Lease or Buy? Quality differences and asymmetric information in California solar panels

January 4, 2016

Abstract

The market for rooftop solar panel systems are particularly vulnerable to information asymmetry issues of quality. In such a market, poor quality solar panel systems could be expected to push out high quality ones. However, solar systems that are leased can be expected to overcome this asymmetric information problem. Using production data on approximately 1000 solar panel systems in California and a Bayesian hierarchical regression model, I find that leased systems show a significantly higher level of degradation over time - a proxy for quality - than those bought out-right.

1 Introduction

In an earlier paper [Mauritzen] I show that the boom in solar panel installations in California coincided with an adoption of cheaper Chinese panels by contractors. The adoption of cheaper Chinese panels also coincided with a change to a leasing business model by the largest contractors. A straight-forward explanation is that the use of Chinese panels allowed for a leasing model to be financially feasible.

However, a more subtle economic explanation is that a leasing model also facilitated the use of Chinese panels. Chinese panels face an acute asymmetric information problem in that the Chinese manufacturers were nearly completely new to the US and European market. Homeowners and small business owners can not generally be expected to have the expertise to judge the quality of panels on their own. The resulting reliance on brands, reputation, and tests of past performance could create a substantial barrier to entry for panels from new Chinese manufacturers.

Generally, the market for rooftop solar panels can be expected to be particularly vulnerable to issues of asymmetric information on quality. Solar panels can be characterized as an "experience" good, where an investor needs to learn about the quality through use. In particular, poor quality panels will tend to show a degradation of output over time. Even here, solar panel owners may find it difficult to measure the degradation as it can happen gradually, over many years.

More so, solar panel systems are expected to last at least 20 years, thus for all practical purposes, their purchase can be considered a one-shot investment. This eliminates repeat buying as a mechanism for ensuring quality.

In the literature, warranties are often suggested as a strong signal of quality. However, warranties may be a relatively weak assurance of quality in the market for solar panels as both contractors and manufacturers are relatively new, tend to be heavily indebted and have recently shown a tendency to go bankrupt.

Given the inability to judge quality and lack of market mechanisms to signal quality, the established economic theory on the subject would suggest that the market will tend to provide low-quality panels Tirole [1988].

The introduction of a leasing model could potentially get over this information asymmetry problem by shifting ownership to the large contractors that install and finance the solar systems. These contractors can then in turn take steps such as testing panels and visiting manufacturing sights to ensure quality of Chinese suppliers. A testable implication that emerges is then that the quality of panels that are leased - as measured by degradation of output over time - is better in solar panel systems that are leased.

In this paper, I use a data set of California solar power systems that includes monthly production data. I measure the average degradation of production over time as a proxy for the inherent quality of solar panel system. I use a Bayesian hierarchical model to test whether systems that were bought outright by owners displayed, on average, higher degradation over time than those that were leased.

The main finding is that solar panel systems that were leased tend to show significantly less degradation over time than those that were sold outright.

The issue of how information asymmetries about quality affects market outcomes is one of the main topics in the modern industrial organisation literature. Ch. 2 of Tirole [1988] and accompanying citation list provides a good overview of the theoretical foundations of this topic. Of particular importance are Chan and Leland [1982] and Cooper and Ross Cooper and Ross [1984].

Empirical papers exploring issues of quality and asymmetric information are sparse, and to my knowledge, this is the first paper looking directly at the issue of information asymmetries related to solar panel system quality.

However, a growing literature is growing around solar power investment behavior of homeowners and small businesses For example Dastrup et al. [2012] argue that solar panels can not be considered a pure investment good, but are also bundled as a type of green conspicuous consumption - showing evidence for a "solar price premium" in homes with solar panels. Bollinger and Gillingham [2012] study the the role of peer effects in solar power adoption. They find evidence that the adoption of solar panels by homeowners in a certain zip-code will increase the probability that other households in that zip-code will install solar panels.

2 California Solar Initiative, Data and Empirical Model

The data used in this paper is an intersection of two datasets from the California Solar Initiative (CSI). The California Solar Initiative is a state-wide program that gives incentives for installing grid-connected solar panels systems. The incentives are based on performance. For smaller systems, typical of most residential and small business installations, this incentive was given in the form of an upfront payout based on the capacity and the expected performance of the system.

Larger systems were required to accept a performance based incentive based on actual production over 5 years (60 months). From the beginning of the program in 2007 this was defined as those over 100kw, but was lowered to 50kw from January 2008 and 30kw from January 2010. Solar panel installations of all sizes have the option of getting an incentive based on actual performance.

CSI provides data on all grid-connected solar panel systems installed in California since January 2007. In addition, CSI provides another data set of monthly production data from the solar panel systems that received production incentives. The data is openly available on the website of CSI. [1]. A cleaned and merged data set that I use in the following analysis can be found on my website at `jmaurit.github.io#buy_or_lease`.

Below are the key variables present in the CSI installation data:

- Installation date
- Location: address, zip code, coordinates

[1] `http://www.californiasolarstatistics.ca.gov/current_data_files/`

- Reported cost of system
- System capacity
- Name of system owner, host, and contractor
- Panel and inverter manufacturer.
- Third party owner (leased)

The installation data can be matched with the production data for those installations receiving a production subsidy. I include only those installations that have been producing for at least 4 years (48 monthly observations), where the maximum number of observations in this data is 5 years (60 months) of production data. The data then covers similar vintage of solar panels - those that were installed from 2007 through 2009 and which continued to produce through 2013.

I removed systems from the dataset that may have had reporting errors. for example I removed systems where production was reported to be higher than what would be theoretically possible from a solar power system with a given nameplate capacity. I also removed systems that reported 0 production in a period. This could also be an indication of quality if zero production indicates a malfunction in the system. However it is not possible to identify which zero observations reflect malfunctions versus reporting issues or other non-quality issues.

Figure 1 shows the production for the included solar panel systems over time. The seasonality of the systems over the calendar year is clear, and needs to be accounted for in any statistical analysis.

The main methodological goal of this paper is to make an unbiased estimate of the slope of the average degradation of solar panels controlling for seasonality and idiosyncratic factors between solar cells while also comparing the slopes between groups of systems that are leased and bought outright. A simple example of fitting a straight lines through two production series is shown in figure 2 - here as a simple single variable OLS estimate.

The question of interest and the available data suggests a natural hierarchical structure to the empirical model. The individual production data are grouped by the different solar panel systems. Each system is then in turn grouped into categories of leased or host-owned. As mentioned, seasonality must also be adequately accounted for. A natural choice is to use a hierarchical Bayesian

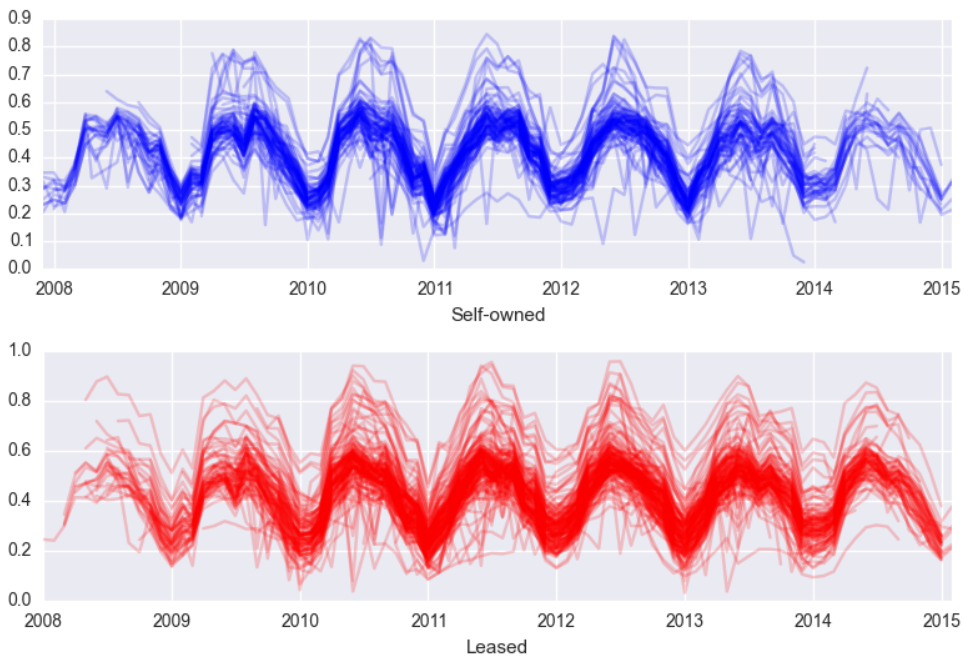

Figure 1: Production over time from California solar panels, Host owned and leased. Production index reflects monthly production in kWh normalized by capacity multiplied by 30.5*12 to reflect a rough estimate of maximum day-light hours in a month.

Figure 2: The main methodological task in this paper is to make an unbiased estimate of the slope of the average degradation of production over time.

model where parameters can be estimated through Markov Chain Monte Carlo (MCMC) simulation techniques.

Figure 3 shows the structure of the model. Starting from the bottom, the individual production data are log transformed and modeled as having a normal distribution with mean y_hat and variance $sigma^2$. $sigma^2$ is in turn modeled as having an uninformative prior. The assumption here is that the variance is constant across the data.

Going up a level, y_hat is modeled as having random group-level intercepts $\beta_{0,j}$, where j represents the j-th of J solar panel system. The $\beta_{0,j}s$ are in turn modeled as being normally distributed with mean μ_{b0} and variance σ^2_{b0}.

In a similar manner J B_j coefficients are modeled as having L mean components $\mu_{\beta 1,l}$ and J random components $\zeta_{\beta 1,j}$ centered around 0. Here L = 2 and corresponds to the average degradation over time for leased systems (l=1) and those sold out-right (l=2). The question of interest - whether leased solar systems display a lower degree of degradation over time than those sold outright - can then be expressed as the distribution of the difference of these parameters: $\mu_{\beta 1,l=1} - \mu_{\beta 1,l=2}$.

Finally, to take into account the seasonality of the monthly data, random effects for each of the 12 month are estimated with a $Normal(1, 10)$ prior on each parameter. The remaining hyper-parameters are given non-informative ("flat") priors.

To estimate the parameters of the model, I use the Stan Bayesian programming language and simulator **??** which uses Hamiltonian MCMC to estimate a joint posterior distribution of the parameters. **??** provides an accessible explanation of the Hamiltonian MCMC algorithm, while more detailed technical descriptions are provided by **??** and **??**.

Bayesian (simulation) methods are still emerging - especially in Economics. So it is worth spending a moment to motivate their use in this case over more common asymptotic methods. As mentioned, the main reason for using the Bayesian model is the flexible treatment it allows for modeling the inherent hierarchy of the data and question of interest.

Bayesian simulation modeling also has some other significant advantages. First, modern simulation methods allow for a wide range of probability distributions in the prior distributions of the parameters and likelihood functions of the data. In turn, the posterior distribution is not constrained to be normal. The fit of this model was adequate with many of the parameters given normal prior distributions, but this can easily be adjusted.

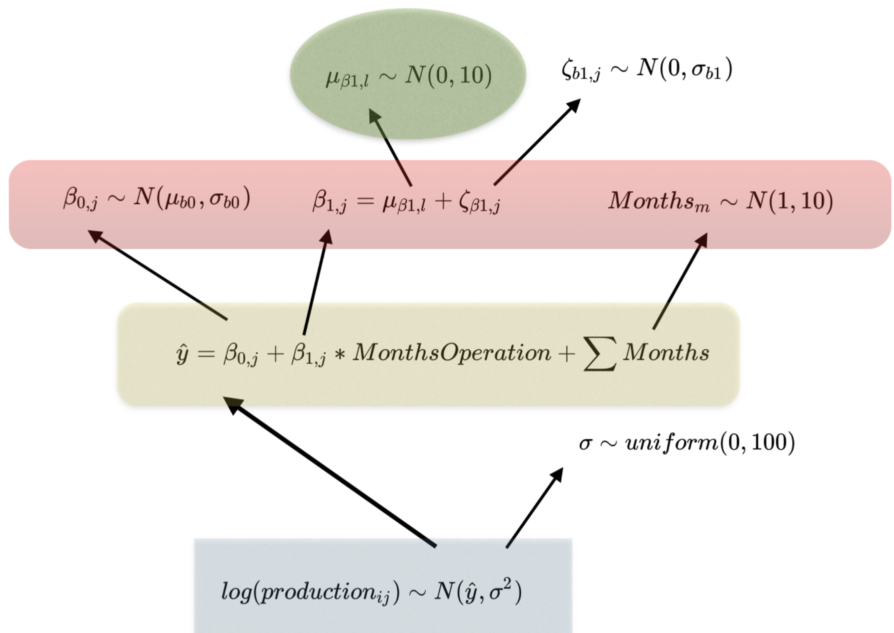

Figure 3: The data and the research question combined suggest a hierarchical Bayesian model. The figure shows the structure of the model. Production data is grouped by individual solar panel systems (index i), and in turn the systems are grouped by categories of leased or sold outright (index l).

Another advantage is that the posterior distributions of the parameters have a natural and direct interpretation as probabilities, as opposed to standard hypothesis testing where the concepts of p-values and significance are often ill-defined. See **??** or **??** for further discussion.

3 Results

I run 1000 iterations with 4 chains of the Hamiltonian MCMC algorithm in order to estimate the joint posterior probability distribution of the model. I can then extract the the estimates of the marginal posterior distributions of the individual parameters in the form of a list of simulated parameter estimates.

Summary statistics for the posterior distributions of the main parameters of the model are shown in table **??** in the appendix. The parameters of interest are $\mu_{b1,l=1}, \mu_{b1,l=2}$ which are the average slope of production over time in, respectively, systems that were sold outright and systems that were leased.

I display the histograms representing the posterior distributions of these two parameters in the top panel of figure 4. The histograms indicate that the posterior distributions appear to have means that are different from each other.

In the lower panel, I show a histogram of the posterior of the difference between the two parameters, $\mu_{b1,l=1} - \mu_{b1,l=2}$ which is the direct test of the question of interest. 95% of the probability lies below the vertical line, which stands at approximately zero. A direct interpretation is then that there is a 95 % probability that on average solar power systems that are sold outright have a higher rate of degradation than those that are leased.

The magnitude of the effect can best be seen by plotting average predicted values over time, which are shown for systems that are leased and sold-outright in figure 5. The solid lines represent the average value of the posterior distribution on the mu_{b1}, l parameters, where the light lines represent random draws from the respective posterior distributions to give an idea of the uncertainty. The model suggests that a leased solar panel system will on average experience degradation of about 1 to 2 percent after 5 years. A system that was sold out right will on average experience degradation of approximately 3-4 percent.

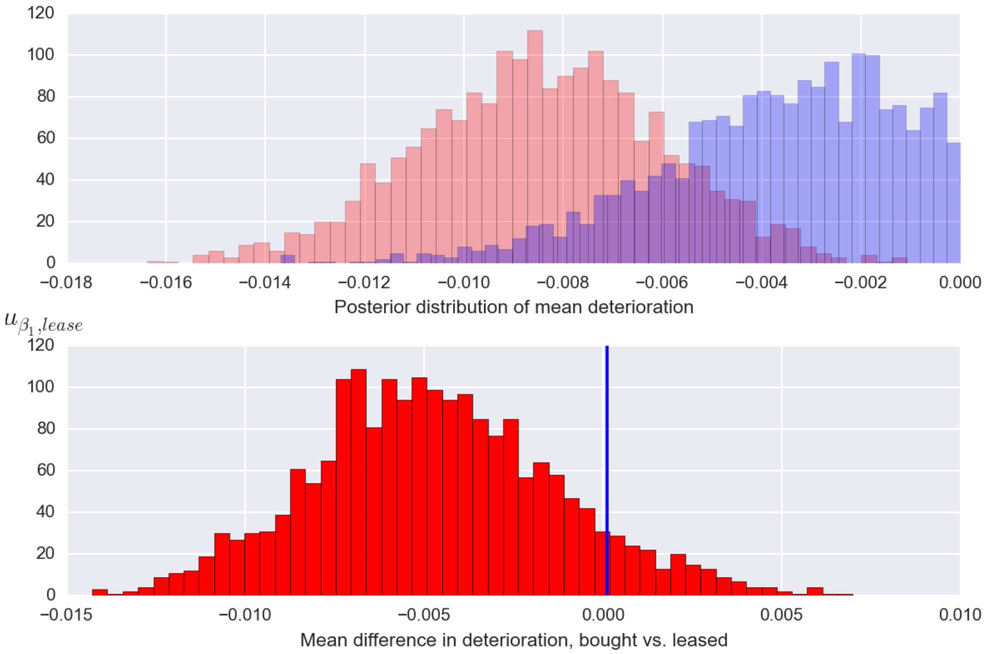

Figure 4:

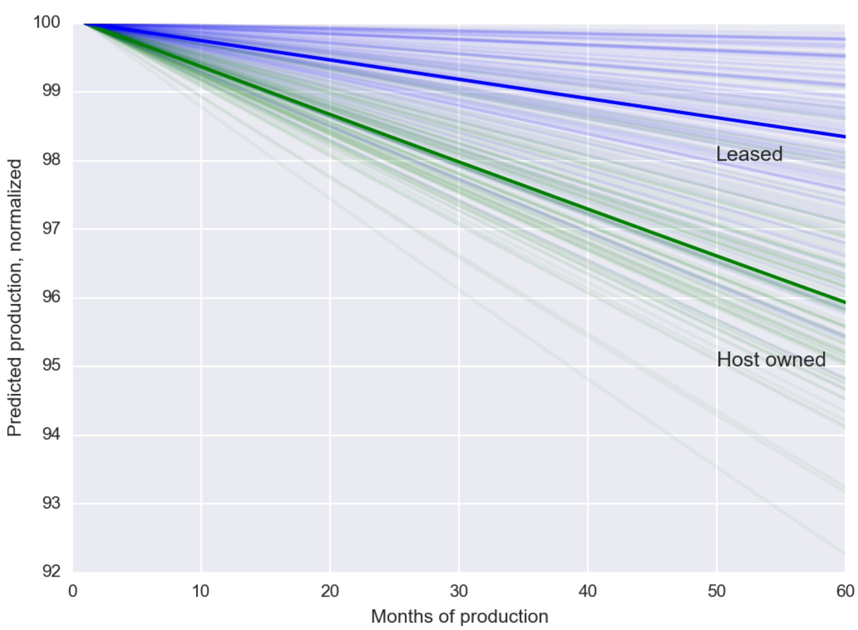

Figure 5: The figure shows the average predicted values of degradation over time of solar panel systems that are leased and sold out-right. The solid lines represent the mean value of the posterior where the lighter lines represent random draws from the posterior distribution

4 Discussion and Conclusion

The structure of the emerging solar industry in the US, as well as other parts of the world suggests that issues of information asymmetry may play an important role. The established theory on the subject suggests that under information asymmetry of quality, poor quality products will push out good quality products. In this article I have tested an implication of that theory for the case of solar panels in California. The results are consistent with the presence of significant information asymmetry in the market.

The theoretical literature on information asymmetry and quality is established and deep. The empirical literature testing the implications in actual markets is, on the other hand, sparse. This article provides one of few examples of both a market that would be expected to have issues with information asymmetry of quality, and shown direct evidence of the presence of information asymmetry.

Beyond providing a case study and confirmation of well-known economic theory, the article also has policy implications for the emerging solar power industry. Taking the political goals of installing more renewable energy as a given, this article speaks to the strength of making policy instruments flexible in terms of who actually owns the generating investment. Subsidies in other countries - like Germany, and US states - like North Carolina, have required the host of the solar panel system to also own the system in order to receive the subsidy.

This article should also be considered as part of a broader literature on the special characteristics of new distributed energy generation technologies. The investment behavior of home owners, farmers, and small cooperatives are bound to be substantially different than those of large, sophisticated energy companies that have traditionally done most of the investment in electricity generation.

References

Bryan Bollinger and Kenneth Gillingham. Peer Effects in the Diffusion of Solar Photovoltaic Panels. *Marketing Science*, 31(6):900–912, September 2012. ISSN 0732-2399. doi: 10.1287/mksc.1120.0727. URL `http://pubsonline.informs.org/doi/abs/10.1287/mksc.1120.0727`.

Yuk-Shee Chan and Hayne Leland. Prices and Qualities in Markets with Costly Information. *The Review of Economic Studies*, 49(4):499–516, October 1982. ISSN 0034-6527, 1467-937X. doi: 10.2307/2297283. URL `http://restud.oxfordjournals.org/content/49/4/499`.

Russell Cooper and Thomas W. Ross. Prices, Product Qualities and Asymmetric Information: The Competitive Case. *The Review of Economic Studies*, 51(2):197–207, April 1984. ISSN 0034-6527. doi: 10.2307/2297687. URL `http://www.jstor.org/stable/2297687`.

Samuel R. Dastrup, Joshua Graff Zivin, Dora L. Costa, and Matthew E. Kahn. Understanding the Solar Home price premium: Electricity generation and Green social status. *European Economic Review*, 56(5):961–973, July 2012. ISSN 0014-2921. doi: 10.1016/j.euroecorev.2012.02.006. URL `http://www.sciencedirect.com/science/article/pii/S0014292112000244`.

Johannes Mauritzen. What's behind the surge in the California rooftop solar market? The role of Chinese panels, uncertainty and new business models. *Working Paper*. URL `http://jmaurit.github.io/#solar_lemons`.

Jean Tirole. *The Theory of Industrial Organization*. The MIT Press, Cambridge, Mass, 1st edition edition, 1988. ISBN 9780262200714.

A Appendix: Tables

params	high_ci	low_ci	means	medians
mu_b1_own	-0.000393	-0.001244	-0.000802	-0.000798
mu_b1_lease	-0.000015	-0.001000	-0.000310	-0.000243
sigma 0.214203	0.207786	0.210951	0.210925	
jan	7.113382	6.198759	6.792520	6.842171
feb	7.298859	6.388751	6.981517	7.029715
mar	7.479400	6.557983	7.156756	7.205567
apr	7.765322	6.853390	7.445405	7.494745
may	7.914406	6.995134	7.592065	7.642249
jun	7.984840	7.068018	7.665142	7.714351
jul	7.994245	7.076688	7.671779	7.721463
aug	7.989234	7.079094	7.669284	7.717939
sep	7.919979	7.011260	7.600108	7.646793
oct	7.732073	6.816785	7.412012	7.460158
nov	7.550459	6.635681	7.229874	7.277312
dec	7.270718	6.360748	6.954599	7.002773