Heterogeneous Impacts of State-Level Residential Solar Rebate Programs in the U.S.

Bixuan Sun, University of Minnesota, sunxx731@umn.edu

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

Heterogeneous Impacts of State-Level Residential Solar Rebate Programs in the U.S.

Bixuan Sun (sunxx731@umn.edu), University of Minnesota

Motivation

Growing amount of literature examines the determinants of increased residential solar photovoltaic (PV) capacity.

- Policies: Solar Renewable Energy Credit, Renewable Portfolio Standards, California Solar Initiative subsidy program.
- Non-policy factors: peer effect, demographic characteristics, such as family size, education and income.

Little is known about the interactions among policies and non-policy factors. Empirically testing a model with a set of predetermined interaction terms might omit some important relationships and yield misleading results.

Research Question

How do the impacts of state-level solar rebate programs on the number of residential solar installations vary by other policies and non-policy factors?

* Rebate programs offer cash subsidies towards the solar installation cost, usually on a dollar-per-KW-capacity basis.

Data

Main Data Sources:

- The Open PV project (National Renewable Energy Lab): roughly 85% of solar projects in the U.S. by 2015.
- Database of State Incentives for Renewables & Efficiency: solar renewable energy credit, solar system tax credit, production tax credit, Renewable Energy Portfolio (RPS).
- Free the Grid report: state level interconnection and net metering scores.
- American Community Survey: demographic variables and housing characteristics (county level)

Final Data Set: zip code level panel data from 2007 to 2015, covering 10,283 zip codes in 48 states. The number of observations is 45,152.

Method: Causal Regression Trees

Causal regression trees is a data-driven approach developed by Athey and Imbens (2016)* that partitions data into subgroups that differ in the magnitude of program impact. This method allows the identification of heterogenous treatment effects without pre-specifying any functional forms.

Causal regression trees is developed based on the assumption of unconfoundedness (i.e., randomized treatment). I use inverse propensity score weighting to adjust the selection into treatment.

* Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353-7360.

Descriptive Statistics

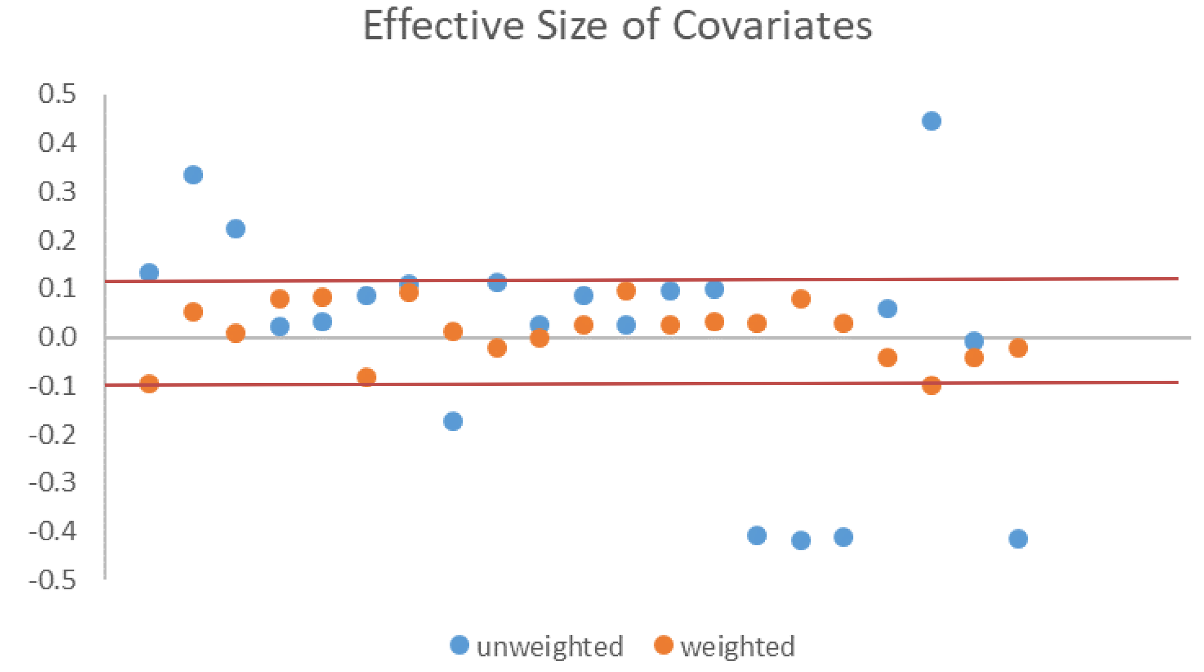
Variable	Mean	Std. Dev.	Variable	Mean	Std. Dev.
Number of solar installations	19.76	51.17	Median income (county)	62373.04	14905.82
Have solar rebate program	0.81	0.39	Percent of bachelor degree (county)	0.18	0.05
Rebate level ($/W)	1.21	1.24	Average household size (county)	2.76	0.26
Net metering score (0-5)	4.52	1.10	Percent of GDP from mining industry (state)	0.05	1.15
Interconnection score (0-5)	3.65	1.19	Have RPS	0.98	0.14
Have SREC program	0.33	0.47	RPS percent (if RPS=1)	25.54	7.50
Have system tax rebate program	0.23	0.42	RPS ending year (if RPS=1)	2019	41.95
Have production tax rebate program	0.16	0.37	RPS establishing year (if RPS=1)	2001	41.60
Average annual solar insolation (1998-2009)	5.10	0.75	RPS has solar carve out	0.38	0.48
Cost of solar installation ($/W)	6.03	1.89	Solar carve out percent (if SCO=1)	1.98	1.77
Residential electricity rate ($/kWh)	0.15	0.03	Solar carve out ending year (if SCO=1)	2020	4.99
Housing density (county)	506.74	1385.21	Senate conservation voting score (1-100)	79.37	27.64
Population density (county)	1071.19	2839.93	House conservation voting score (1-100)	66.55	19.57
Median house value (county)	300047.50	149085.20	Wind energy capacity (80 meters, 2008)	51992.27	239117.40

Balancing Test

Generalized boosted regression, a machine learning method, is used to estimate the propensity scores.

Effective size (or absolute standardized mean difference) is used to evaluate the balancing propensity:

$$\frac{\bar{X}_{treated} - \bar{X}_{control}}{\sqrt{({s_{treated}}^2 + {s_{control}}^2)/2}}$$



Results

Distribution of ATTs (Average Treatment Effects on the Treated)
The honest causal regression trees method uses half of the data to build regression trees and the other half to estimate ATTs. The estimation sample size is 22,576.

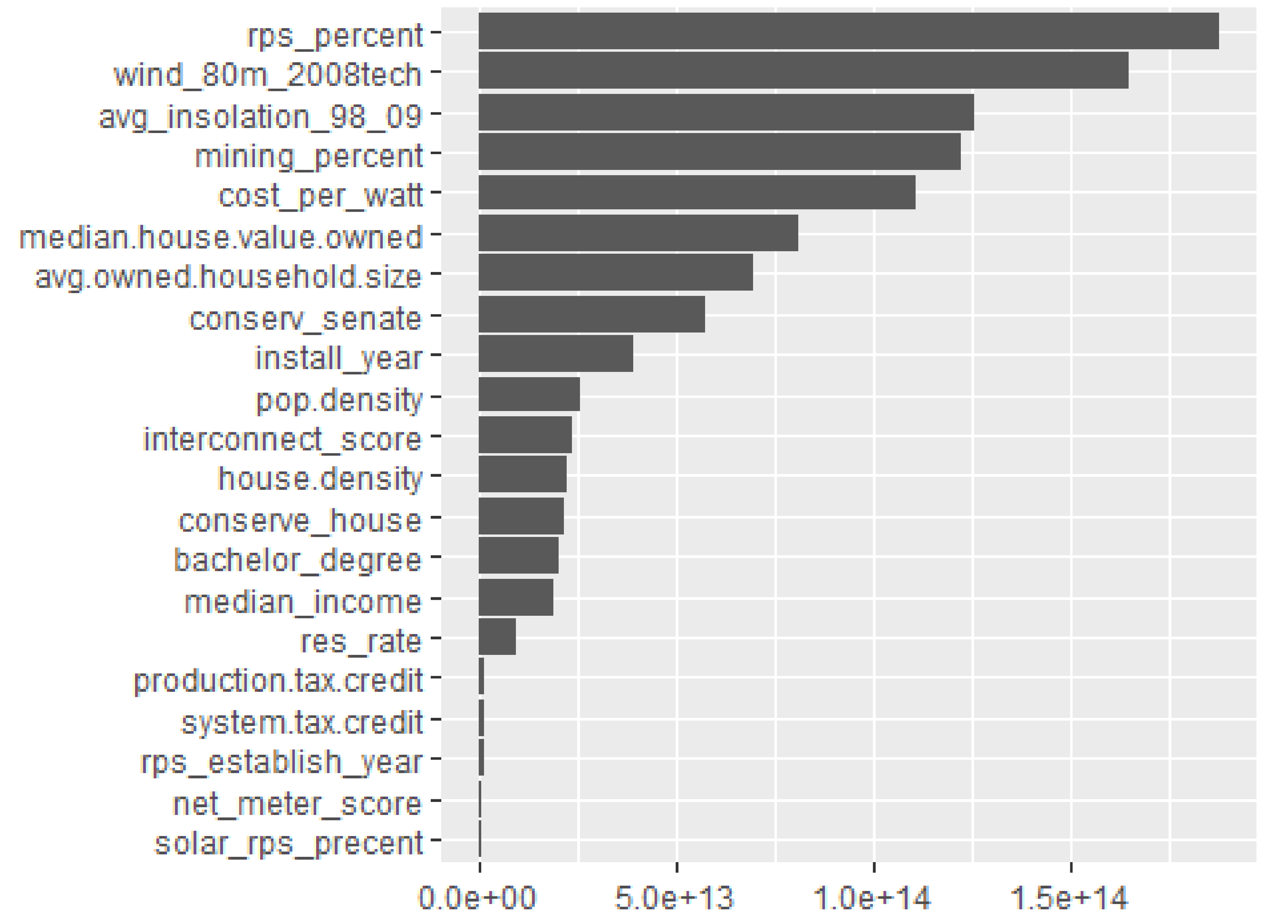
27 subgroups (a.k.a. leaves) are identified that have distinct treatment effects.

The largest ATT is 337.76, while the only negative ATT is -38.82. The ATT for the whole estimation sample is 13.51.

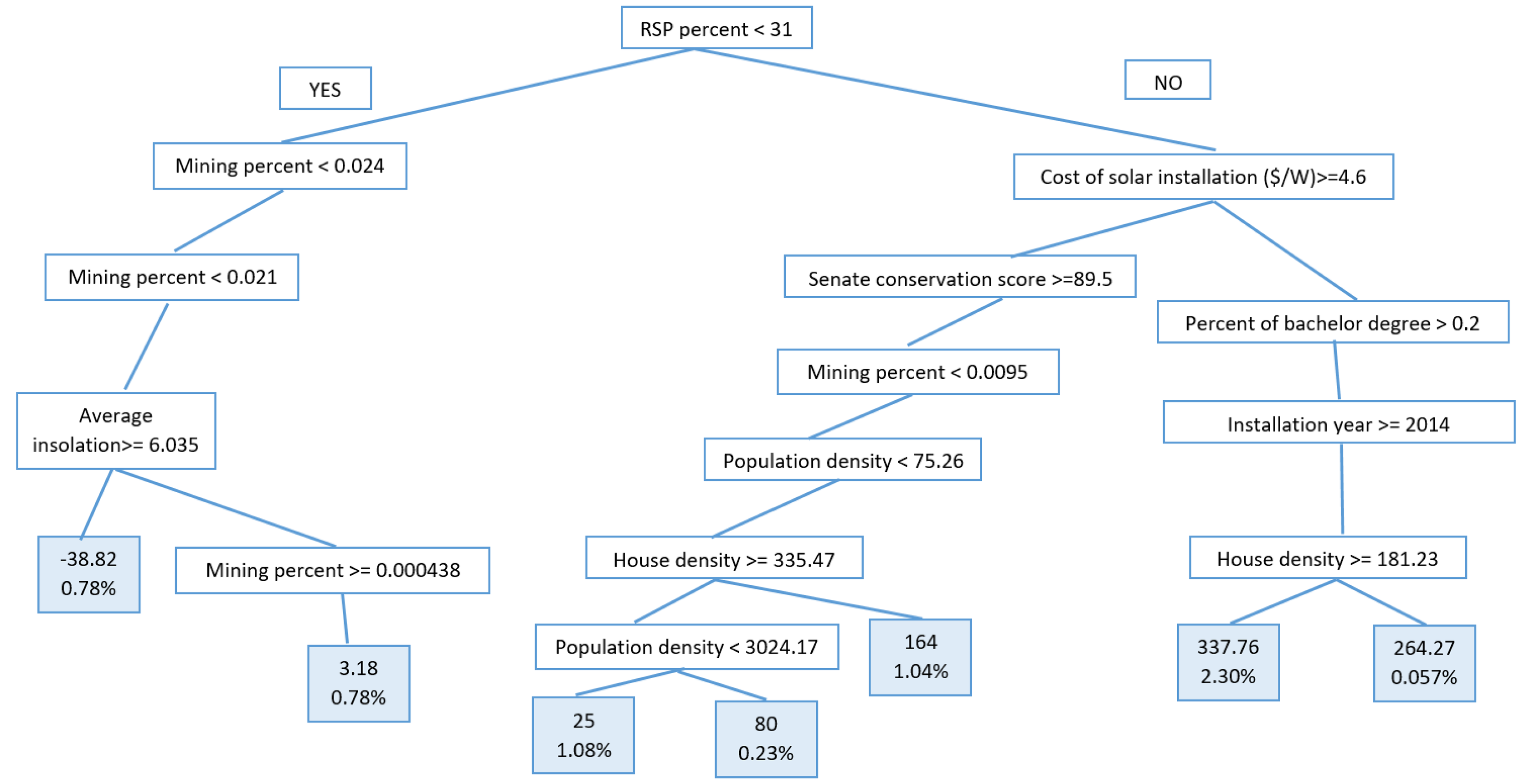
N	estimated ATT	N	estimated ATT
12256	3.18	96	53.18
3805	35.42	87	9.90
1645	5.74	79	193.15
880	24.97	53	103.79
862	5.59	52	80.42
712	20.17	52	337.76
342	22.92	50	122.83
291	59.42	38	292.22
255	64.95	21	170.88
244	25.18	20	76.68
237	164.49	13	246.27
193	93.16	10	233.76
177	-38.82	8	139.43
98	120.46	**whole sample**	**13.51**

Results

Relative influence of variables on predicting ATT



Partial Causal Regression Tree



Conclusions

The percentage of renewable energy mandated by the Renewable Energy Portfolio is the most influential factor determining the effect of solar rebate programs on the number of residential solar installations. Wind, solar and mining resources are also influential factors determining the rebate program impact.

Causal regression trees can identify the subgroups of which the solar rebate programs are most effective at promoting residential solar installations.

Additional steps are needed for statistical inference (i.e., constructing confidence intervals) for the ATT estimates.