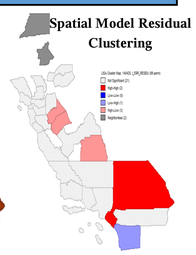


This bubble chart of 2009 data illustrates the basic expectations of an effective allocation of solar resource. Rationally large bubbles should appear in the top left corner of the graph, representing high DNI values and low temperatures. However, the graph does not show any visible trends in that direction or any other, indicating deviation from the optimal.

Results

In 2014, the spatial error model yielded further interesting results. A relatively significant (p=0.18) coefficient on q01 of 68kW adds to evidence of surveyed opinion overly factoring into capacity installations. The insignificant temperature (mntm) regressor furthers this idea - although solar resource again outshines both with the large and significant coefficient of 310.7kW of installed capacity. The strongly significant lambda value of 0.702 implies that spatial autocorrelation is strongly at play in 2014. This could be a result of realistically blurred boundaries when it comes to true county characteristics (overlapping utility service territories, geographic characteristics, etc.), or more likely evidence of omitted explanatory variables.





The first map to the right shows the residual errors in the 2014 spatial error regression, the only regression for which the spatial correction is strongly significant. The clustering of these errors implied in the map to the far right tell a story of the model is failing to predict capacity accurately. The large High-High and Low-High clusters in the south could be explained by the exclusion of large projects over 10kW from the dependent variable of total installed capacity, which could have a negative effect on the proliferation of smaller installations in these southern counties where large desert solar fields are more likely. The High-Low counties in the middle of the state, Tulare, Merced, and Stanislaus, could be experiencing high errors in a field of low due to their low number of available neighbors (2 each) or the exclusion of unknown variables that explain their installed capacities outside the parameters of the model.

Yearly Drivers of Installed Capacity by County
Spatial Error Model

	tot100kwcap 2009	tot100kwcap 2010	tot100kwcap 2011	tot100kwcap 2012	tot100kwcap 2013	tot100kwcap 2014
q01	0.2668 (0.0185)***	0.04597 (0.7225)	0.2495 (0.1488)*	0.1781 (0.4388)	-0.6935 (0.0373)***	0.6841 (0.181)*
dnikw	6.1238 (0.001)***	10.2341 (0.001)***	15.6257 (0.000)***	27.2059 (0.000)***	38.415 (0.000)***	31.0709 (0.0062)***
mntm	-0.2707 (1.04202)	0.2645 (0.6178)	0.1026 (0.8735)	1.9258 (0.000)***	4.2342 (0.000)***	0.0371 (0.9282)
Lambda	-0.2306 (0.3442)	0.1115 (0.6275)	0.145 (0.5219)	-0.1793 (0.462)	0.3262 (0.1191)*	0.7021 (0.000)***
R2	0.76	0.78	0.82	0.90	0.92	0.84
N	33	32	33	33	32	29

* *p*<0.05; ** *p*<0.01
Note: *P* values reported in parentheses
Note: Other regressors not included in this display; costperw, medinc1000, tot10kpop, tot_housin10k

Conclusions and Future Work

The core question of this study, *what exactly is the economic harm done by inefficient renewable energy policy,* remains unanswered by the data and current methodology. The theoretical implementation of simplified energy policy, such as the elimination of fossil fuel subsidies and a fair carbon price, might result in a more closely correlated relationship between the solar resource, temperature, and the installed capacity in any given county, as opposed to current project allocations which rely on the externalities of voter preference for reduced emissions and other factors.

Due to the restricted data and panel nature of this model, the highest fit achieved of any iteration was R^2 = 0.64 with about 21 degrees of freedom. One option for a different version of this test might be foregoing a time trend analysis and looking at county numbers over a few aggregated years only. This way, survey questions that were not asked every year that had better response rates could be used, or even voting records on issues like the extension of the CA AB32 program that more closely relate to the proliferation of solar financial support in California. This method would also reduce the errors from the omission of time-reliant policies that expire or go into effect sporadically over the years included, skewing data in certain counties in any given time period. Solar projects are also largely placed in counties with higher electricity prices, where distributed generation can make a larger dent in electricity bills more quickly.

The exclusion of this information from the model, due to the lack of geographic coding currently available, could be a main driver of the above results' mediocre explanatory power.

Data Sources:

SDTS Point and Vector Object Type: G-polygon
Point and Vector Object Count: 58
Latitude Resolution: 0.000000
Longitude Resolution: 0.000000
Geographic Coordinate Units: Decimal degrees\
Horizontal Datum Name: North American Datum of 1983
Ellipsoid Name: Geodetic Reference System 80
Semi-major Axis: 6378137.000000
Denominator of Flattening Ratio: 298.257222

Temperature Data: NOAA Weather Stations
Survey Data: Public Policy Institute of California (PPIC)
Solar resource data: NREL
Solar project data: California Solar Initiative (CSI)
County Data: American Community Survey (ACS)

Acknowledgements:

Prof. Sumeeta Srinivasan
Prof. Ujjayant Chakravorty
Prof. Jeffrey Zabel
Prof. Ann Rappaport

For advice and support in the production of the Senior Honors Thesis upon which this project is based.