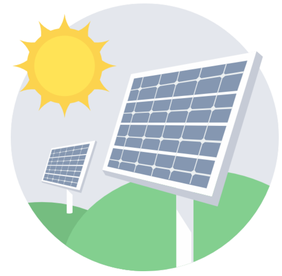
Is the average household income associated with solar panel installation rates in the US?

Dhruv Joshi ‘21, Zhijun Lei ‘20, Roujia Zhong ‘22



Question:
Is the average household income associated with solar panel installation rates in the US?

Primary explanatory variable:
Average household income in census tract units.
Primary response variable:
Solar panel installation rates in census tract units.

Introduction

Solar energy is an inexhaustible resource that can supply a significant portion of global electricity needs (Barbose et al. 2012). The use of solar energy can decrease the greenhouse gas emission and increase national energy independence, work opportunities, and rural electrification rates (Tsoutous et al. 2005). It is necessary to understand whether the household income is associated with solar panel installation rates so that the government can offer new subsidies for household solar panels based on average household income.

Data

- Data of solar panel installation rates in each census tract unit was obtained from the DeepSolar dataset from Stanford University.
- Researchers used machine learning to determine number of solar panels installed in each census tract unit from the satellite images in the 48 contiguous US states (Yu et al. 2018).
- Average annual household income, number of state-given incentives, and solar radiation data for each census tract are also appended into the DeepSolar dataset.
- We used random sampling method from the population to select our sample data points (n = 5000).

Category	n	Proportion (of 5000)
Household Solar Panel Installation within Census Tract		
No panels	1231	0.25
Some panels	3769	0.75
Number of State-given Incentives		
Low (8-17)	1499	0.30
Low-Medium (17-21)	1059	0.21
Medium-High (22-41)	1361	0.27
High (42-75)	1081	0.22

Table 1. Frequency table of categorical variables (Installation within census tracts and incentive counts).

Methods

- The distribution of average household income and solar panel installation rates were skewed to the right. So the data was log-transformed to match the normality assumption.
- Additional variables include number of incentives and daily solar radiation. We categorized the solar panel installation into 2 categories and the number of incentives into 4 categories (Table1).
- A two-sample t-test for the difference in means was used to determine if the household income is higher for census tracts with some solar panel installation (Figure 1).
- A Chi-square test for association was used to see if solar panel installation and number of incentives are associated (Figure 2).
- An ANOVA test was used on the number of solar systems per household for all levels of incentives. The result showed that at least one mean is different, so pairwise comparisons using t tests with pooled SD and Bonferroni adjustment were conducted between all groups (Figure 3).
- A linear regression was conducted to determine the association between the solar panel installation rates and average household income (Figure 4).
- A multivariable linear regression (MLR) was used to determine the influence of the confounder solar radiation on the relationship between two explanatory variables and the response variable (Table 2).

Results

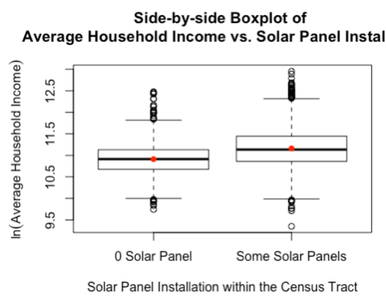


Figure 1 Side-by-side boxplot displaying the distributions of average household income by the solar panel installation within census tract units.

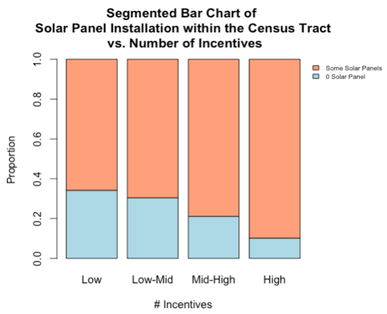


Figure 2 Segmented bar chart displaying the proportion of solar panel installation based on the number of incentives.

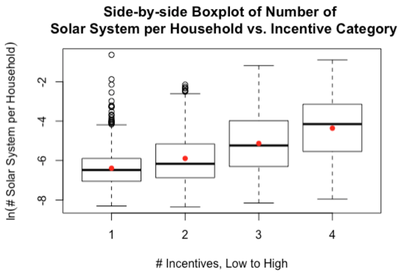


Figure 3 Side-by-side boxplot displaying the distributions of the average number of solar system installed per household by the number of incentives within the census tract (n = 3769).

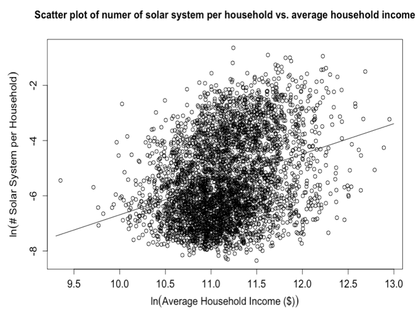


Figure 4. The scatter plot displaying the relationship between the average household income and the number of solar systems per household in census tract units with solar system installations (n = 3769, r = 0.32).
Our model:
ln(solar systems per household) = -17.70 + 1.10 * ln(average household income)

	Coefficient Estimate	Std. Error	t value	Pr(>|t|)
Intercept	13.92001	8.26348	1.685	0.0922
ln(Income)	-2.43592	1.44258	-1.689	0.0914
ln(Income)^2	0.15366	0.06424	2.392	0.0168
Daily solar radiation	-6.43560	0.57793	-11.136	< 2e-16
Daily solar radiation^2	0.84278	0.06471	13.024	< 2e-16
Low-medium	0.23750	0.05654	4.201	2.72e-05
Medium-high	0.72470	0.05301	13.671	< 2e-16
High	0.81286	0.06505	12.496	< 2e-16

Table 2. A multivariable linear regression is used for the primary explanatory variable ln(income), the secondary explanatory variable level of incentives with the low level as the baseline, and the confounding variable daily solar radiation.

- Figure 1 shows that the census tracts with Some Solar Panels form a higher boxplot than the census tracts with 0 Solar Panel, suggesting the average household income is higher in the census tracts with some solar panel installations ($p < 0.0001$).
- Figure 2 shows there is an association between the incentive counts on solar panel installations and solar panel installation in census tract units ($p < 0.0001$).
- Figure 3 shows that the means of the number of solar panels installed per household are pairwisely different for census tracts with different levels of incentives ($p_i < 0.0001$).
- Figure 4 shows that there is a weak positive linear relationship between the average household income and solar panel installation rates. ($p < 0.0001$).
- Table 2 is the output of the MLR. There is strong evidence showing correlations between solar panels installed and all the variables listed except for ln(income). The confounding variable terms (radiation and radiation squared) have a stronger effect among all variables in the model.

Discussion

- **We found that there is a weak positive linear relationship between household solar panel installation rates (log scale) and the average household income (log scale) after controlling for solar radiation.**
- **Solar panel installation rates are also weakly positively associated with state-given incentives on solar panels and strongly positively correlated with daily solar radiation received.**
- **The results indicate that the government can offer monetary subsidies to low-income households to increase the use of solar energy.**

Limitations: The ANOVA and linear regression models only concerned census tracts with solar panel installations because the log-transformation ruled out areas without solar panel installation.

Future research: We would like to examine other variables that might impact the number of solar panel installations, including the average number of years of residents’ education, ratio of votes for Dem to votes for GOP, and electricity price within census tract units.

Recommendations: Since this study utilized a large size of SRS (n=5000), the results of the study are representative of all the census tracts within the 48 contiguous states of the US. Our study indicates the possibility for state governments to incentivize households with lower income to install more solar panels. For example, the government can offer cash rebates and state tax credits and encourage households to sell their carbon offsets to make solar more accessible for today’s homeowners.

Literature cited

Barbose, Galen, et al. “Tracking the Sun III: The Installed Cost of Photovoltaics in the U.S. from 1998-2009.” *Photovoltaics: Local Industry Development, Installed Cost Trends and Mineral Commodities Used*, 2012, pp. 68–120, doi:10.2172/1003825.
Tsoutsos, Theocharis, et al. “Environmental Impacts from the Solar Energy Technologies.” *Energy Policy*, vol. 33, no. 3, Elsevier, Feb. 2005, pp. 289–96, doi:10.1016/S0301-4215(03)00241-6.
Yu, Jiafan, et al. “DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States.” *Joule*, vol. 2, no. 12, Dec. 2018, pp. 2605–17, doi:10.1016/j.joule.2018.11.021