Weather and the Decision to Go Solar: Evidence on Costly Cancellations

Yanjun Liao*

May 2019

Abstract

This paper studies the effect of short-run weather fluctuations on solar panel adoption in California. This decision appears to respond strongly to weather patterns associated with solar panel productivity: I find that customers whose sign-up for solar panels is followed by unfavorable weather are more likely to cancel their contracts. In contrast, non-residential customers are not subject to the same effect. Together, these results suggest that short-run weather conditions affect customers' valuation of solar panels. The most plausible mechanisms are psychological biases such as projection bias or a salience effect, leading the decision-maker to rely too heavily on transient conditions when predicting long-run utility. This paper is among the earliest to document evidence of behavioral anomalies in the solar market.

JEL Classification: D03, D12, D81, Q42, Q48.

Keywords: projection bias, salience, solar adoption, weather.

*University of California, San Diego. Email: yjpenny@gmail.com. I am especially grateful to Mark Jacobsen, Joshua Graff Zivin, Julie Cullen, and Charles Sprenger for their invaluable advice and support. This paper has also benefited from comments and suggestions by anonymous referees, Richard Carson, Tom Chang, Kenneth Gillingham, Gordon McCord, Sally Sadoff, David Victor, and seminar participants at UCSD and ASSA. I am also grateful to the California Solar Initiative (CSI) for making their data publicly available. All errors are my own.

1 Introduction

Many important economic decisions require the consumer to estimate future utility. When purchasing a durable good, the consumer needs to trade off her future utility against the present cost of purchase. Standard economic theory assumes that these estimates are accurately and consistently made, but emerging evidence shows that psychological mechanisms can often interfere.[1] For example, people with projection bias tend to exaggerate the degree to which future state will resemble the current state. People may also be affected by the salience effect, where current conditions accentuate certain attributes of a good and lead to a change in valuation.[2]

In this paper, I study whether residential solar customers are affected by short-run weather in their decision to acquire a solar photovoltaic (PV) system. If the current weather is favorable for real-time solar power generation, we might see more sign-ups because the benefit of a solar PV system is more salient, while bad weather would have the opposite effect. Moreover, because solar panels are not immediately installed after sign-up, weather conditions during the waiting period may also affect the customer's valuation of the system and lead to cancellation. These implications, with a focus on cancellations, are empirically tested in this paper.

Using transaction-level administrative data from the California Solar Initiative (CSI), a statewide subsidy program, I find that residential customers are more likely to cancel their solar contracts when they experience worse weather after signing up. Specifically, a one-standard-deviation decrease in solar radiation is associated with a 7.1-9.8% increase in cancellations, which, in turn, account for 12.06% of all CSI applications. This effect is identified using variation within each season while controlling for individual system characteristics and monthly economic conditions. It is consistent with the aforementioned psychological mechanisms and robust across a variety of other specifications, including county-quarter-year fixed effects, month fixed effects, nonlinear models, and alternative weather controls.[3] Moreover, I find that non-residential customers are not subject to the same effect. This is likely because their decision-making process is less susceptible to psychological factors in

[1] See DellaVigna (2009) for a review.

[2] Empirical evidence on projection bias: Conlin et al. (2007), Simonsohn (2010), Busse et al. (2015), and Magistris and Gracia (2016). Empirical evidence on salience effects: Kőszegi and Szeidl (2012), Bordalo et al. (2013), and Hastings and Shapiro (2013).

[3] By controlling for a rich set of fixed effects, I use the remaining random variations in weather to identify the causal effect on cancellations. This fixed effects framework is first used in Deschenes and Greenstone (2007) and later adopted in numerous papers in the climate-economy literature. See Dell et al. (2014) for a review.

general.

Solar PV investment has some unique features that make it difficult to reconcile the above results with rational mechanisms. First, solar panels typically last more than 25 years. Short-run weather fluctuations prior to installation do not affect the return to such a long-run investment. Second, solar companies commonly provide estimates of electricity savings based on micro-climate to potential buyers. Short-run weather fluctuations do not contain useful information in addition to those estimates. For these reasons, the results also cannot be explained by present bias unless the discount rate is very large and the weather is very persistent.

As an outcome of interest, the *cancellation* of a signed contract has a number of advantages. In terms of interpretation, a cancellation directly reflects that the customer decides ex post that the contract was a mistake. Projection bias can thus be directly tested using this decision margin, as discussed in the literature (Conlin et al. 2007; Busse et al. 2015). Furthermore, the cancellation often entails a penalty which supports the interpretation of a change in valuation.[4] An obvious alternative is *purchase*: consumers with projection bias are more likely to purchase solar panels on sunnier days. It is tempting, therefore, to estimate whether there are more purchases on sunny days.[5] However, the implications of such a pattern of purchase are less clear than cancellation. For instance, a solar purchase might simply be due to increased awareness of solar power rather than higher valuation by the customer. Also, a sunny day might "harvest" solar purchases that would have occurred rationally, which implies minimal welfare impacts. In contrast, cancellations necessarily carry losses in overall welfare in the form of wasted effort by both the contractor and the customer in coordination, negotiation, site inspection, and other contracting costs. Another key advantage of cancellation over purchase is data availability. Cancellation can be analyzed at the individual level, whereas purchase can only be examined at an aggregate level due to lack of information on all potential buyers.

I examine how current weather can impact a customer's decision by explicitly modeling her over-dependence on current conditions along two dimensions. The model allows weather to affect the customer's valuation by (1) a "solar production" channel that affects her prediction of future solar generation, and (2) an "increasing rate" channel where solar generation displaces electricity on different price tiers. Applying engineering models, I transform the weather variables into two indices that represent the two channels. I then estimate the ef-

[4] Section 3 provides more details on the cancellation process and the associated costs.

[5] A concurrent paper (Lamp 2018) examines this outcome in Germany. As supplemental evidence, I also look at solar sign-ups and find results that are similar in direction and scale.

fects of these indices jointly. The results show that cancellation responds most strongly to variations in short-run solar panel productivity. I also show evidence that negative updates in weather conditions relative to the pre-contract period play an important role, and the responses are particularly strong for customers who sign up following a period of above-average weather. These results lend further support to the psychological mechanisms.

Lastly, I explore heterogeneous responses along various dimensions. I first examine whether customers from neighborhoods with higher solar-market penetration respond differently to weather. I find that they are less likely to cancel, which is consistent with the peer effects identified in Bollinger and Gillingham (2012). However, they are not less responsive to weather. There are heterogeneous effects on other margins: the weather effect is larger for customers whose system is owned by a third party, whose neighborhood of residence is more urban, and whose sign-up was followed by an imminent decrease in the incentive rate.

This paper makes several contributions. It is closely related to a growing literature in behavioral economics that presents "smoking gun" evidence of projection bias and salience effects in the field (Conlin et al. 2007; Simonsohn 2010; Hastings and Shapiro 2013; Busse et al. 2015; Chang et al. 2018; Lamp 2018). Rather than studying the effect of weather on the purchase decision as most previous studies do, this paper highlights the importance of weather updates, especially negative ones, in changing the customers' valuation. Furthermore, because a solar PV system is a big-ticket item with an instrumental nature, the findings suggest that such biases are more pervasive and economically important than previously understood.

This paper also adds to the literature that documents the behavioral anomalies in energy efficiency investments. Consumers are often found to mis-optimize due to inattention, imperfect information, or myopia (Hausman 1979; Attari et al. 2010; Allcott 2011, 2013; Busse et al. 2013; Ito 2014; Sallee 2014). The decision to invest in solar PV is, in many ways, similar to the decision for energy efficiency. Therefore, the behavioral mechanism identified in this paper is also relevant for any energy efficiency investment whose real-time productivity depends on transient conditions.

Such behavioral anomalies have profound policy implications. They provide the basis for paternalistic intervention in energy efficiency (Allcott et al. 2014; Allcott 2016). They can also be targeted in new behavioral programs to achieve much higher cost-effectiveness than traditional economic incentives (Allcott and Mullainathan 2010; Gillingham et al. 2018). These lessons also apply to solar policies. In the U.S., there are many solar subsidies at the federal, state, and local levels. However, prior evaluations show that these programs often cost more than the environmental and industrial benefits they generate.[6] It will be fruitful for future policy to incorporate insights from the behavioral sciences. As behavioral research on consumer choice regarding solar is limited, the current study is an early effort in that direction. The findings reveal opportunities to improve welfare by presenting potential customers with accurate information and cautioning them against over-dependence on short-run weather conditions, thereby reducing costly cancellations.

2 Weather in the Decision to Go Solar

This section presents a simple model to illustrate how specific weather patterns might affect customers' valuation of solar panels. It is adapted from existing theory on projection bias, but the implications are applicable to other psychological channels where consumers are over-influenced by current conditions.

2.1 An Illustrative Model on Projection Bias

Projection bias refers to the tendency for people to exaggerate the degree to which the future state will resemble the current state. As modeled in Loewenstein et al. (2003), this framework distinguishes between $u(G, s_1)$, the actual utility from using good G in period one (with state s_1), and $\tilde{u}(G, s_1|s_0)$, the "projected" utility based on the current period (with state s_0). A simple representation of over-influence by the current state is

$$\tilde{u}(G, s_1|s_0) = (1-\alpha)u(G, s_1) + \alpha u(G, s_0), \tag{1}$$

where α is between 0 and 1. When α is nonzero, the projected utility is different from the actual utility and biased towards the period-zero utility. At the current period, the projected total utility from purchasing a good that lasts from period 1 through T is given by

$$\begin{aligned} \tilde{U}(G|s_0) &\equiv \sum_{\tau=1}^{T} \delta^{\tau} \tilde{u}(G, s_{\tau}|s_0) \\ &= (1-\alpha)U(G) + \alpha\delta\Big(\frac{1-\delta^T}{1-\delta}\Big)u(G, s_0), \end{aligned} \tag{2}$$

[6] Hughes and Podolefsky (2015) and Crago and Chernyakhovskiy (2017) both find that the cost of carbon mitigation in subsidy programs is much higher than the social cost of carbon. Bollinger and Gillingham (2014) find that non-appropriable learning-by-doing in the solar industry similarly do not justify the cost. Furthermore, Borenstein and Davis (2015) notes the distribution of benefits of such programs are highly regressive. More generally, see Baker et al. (2013) for a review on the cost and benefit of solar electricity.

where $U(G) = \sum_{\tau=1}^{T} \delta^\tau u(G, s_\tau)$ is the discounted sum of the stream of actual utilities. The second equality follows from substituting in equation (1). Therefore, the current state biases the total projected utility.

In the context of solar adoption, I consider the relevant states as different weather conditions. Households decide whether to go solar by weighing the projected total utility against a fixed cost. Therefore, although current weather does not affect the true utility ($U(G)$) at all, it affects the projected utility and hence the decision of the marginal customer. This has two testable implications:

(I1) Better weather leads to a higher probability of signing up for solar.

(I2) Worse weather after the sign-up leads to a higher probability of canceling the contract.

While we can test both implications, there are conceptual and practical differences between them. Section 3.2 provides a detailed discussion on these issues.

2.2 Channels: Solar Production vs. Increasing Rate

The return to a solar PV system in each period depends on the total electricity bill savings in that period, which depends on the amount of power it generates and the cost of the displaced electricity.[7] I assume that the current period utility equals to the reduction in electricity bill under net metering, which takes the following form:

$$u(G, s_0) = \int_{D(s_0)-E(s_0)}^{D(s_0)} P(v)dv. \tag{3}$$

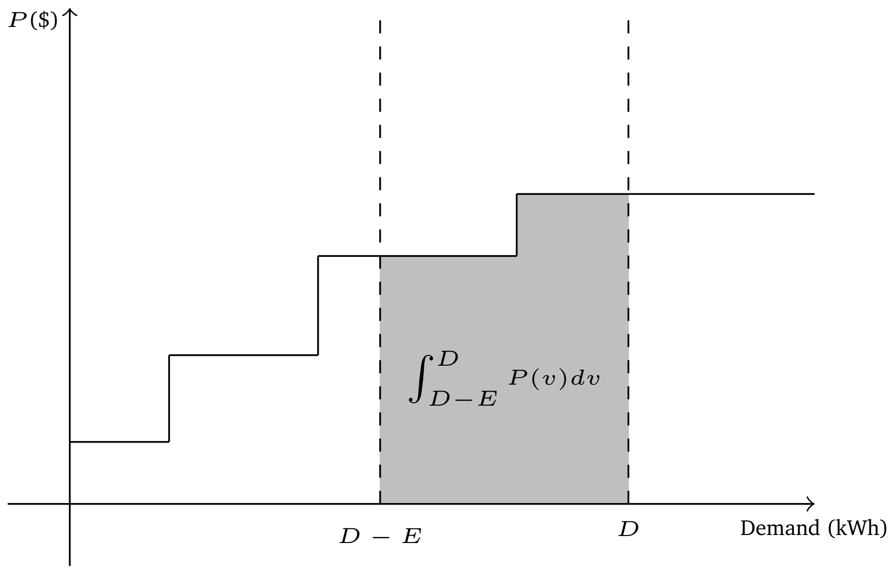
$D(s_0)$ is gross electricity demand, $E(s_0)$ is solar power production, and $D(s_0) - E(s_0)$ is the net demand that eventually enters into the price schedule. $P(\cdot)$ is the marginal price schedule of electricity and is weakly increasing in electricity usage due to the increasing-block pricing scheme in California.[8] As illustrated in figure 1, equation (3) is essentially measuring the shaded area for given electricity demand D and solar power production E.

For simplicity, assume the scalar s_0 is a sufficient statistic that summarizes the state of the world in period zero. It is clear that s_0 affects the projected total utility through both

[7] Some solar customers might have "green" preference. The model accounts for such preference so long as "green" utility is proportional to solar power generation. Other forms of green preference are ignored for simplicity.

[8] See Borenstein (2012) for details on electricity pricing in California. This rate structure is a major option for electricity pricing. It is also adopted in other important solar markets such as China.

Figure 1: Benefit from Solar PV System



$D(s_0)$ and $E(s_0)$. More formally, we can derive

$$\frac{d\tilde{U}(G|s_0)}{ds_0} = \alpha\delta\Big(\frac{1-\delta^T}{1-\delta}\Big)\Big[\underbrace{P|_{D(s_0)-E(s_0)} \cdot E'(s_0)}_{\textit{Solar Production Channel}} + \underbrace{\big(P|_{D(s_0)} - P|_{D(s_0)-E(s_0)}\big) \cdot D'(s_0)}_{\textit{Increasing Rate Channel}}\Big]. \quad (4)$$

Equation (4) highlights the two channels through which weather might affect the decision to go solar. The first term represents the solar production channel. When the current weather is unfavorable for solar power production, customers might underestimate the lifetime productivity of the solar PV system. The second term represents the increasing rate channel. During periods of greater energy demand, such as particularly hot or cold days, customers might overestimate their future energy demand. In turn, they will perceive greater savings from the solar panels under an increasing rate schedule.

Theoretically, the production channel will shut down when the solar PV system generates sufficient power to offset all electricity consumption in a year.[9] On the other hand, the increasing rate channel will shut down when solar generation is not large enough to move the net demand to a lower price tier. In reality, however, these conditions are rare. It is worth noting that these two channels do not necessarily go in the same direction. For example, extreme heat reduces solar panel productivity but increases electricity demand for cooling. Another real-life complication is that customers have been shown in this context to respond

[9] The billing cycle for California net metering customers is 12 months.

to average rather than marginal price (Ito 2014). We can account for this by replacing $P(\cdot)$ with a perceived price schedule that is constantly increasing. Under the new schedule, the qualitative result in equation (4) remains valid.

3 Policy Background and Data

3.1 The California Solar Initiative (CSI)

The CSI is a solar rebate program in California for the customers of the three investor-owned utilities (IOUs) covering most of California. It funds solar on commercial, agricultural, government, and non-profit buildings as well as existing homes. Overseen by the California Public Utilities Commission (CPUC), the program had a total budget of $2.167 billion between 2007 and 2016 and a goal to install approximately 1,940 MW of new solar generation capacity. Over the years, California continues to be the leading solar market in the United States, accounting for about half of the nationwide installed capacity.

An important feature of the program is that the rebate rate is designed to decrease automatically based on the total volume of solar megawatts (MWs) with confirmed project reservations.[10] The exact timing of each step decrease is uncertain to solar market participants. This design is intended to encourage learning-by-doing and lower the cost of solar over time. During the sample period, the average cost has indeed been rapidly decreasing.

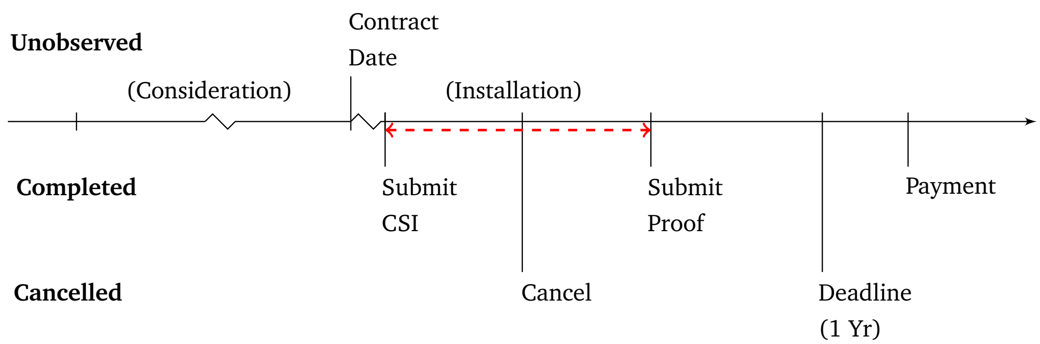
The main data source in this paper is the administrative dataset from the CSI. It contains the universe of CSI applications, regardless of whether they are eventually completed. They represent about 80% of all solar projects in California before 2012 and much less afterward, as the incentive weakens. For each application, the dataset provides detailed information on system characteristics such as the CSI rating (generation capacity), total cost, incentive amount, the sectors for the host and the owner, and status of third-party ownership.[11] More importantly, it contains the zip code and relevant dates for each application, which can be matched to meteorological data at a fine geographical scale.

Figure 2 shows the typical timeline for a CSI application, where unobserved events are plotted above the line and observed ones below. The process begins with the household spending an unknown amount of time to consider going solar. In this process, they might

[10] For the incentive schedule, see figure C1 in the online appendix.

[11] The sectors include residential, commercial, government or non-profit. The host is usually the customer, or where the system is physically located. Third-party ownership (TPO) means that the system is not owned by the host but a third party such as a solar developer or investor.

Figure 2: Timeline of a CSI application



be in contact with one or more solar contractors. Each contractor provides some relevant information such as estimated bill savings, design options, and quotes. The consideration period ends when the household signs a contract. This sign-up date is not observed. However, CSI records the date when a rebate application is electronically submitted, at which point a signed contract is required by program rules. The application might be submitted by either the homeowner or the contractor, but the subsidy is mainly captured by the homeowner (Pless and Benthem, Forthcoming). Because a rebate is reserved for the homeowner at the current rate upon submission, the homeowner has the incentive to submit as soon as possible. This makes the submission date an effective proxy for the sign-up date.[12]

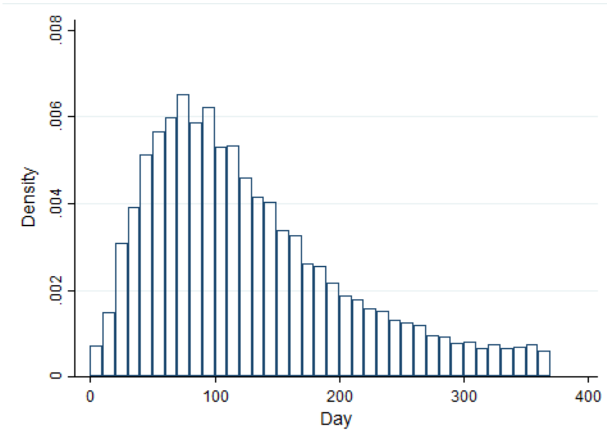
The majority of the applicants completed the installation and submitted the proof within one year. For these projects, I calculate the complete duration as the number of days between the initial application and the proof submission. On the other hand, 12.06% of the applications ended up being canceled.[13] Likewise, a cancellation duration can be calculated as the period between the initial application and the recorded cancellation date.

In figure 3, I plot separate distributions for complete duration and cancellation duration. Because applications are subject to a one-year deadline, observations beyond one year are excluded. The distribution of complete duration is single-peaked, relatively dispersed and smooth with a 96.2% valid rate, which is what we would expect if most dates are reported accurately. On the other hand, the distribution of valid cancellation duration has a highly concentrated peak around a month, but it accounts for only 36.4% of all cancellations. In

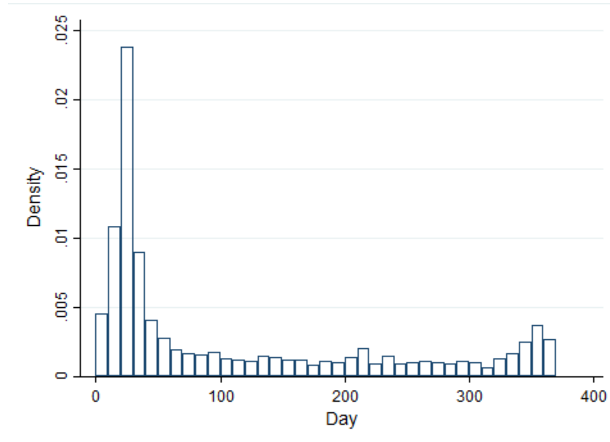
[12] This is also confirmed by the author's communication with a CSI staffer.

[13] According to the author's interview with a major solar company, the cancellation rate observed in this dataset is lower than their own records. The primary cancellation reason for their customers is that further assessment lowers estimated capacity due to changes in parameters such as shading or azimuth.

Figure 3: Distribution of Complete and Cancellation Duration



(a) Complete duration (within 1 year)



(b) Cancellation duration (within 1 year)

Note. Panel (a) shows the distribution of completion duration, and panel (b) shows cancellation duration. Only observations before the 1-year installation deadline are included, which account for 96.2% of the 136,002 completed projects and 36.4% of the 18,653 canceled ones.

other words, the majority of canceled projects have a recorded cancellation date beyond one year.

When a project is canceled, there is no incentive for either the solar company or the customer to report to the CSI administrator on time. Thus, many cancellations might have gone unreported until the deadline, when the administrative process requires the project status. Indeed, when we look at the uncensored distribution of recorded cancellation duration (see figure B3 in the online appendix), there is strong evidence of bunching right after the deadline. Therefore, these dates are not likely to reflect true cancellation time frame accurately. In section B of the online appendix, I investigate this reporting issue in more details and find that this is a pervasive issue across program administrators and contractors.

3.2 Purchase and Cancellation

The theory above generates testable implications of weather fluctuations on purchase (I1) and cancellation (I2). While both implications will be empirically tested, the focus will be on cancellation for a number of reasons.

First, cancellation can be analyzed at the individual level and its interpretation is clear. As first proposed by Conlin et al. (2007) and later discussed by Busse et al. (2015), projection bias can be directly tested by finding evidence of consumers deciding *ex post* that a decision was a mistake. In contrast, the sign-up decision can only be examined in the aggregate. This is due to a data limitation: we only observe realized transactions but not the pool of potential buyers. As a result, instead of examining the binary purchase outcome, we test whether there are more purchases in a zip code when conditions become favorable. This is similar to prior research on health insurance take-up (Chang et al. 2018). However, it is difficult to interpret the effect as a change in customer valuation of solar panels, as this interpretation assumes that the pool of potential buyers does not change based on weather conditions. This assumption might fail if, for example, a sunny period increases the pool of potential buyers by reminding them of solar power. The test of (I1) does not rule out this possibility.

Second, cancellation reflects larger economic consequences than purchase. The California state law specifies a cool-down period of three business days for the customers to freely cancel their contracts. Outside of that period, a cancellation penalty can be charged legally, depending on the solar contractor and the progress of the installation.[14] Moreover, cancellation also means wasted time and effort by both the contractor and the customer. In contrast, a positive correlation between purchase and sunshine might simply reflect short-run postponement of sign-up with minimal welfare implications.

Because many of the recorded cancellation dates are invalid, we need to make alternative assumptions about the relevant time frame of the cancellation decision. I use two separate assumptions in the main analysis. First, I assume that the cancellation decision is made within the same time frame as installation. Thus, I use the median complete duration (113 days after the sign-up) as the relevant window for every project. Alternatively, the reported cancellation dates are more likely to be accurate if the duration is of a reasonable length (for example, less than 180 days). For this subset of observations, I assume the relevant time frame is the 30 or 60 days before the cancellation or completion date.

While both are imperfect proxies of the true decision time frame, the two assumptions complement each other in their strengths and weaknesses. In the first assumption, the 113-day period is consistent with the typical installation time frame specified in solar contracts. However, it is too long for some projects and too short for others. This introduces measurement errors in the relevant weather variables and attenuates the estimates. In the analysis, I test the robustness to alternative period lengths and whether the results are further at-

[14] Anecdotes suggest that some companies impose no penalty if the work has not started, while others are notorious for their strict cancellation policy. The penalty could include the deposit, which is typically 500-1000 USD, a restocking fee if equipments have been ordered, and potentially more if the installation is in later stages. Some customers sued their contractors in order to exit from a contract, which incurs additional legal fees.

tenuated as we move further away from the preferred length. The second assumption might capture the true decision time frame better. There are reasons to believe that more recent experiences of weather are more important for the decision, such as an increased urgency of the decision and recency bias. However, this assumption excludes the majority of cancellations in the data while the first allows us to examine all applications. Thus, the second assumption delivers stronger evidence while the first provides a more complete picture.

Besides cancellation, I also examine the purchase decision. In particular, I am concerned that favorable weather might lead some customers to overestimate the profitability of solar and sign contracts that will eventually get canceled. In that case, mean-reverting weather trends might introduce mechanical correlations between cancellation and post-contract weather. While this phenomenon is still consistent with projection bias, the main estimates on cancellation and the welfare consequences will be harder to interpret. In the empirical analysis, I address this issue by examining whether favorable weather is associated with more sign-ups or larger systems. I also explore how the dynamics of pre- and post-period weather jointly affect cancellations.

3.3 Solar and Weather Data

The weather data in this paper comes from two main sources. I obtain station-level data on daily maximum temperature and wind speed from the Global Surface Summary of the Day (GSOD) dataset at NOAA. I exclude weather stations with elevation above 1,500 meters, which leaves 191 stations operating in the sample period to be matched with 1,428 zip codes. Based on the center coordinates of each zip code, I find all the weather stations within 100 kilometers (62 miles) and calculate the mean of observations from these stations, weighted by inverse distance. These measures might not be precise for each application. This problem is less severe for temperature – which is more smoothly distributed across space – but more severe for wind speed, which can be highly localized.

Surface solar radiation is a key weather element for solar power generation. I obtain daily observations of solar insolation from the Prediction of Worldwide Energy Resource (POWER) Project at NASA.[15] This data is derived using satellite observations and provided at a 1° latitude by 1° longitude resolution. For each zip code, I extract data using its center coordinates. As a zip code is much smaller than a grid cell, it is likely to lie entirely in the

[15] Solar insolation is the amount of electromagnetic energy (solar radiation) incident on the surface of the earth, measured in MJ/m^2 per day. See Stackhouse Jr et al. (2015) for methodology and accuracy of this measure.

same grid cell as its center coordinates.

Finally, I convert daily observations of weather into summaries over the relevant period for each application: average solar insolation, average wind speed, number of days with maximum temperature below 40°F and the same for above 100°F.

3.4 Summary Indices for Solar Production and Energy Demand

The weather affects both energy demand and solar production, and both are related to the value of a solar PV system. To understand which factor plays a bigger role, I construct a proxy for each of the two factors and analyze their effects on cancellation simultaneously. These measures are calculated using data for solar insolation, wind speed, and temperature. As regressors, they allow for easier economic interpretations compared to the original weather elements.

Increasing rate (energy demand) channel. Weather affects energy demand mainly through heating (cooling) in cold (hot) days. A common practice is to calculate heating and cooling degree days (HDD and CDD):

$$HDD = 1\{Temp < 65\} \cdot (65 - Temp) \quad \text{and} \quad CDD = 1\{Temp > 65\} \cdot (Temp - 65).$$

Heat is a major driver of electricity use in California, as air conditioners and fans all run on electricity. On the other hand, cold weather might be less relevant for solar in California because of its mild winters. Furthermore, electricity is not the primary choice for heating in Californian homes, accounting for only 8.5% of the heating demand (EIA 2016). As a result, the time series of peak demand in the California ISO system is characterized by large spikes in the summer (see figure C2 in the online appendix). This suggests that CDD is a good proxy for energy demand while HDD is not. I also define total degree days (TDD), the sum of CDD and HDD, as an alternative to further account for the smaller spikes in the winter. Finally, for each CSI application i, I calculate the daily CDD and TDD and then average them over the relevant period (length T):

$$\overline{CDD}_i = \frac{1}{T}\sum_{\tau=1}^{T} CDD_{i\tau} \quad \text{and} \quad \overline{TDD}_i = \frac{1}{T}\sum_{\tau=1}^{T} TDD_{i\tau}.$$

Solar production channel. Real-time solar generation is a nonlinear function of several weather elements. Solar energy generation is primarily affected by the intensity of solar radiation, but the conversion rate is also affected by heat negatively and wind positively. I use a PV performance model developed by Kleissl (2013).[16] This model simulates solar panel output for given inputs of weather and system characteristics. For each CSI application i, I generate a series of daily indices by plugging in daily weather variables:

$$ProdIndex_{i\tau} = F(Solar_{i\tau}, Temp_{i\tau}, WindSpeed_{i\tau}), \text{ where } \tau = 1, ..., T.$$

Similar as before, I calculate the average in the relevant period:

$$\overline{ProdIndex_i} = \frac{1}{T}\sum_{\tau=1}^{T} ProdIndex_{i\tau}.$$

Since I keep system characteristics the same across applications, the production index measures how favorable the current weather is for solar power generation.

3.5 Economic Controls and Demographic Characteristics

Variations in weather are usually orthogonal to most other variables, but it is less certain in this study because weather variables are averaged over a period. Therefore, I also control for economic conditions that might drive solar demand. I collected the following monthly economic variables: leading index and unemployment rate from the FRED database[17], the prime interest rate, and index of consumer sentiment and buying conditions for the West region from the Survey of Consumers.

The CSI dataset does not provide individual demographics. Instead, I collect zip-code average demographics including race, education, income, housing cost, household size, etc. They are either from the 2010 Decennial Census or 5-year estimates from the 2011 American Community Survey (ACS). I use them to analyze whether responses to weather are heterogeneous across different communities.

4 Econometric Framework

The main analysis focuses on testing implication (I2), or how weather fluctuations in the post-contract period affect the probability of cancellation. The baseline estimating equation

[16] More information on the PV performance model is provided in the online appendix (section A).

[17] The leading index for each state predicts the six-month growth rate of the state's coincident index, which in turn measures the current state of economic activity.

is a linear probability model[18]:

$$1(Cancel)_{izt} = \beta_1 Solar_{izt} + \beta_2 Weather_{izt} + \beta_3 System_i + \beta_4 Econ_{zt} + \delta_z + \delta_y + \delta_q + \epsilon_{izt}, \quad (5)$$

where i denotes an application (project), z denotes the zip code, and t denotes the application date. $1(Cancel)_{izt}$ is an indicator of whether project i is eventually canceled. $Solar_{izt}$ is the average solar insolation in the relevant period for project i. $Weather_{izt}$ is a vector of temperature and wind speed variables to which households might also respond, corresponding to the same post-contract period. The coefficients on both solar and weather variables are of interest. $System_i$ is a vector of system characteristics in each application, including CSI rating (capacity), third-party ownership, and out-of-pocket unit cost. $Econ_{zt}$ are economic conditions in the month of date t, including the leading index, unemployment rate, prime interest rate, indices of consumer sentiment and buying conditions. δ_z, δ_y and δ_q denote zip code, year and quarter-in-year fixed effects, where the year and quarter are associated with the application date.

This specification incorporates a number of important considerations. First, it includes the relevant solar and weather variables simultaneously. Solar insolation is technically the most important factor in solar power generation and likely the most salient for customers. Wind speed and extreme temperature bins are also technically related to solar production and energy demand. Including them simultaneously addresses the concern of confounding weather patterns (Auffhammer et al. 2013). It also allows us to jointly examine the solar production channel and the increasing rate channel.

Secondly, this specification includes a rich set of fixed effects (FE) to account for non-weather drivers of cancellation that might be correlated with the weather. The zip code FEs control for cross-sectional correlations between weather amenities and neighborhood characteristics on a fine scale. The year FEs control for any statewide economic and policy shocks as well as climatic cycles such as the El Niño Southern Oscillation. The quarter-of-year FEs control for seasonal labor market conditions, which might lead to changes in demand for solar PV systems.

There is a trade-off in the choice of FEs: a richer set of FEs helps to rule out confounders, but it also eliminates meaningful variations. Adding more FEs might also exacerbate the attenuation bias due to measurement errors in the weather variables. To examine this trade-off

[18] Nonlinear models such as Probit or Logit are valid alternatives. However, a Probit model with zip code fixed effects might yield inconsistent estimates due to the incidental parameter problem (Greene 2004). Results based on Logit models are reported in the Appendix, but marginal effects cannot be calculated because fixed effects are included.

in a transparent manner, I present four other specifications in the main analysis as refinements of the baseline. In my preferred specification, I control for seasonal economic dynamics specific to the county by quarter-by-county FEs, and any trend in solar PV installations – or common economic and policy shocks to the entire state – by year FEs. Another specification includes county-quarter-year FEs, which addresses the concern that local policy or economic shocks might be correlated with the weather deviations. I also test the robustness of the results by adding zip-code-specific linear time trends where appropriate.

To further enhance identification, I control for economic conditions at the monthly level. I also include individual system characteristics as they might drive cancellation. In the later analysis of channels and other outcomes, I use the same set of FEs and controls because the threats to identification are similar.

Lastly, since there are overlaps in the weather stations matched with different zip codes, the standard errors are clustered by county in most tables. The number of clusters is 54.

5 Results

Table 1 presents the summary statistics. For the main analysis, I exclude applications whose host (customer) sector is non-residential, whose CSI rating is above 100 (less than 0.03%), and whose application submission starts after September 1, 2015. After cleaning, there are a total of 154,655 residential applications, of which 18,653 (12.06%) were eventually canceled. The average system cost is $35,377.04, which is more expensive than most home appliances. The average incentive amount is about one-tenth of system cost.

Canceled projects are larger and more expensive than completed ones. They are slightly more likely to be third-party-owned (TPO), that is, hosted by the customer but owned by the solar company. Canceled projects also appear to be located in zip codes with a slightly smaller fraction of whites and college graduates, and with lower median income.

In the formal analysis below, I first present a set of results and robustness checks on equation (5). This establishes that residential customers indeed respond to weather fluctuations in the post-contract period. Then, I extend the analysis by further disentangling the solar production and increasing rate channel, and by examining the dynamics of purchase and cancellation in response to weather shocks before and after the contract. I also explore potential heterogeneous responses from residential vs. non-residential systems, and across a host of different system and neighborhood characteristics. Last but not least, I discuss the implications of the empirical results.

Table 1: Summary Statistics

Variable	All Applications (N=154,655) Mean	All Applications (N=154,655) SD	Completed (N=136,002) Mean	Canceled (N=18,653) Mean
A. Application Characteristics				
CSI rating	4.86	3.02	4.82	5.22***
total cost ($)	35,377.04	23,829.66	35,090.3	37,467.7***
incentive amount ($)	3,551.24	4,838.65	3,526.86	3,729.01***
net unit cost ($)	6,897.71	3,179.06	6,901.94	6,866.86
third-party ownership	0.500	0.500	0.495	0.541***
year	2011.11	1.83	2011.09	2011.27***
B. Weather				
solar insolation	18.96	5.85	18.98	18.81***
max temp (F)	74.50	9.40	74.53	74.24***
#days(max $\leq$ 40)	0.0160	0.214	0.0157	0.0177
#days(max $>$ 100)	3.24	9.50	3.26	3.14
wind speed (knots)	4.99	1.48	5.00	4.93***
#(precipitation$>$ 0)	23.43	17.46	23.43	23.45
production index	98.51	31.45	98.63	97.66***
production index	98.51	31.45	98.63	97.66***
cooling degree days	2.81	3.94	2.82	2.76*
heating degree days	5.92	4.74	5.91	5.97
total degree days	8.73	3.79	8.73	8.73
C. Economic Conditions				
leading index	1.41	1.42	1.39	1.55***
unemployment	10.02	1.78	10.01	10.15***
prime interest rate	3.56	1.04	3.57	3.46***
index of consumer sentiment	74.2	8.13	74.19	74.27
index of buying condition	128.07	13.12	128.07	128.17
D. Zip Code Demographics and Interconnection Characteristics				
white (%)	67.30	15.88	67.53	65.65***
bachelor degree (%)	36.15	17.87	36.29	35.12***
median income ($)	79,414.95	27,052.83	79,617.02	77,940.58***
monthly housing cost ($)	2,170.04	655.63	2,172.93	2,148.93***
mean household size	2.86	.460	2.85	2.88***
urban (%)	91.5	18.7	91.5	91.7*
current installed base	198.68	191.54	200.94	182.12***
installed penetration	16.61	15.06	16.77	15.38***

Note. The weather variables in panel B are calculated over the 113-day period following the application submission. See table 9 for a description of the variables. Statistical significance of the differences between the completed and canceled subsamples are indicated in the last column. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

5.1 Do Cancellations Respond to Weather?

In this section, I estimate equation (5) to examine whether solar customers respond to weather shocks. The primary variable of interest is the mean solar insolation in the relevant time frame after sign-up. For the same period, I also include the following weather variables: the number of days with maximum temperature below 40°F, the same for above 100°F, and average wind speed. Including them simultaneously addresses the concern of confounding weather patterns and provides a first sense of the relative importance of the solar production vs. the increasing rate channel.

As discussed in section 3.2, I analyze two mutually-nonexclusive time frames: (a) 113 days after the sign-up and (b) 30 or 60 days before the completion/cancellation. Table 2 reports these results. Columns (1)-(2) present estimates under time frame (a), columns (3)-(4) show those under time frame (b) with a 60-day period, while (5)-(6) shorten the period to 30 days. In each pair of columns, the first column uses the baseline specification with zip code, year, and quarter-of-year fixed effects (FE), and the second allows the quarter-of-year FEs to differ across counties. As discussed in section 4, the latter is my preferred specification because it is well-balanced between keeping meaningful variations and eliminating threats to identification.

The coefficient on solar insolation is negative and statistically significant throughout. This implies that a residential customer exposed to less sunshine after sign-up is more likely to cancel the contract. The baseline estimate suggests that a one-standard-deviation decrease in solar insolation is associated with a 7.1% increase in cancellations, while the cancellation rate is 12.06% among all applications.[19] The preferred estimate in column (2) shows a 9.8% increase. Despite using different variation in the data, the estimates in columns (3)-(6) are remarkably similar in scale to column (2). Other aspects of the weather also appear to play a role. Higher wind speed is positively associated with cancellation, and an additional hot day has the opposite effect. On the other hand, the effect of an additional cold day is not stable in either direction or scale. In California, such days are very rare. This might have been a major contributing factor to unstable results.

The online appendix reports a variety of robustness checks. The key estimate on solar insolation is robust to three alternative fixed effects (FE) specifications: the inclusion of county-year FE, county-quarter-year FE, and month-of-year FEs (results using time frame (a) and (b) are reported in tables D1 and D2, respectively). It is also robust to varying the set of weather variables included in the model (table D3), and adding zip-code-specific

[19] Calculation: $\beta_1 \times SD(Inso)/D.V.Mean = 0.146 \times 5.85/12.06 = 7.1\%$.

Table 2: Cancellation and Solar Insolation

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
solar insolation	−0.146	−0.203	−0.186	−0.192	−0.186	−0.205
	[0.055]**	[0.045]***	[0.045]***	[0.045]***	[0.048]***	[0.050]***
wind speed	0.262	0.799	0.815	0.754	0.617	0.593
	[0.241]	[0.271]***	[0.173]***	[0.162]***	[0.139]***	[0.143]***
#days(tmax < 40)	1.055	1.406	−1.278	−1.572	−0.698	−0.691
	[0.383]***	[0.520]***	[0.489]***	[0.361]***	[0.450]	[0.490]
#days(tmax≥100)	−0.020	−0.024	−0.256	−0.330	−0.149	−0.214
	[0.023]	[0.021]	[0.030]***	[0.065]***	[0.023]***	[0.048]***
R^2 (within)	0.033	0.035	0.072	0.076	0.071	0.074
N	154,519	154,518	62,069	62,062	62,069	62,062
Sample						
Time Frame	First 113	First 113	Last 60	Last 60	Last 30	Last 30
Duration	All	All	0-100	0-100	0-100	0-100
D.V. Mean	12.06	12.06	6.70	6.70	6.70	6.70
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Note. This table reports estimates from equation (5). Columns (1)-(2) use the full sample and a time frame of 113 days after the sign-up. Columns (3)-(4) restrict the sample to observations with a reported duration of less than 100 days and use a time frame of 60 days before the completion/cancellation. Columns (5)-(6) are similar but shortens the time frame to 30 days. All regressions control for system characteristics and monthly economic conditions. All coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are by county. Full results and additional specifications are reported in tables D1 and D2 in the online appendix. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

annual or quarterly time trends (table D4). Finally, table D5 reports the results from Logit regressions with similar fixed effects. Although the fixed effects prevent valid calculations of marginal effects, the qualitative patterns are very similar and roughly proportional to the linear model.

I also look into the potential weaknesses of each time frame. The main concern for time frame (a) is that it might fail to capture the relevant period if the period length is not specified optimally. Table D6 shortens the post-period length to 70 or 90 days. All of the results stand similar to before, as is expected from period lengths that are close to the optimal. On the other hand, time frame (b) relies on using a subsample with valid cancellation dates. In the main results, this is done by imposing a sample restriction to

Table 3: Solar Production vs. Increasing Rate Channel

Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
ProdIndex	−0.761	−0.792	−0.650	−0.687	−0.830	−0.793
	[0.263]***	[0.259]***	[0.323]**	[0.362]*	[0.247]***	[0.252]***
CDD			−0.260	−0.223		
			[0.297]	[0.341]		
TDD					−0.126	−0.003
					[0.279]	[0.314]
R^2 (within)	0.034	0.035	0.034	0.035	0.034	0.035
N	154,519	154,518	154,519	154,518	154,519	154,518
Fixed Effects						
Zip Code	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Note. All regressions control for system characteristics and monthly economic conditions. The mean of the dependent variable is 12.06%. All indices are normalized. All coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

observations with a reported duration under 100 days. In table D7, I explore the effect of relaxing this restriction to include observations with duration up to 150 and 200 days. The qualitative pattern and statistical significance remain the same, but the scale of the estimates decreases. This is consistent with our expectation that these prolonged reports are less likely to be accurate. To briefly sum up, the results under time frame (a) and those using the most reliable subsample under time frame (b) are consistent with each other.

Next, I further disentangle the two main channels outlined in section 2.2. Specifically, I replace the solar and weather variables in equation (5) with two indices corresponding to the solar production channel and the increasing rate channel, respectively. This specification directly estimates and compares the effects of the two channels following the logic of an encompassing test (Davidson and MacKinnon 1993). The identifying assumption is that the other unobservable concerns are orthogonal to these two channels.

The construction of the two indices is detailed in section 3.4. The index for solar production channel is generated from an engineering model. As expected, it is highly correlated with solar insolation ($\rho = 0.98$). As the increasing rate channel works through changes in energy demand, it is represented by either cooling degree days (CDD) or the sum of heating and cooling degree days (TDD). Both indices are normalized so that the estimates correspond to percentage point change in cancellation rate associated with an increase in the index by one standard deviation.

Table 3 reports these results under time frame (a). In columns (1)-(2) where the production index is the only regressor, its coefficients are negative and statistically significant. Columns (3)-(4) account for CDD, resulting in slightly smaller and less precise estimates on the production index. This might be driven by the correlations between the two indices. Nevertheless, they are still much larger in scale and more significant than the coefficients on CDD. The last two columns switch to TDD to represent energy demand. The coefficients on TDD are even smaller, while those on the production index are similar to columns (1)-(2) in scale and statistical significance. Overall, these estimates suggest that the solar production aspect plays a greater role in the customers' decisions.

There are several possible explanations for the weak response through the increasing rate channel. For one, the channel does not apply to customers whose marginal electricity rate stays the same with or without the solar PV system. Moreover, customers tend to respond to average rather than marginal rate, which attenuates the rate increase. In fact, the increasing rate schedule might not even be salient in this decision. Lastly, for those customers who are primarily motivated by green preference, the billing aspect is simply unimportant.

These results are robust to using the the alternative time frame (b) (see table D8 in the online appendix). Using the production index, I also examine the validity of using a linear specification. This is done by plotting the non-parametric relationship between cancellation and the index using a fitted local polynomial and a binned scatter plot (see figure C3 in the online appendix). Both show a mostly linear relationship. Last but not least, I examine the attenuation bias regarding assumption (a) further. As discussed in section 3.2, using a period length that is further away from the optimal point will result in larger attenuation bias and smaller estimates. I test this prediction by estimating the same model using alternative windows of 10, 30, 50, 70, and 90 days to construct the production index (see table D9 in the online appendix). The estimates are increasing in the length of period and all except for the 90-day one are statistically different from the main (113-day) estimate. This pattern is consistent with the assumption that 113 days are close to the optimal length.

5.2 Do Sign-Ups Respond to Weather?

This section tests implication (I1), that more favorable weather leads to an increase in purchase. This, in fact, is a recurring finding in the previous studies. I provide evidence in a similar vein. Unlike cancellation, this outcome cannot be analyzed at the individual level as

Table 4: Pre-Period Weather and Purchase

	Number of Applications			log(Number of Applications)		
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-period length	1Mo	2Mo	3Mo	1Mo	2Mo	3Mo
ProdIndex (pre)	0.551	0.931	0.891	0.116	0.215	0.212
	[0.188]***	[0.194]***	[0.274]***	[0.034]***	[0.033]***	[0.047]***
D.V. Mean	3.36	3.36	3.36	-	-	-
R^2	0.410	0.410	0.410	0.481	0.482	0.481
N	45,875	45,858	45,768	45,875	45,858	45,768

Note. The dependent variable is the number of applications in columns (1)-(3) and its log level in columns (4)-(6). The unit of analysis is zip-code-by-month. The length of the pre-period used to calculate the production index is indicated in the header. All regressions control for monthly economic conditions and zip code, year, and month fixed effects. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

we do not observe all potential solar customers. The unit of analysis is zip-code-by-month instead. Moreover, since the relevant period for the sign-up decision is harder to infer, I show results with various pre-contract windows.

Table 4 shows these results with zip code, year, and month-in-year fixed effects. Columns (1)-(3) use the number of applications as the dependent variable, and columns (4)-(6) use its log. Three estimates are shown for each dependent variable, corresponding to an assumed window of 1, 2, and 3 months for calculating the production index.[20] The estimates are all positive and statistically significant for both level and log. When the index increases by one standard deviation, the level estimates suggest 0.6-0.9 additional applications and the log estimates suggest a 11.6-21.5% increase. I further test for harvesting effects by augmenting column (1) with one lead and six lags of the production index. These estimates are plotted in figure C5 in the online appendix. Full harvesting implies that the positive effects of favorable weather should be followed by negative effects of similar magnitudes so that the net effect is zero. The results show that the positive effect peaks at a month later and reverses signs afterwards, but the overall effect is positive. This suggests that harvesting drives part of the dynamics in sign-up, but not entirely. These results align with Lamp (2018)'s main finding in Germany and complements the above findings on cancellation. On the intensive margin, favorable weather does not appear to induce systematic selections into larger system sizes (see table D10 in the online appendix).

[20]The window starts from the current month. For example, column (2) uses the average production index in the current and previous month.

5.3 Response to Weather Updates

This section examines how the cancellation decision is affected by the dynamics of pre- and post-contract weather. Because cancellation reflects a change in valuation, it might be a response to an update in post-period weather relative to the pre-period.[21] Therefore, analyzing the response to weather updates can provide insight into how customers are processing weather information at different points in time.

In table 5, I examine whether the customers are responding to the change in post-period weather relative to the pre-period ("weather update" henceforth). I use a 30-day pre-period, a 113-day post-period, and the baseline FE specification in this analysis.[22] Column (1) estimates the relationship between cancellation and the pre-period production index. The coefficient is positive but insignificant. In column (2) I add weather update, which yields a negative and statistically significant coefficient. This suggests that conditional on the pre-period weather, customers respond to weather updates.

Column (3) tests whether a negative update has a different effect on cancellation than a positive one. The motivation is that a positive update might not be as informative given that these customers have already self-selected into a contract. I include an indicator of whether the update is negative and an interaction term of the indicator and the update variable, so that the latter will capture the differential effect of a negative update relative to a positive one. Indeed, the interaction term completely absorbs the effect of the *Update* variable and becomes twice as large. This provides strong support for different information contents between positive and negative updates.

In the next two columns, I estimate a similar model but change the content of the indicator: it indicates whether the pre-period index is higher than zip-code average in column (4); and whether the post-period index is lower than zip-code average in column (5). Larger impacts are found on both margins. These results suggest that the customers fail to recognize the transient nature of weather shocks.

5.4 Heterogeneous Effects

This section explores whether responses are heterogeneous across a variety of customer and neighborhood characteristics. I classify each observation under one of two categories based

[21] Two other studies have touched on this point. Conlin et al. (2007) control for weather in both periods without explicitly examining updates. Chang et al. (2018) show evidence that the comparison is important.

[22] Using a 60-day pre-period yields qualitatively similar results (see table D11 in the online appendix). Results under alternative FE specifications are also very similar and available upon request.

Table 5: Responses to Weather Updates

Cancel = 1	(1)	(2)	(3)	(4)	(5)
ProdIndex (pre)	0.253	−0.472	−0.055		
	[0.235]	[0.313]	[0.287]		
Update		−0.802	−0.124	−0.257	0.368
		[0.269]***	[0.601]	[0.203]	[0.443]
Indicator:					
Update < 0			−2.027		
			[0.404]***		
ProdIndex (pre) > $\overline{\text{ProdIndex}}_z$				−0.383	
				[0.687]	
ProdIndex (post) < $\overline{\text{ProdIndex}}_z$					0.437
					[0.351]
Update × Indicator			−1.746	−0.556	−0.976
			[0.555]***	[0.314]*	[0.334]***
R^2	0.033	0.033	0.034	0.033	0.033
N	154,519	154,519	154,519	154,519	154,519

Note. "ProdIndex (pre)" is the normalized average production index for a 30-day pre-contract period. "Update" is the average post-period index minus the pre-period one. The first indicator indicates whether the update variable is negative. The second indicates whether the pre-period index is higher than the zip-code average, and the third is whether the post-period one is lower than average. All regressions control for system characteristics and monthly economic conditions, as well as zip code, year, and quarter fixed effects. The mean of the dependent variable is 12.06%. All coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

on the characteristic of interest. I then estimate the differential responses by adding an interaction term of the production index and an indicator for one of the categories.

An informative margin of heterogeneity is residential vs. non-residential customers. The latter group – including small businesses, government agencies, and NGOs – are also eligible for the CSI rebate. Compared to residential customers, their decisions are more likely to involve formal profitability calculations and less prone to psychological effects. There are altogether 8,704 non-residential applications in the cleaned dataset, including 1,730 small systems (<10kW) and 6,974 large ones ($\geq$ 10kW). To apply for CSI rebate, the small systems follow the same 2-step procedure as residential projects while the large systems follow a 3-step process. Rather than a signed contract, CSI requires the payment of an application fee within 30 days of the initial application for large systems. Overall, non-residential applications are much more likely to be canceled (27.90%), potentially due to greater complexity in the financing and installation process.

The regressions largely follow the baseline and preferred FE specifications but change

Table 6: Cancellations of Non-Residential Applications

	All		Small Systems		Large Systems	
Cancel = 1	(1)	(2)	(3)	(4)	(5)	(6)
β_1: ProdIndex	-0.629	-0.669	-0.743	-0.779	-0.677	-0.714
	[0.260]**	[0.257]**	[0.265]***	[0.257]***	[0.255]**	[0.253]***
β_2: ProdIndex $\times$ 1(non-resid)	1.194	1.237	2.228	2.171	1.117	1.206
	[0.569]**	[0.567]**	[1.536]	[1.552]	[0.766]	[0.760]
$\beta_1 + \beta_2$	0.565	0.568	1.485	1.392	0.440	0.492
p-value: $\beta_1 + \beta_2 = 0$	0.38	0.36	0.34	0.38	0.59	0.54
R^2 (within)	0.057	0.058	0.040	0.041	0.055	0.057
N	163,059	163,058	155,947	155,946	161,336	161,335
Non-residential sample						
D.V. Mean	27.90	27.90	21.45	21.45	29.50	29.50
N	8,504	8,704	1,730	1,730	6,974	6,974
Fixed Effects						
Zip-type	Yes	Yes	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes	Yes	Yes
Quarter	Yes		Yes		Yes	
County-quarter		Yes		Yes		Yes

Note. The regressions use pooled samples of residential and non-residential applications. Columns (1)-(2) include all non-residential applications, (3)-(4) include only small systems (<10kW), and (5)-(6) include only large systems ($\geq$ 10kW). The last two rows in the top panel show the value of $\beta_1 + \beta_2$, which is the weather effect for non-residential customers, and the corresponding p-value. All regressions control for system characteristics and monthly economic conditions. The production index is normalized, and all coefficients are multiplied by 100 for legibility. Zip-code-by-type FEs are included instead of zip code FEs. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

the zip-code FEs into zip-code-by-type ones to control for finer cross-sectional variations. Columns (1)-(2) in table 6 pool the residential sample with all 8,704 non-residential applications, while columns (3)-(4) include only small ones, and (5)-(6) include only large ones. The coefficient on $ProdIndex$ captures the effect of weather on residential customers and is similar to the original estimates in table 3. The interaction term has a positive coefficient in all columns. The sum of these two coefficients suggests a slightly positive effect for the non-residential customers, which is the opposite of the effect on residential customers. The positive effect is even larger for small systems that are comparable in size to residential projects. Despite this puzzling observation, the F-test fails to reject having a null effect in all columns. Overall, these results suggest that non-residential customers do not fall under the same behavioral influence.

Table 7: Peer Effects in Cancellations and Response to Weather

	Current Installed Base		Installed Base > 1 Year	
Cancel = 1	(1)	(2)	(3)	(4)
ProdIndex	−1.005	−0.716	−1.037	−.751
	[0.412]**	[0.367]*	[0.401]**	[0.351]**
Penetration	−0.096	−0.086	−0.124	−0.109
	[0.028]***	[0.027]***	[0.047]**	[0.046]**
ProdIndex × Penetration	0.017	0.001	0.027	0.004
	[0.016]	[0.013]	[0.022]	[0.017]
R^2	0.054	0.032	0.054	0.032
N	150,859	151,029	150,859	151,029
Fixed Effects				
Zip Code		Yes		Yes
Zip Code-quarter	Yes		Yes	
County-quarter		Yes		Yes
Year	Yes	Yes	Yes	Yes

Note. Market penetration is defined as the number of interconnected systems per thousand households in columns (1)-(2), and that of systems which have been interconnected for at least one year in columns (3)-(4). All regressions control for system characteristics and monthly economic conditions. The mean of the dependent variable is 12.06%. The production index is normalized. All coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Next, I explore heterogeneous responses among residential applications in locations with different solar market penetration. As peer effects are found to play a role in the adoption of solar panels (Bollinger and Gillingham 2012), they might also affect the cancellation decision. For instance, a solar customer might be less likely to cancel her contract if her neighborhood has more solar PV systems installed. That customer might also learn from her neighbors with older systems to recognize the transient nature of the fluctuations in solar power generation. These hypotheses are tested below.

Identification of peer effects in Bollinger and Gillingham (2012) relies on the fact that installation typically takes place with some delay after the sign-up, and the peer effects begin only after the installation. The increase in the current installed base (past sign-ups) is thus predetermined and uncorrelated with other contemporaneous shocks to the peer group, yielding a consistent estimate. Following this logic, I construct two measures of solar market penetration at the zip-code-by-month level and merge them with each application based on the sign-up month. They are predetermined under the same assumption. The first statistic is the number of all interconnected (installed) solar systems per thousand households, and

Table 8: Heterogeneous Effects by System Characteristics

Cancel = 1 Characteristic	(1) System Size Above Median	(2) Average Cost Above Median	(3) Third-Party Ownership	(4) Near Incentive Step Change
β_1: ProdIndex	−0.753	−0.787	−0.449	−0.595
	[0.309]**	[0.275]***	[0.251]*	[0.280]**
β_2: 1(Characteristic)	1.723	−1.221	0.837	1.863
	[0.383]***	[0.615]*	[0.522]	[0.580]***
β_3: ProdIndex × 1(Characteristic)	0.007	0.081	−0.638	−0.251
	[0.270]	[0.210]	[0.292]**	[0.334]
$\beta_1 + \beta_3$	−0.746	−0.706	−1.087	−0.846
p-value: $\beta_1 + \beta_3 = 0$	0.010***	0.019**	0.003***	0.030**
R^2	0.031	0.033	0.033	0.033
N	154,519	154,520	154,519	154,519

Note. The characteristic of interest is described in the header. In particular, the indicator in column (4) records whether the application is within 47 days to the next incentive step change. The last two rows in the top panel show the value of $\beta_1 + \beta_3$, which is the weather effect for customers with the characteristic of interest, and the corresponding p-value. All regressions control for other system characteristics and monthly economic conditions, as well as zip code, year, and quarter fixed effects. The mean of the dependent variable is 12.06%. The production index is normalized. All coefficients are multiplied by 100 for legibility. Standard errors (in squared brackets) are clustered by county. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

the second is one for systems that are at least one year old. These calculations are based on the net energy metering (NEM) interconnection dataset, which contains all interconnected projects in California with information on zip code and the relevant dates.

Table 7 presents results of regressing cancellation on the production index, market penetration, and their interaction. Columns (1)-(2) are estimated using the current installed base and (3)-(4) are estimated using those older than one year. The coefficient on the production index is similar to before. As for market penetration, one more system per thousand households decreases cancellation probability by around 0.09 percentage point in columns (1)-(2), and 0.12 percentage point in (3)-(4). This finding complements the previous studies: peer effects not only motivates the households to sign up but also prevents them from regretting afterward. However, the coefficient on the interaction term remains small and statistically insignificant throughout, suggesting that having neighbors with older systems might not guard against the psychological effects of current weather.

I also explore whether responses differ with system and application characteristics (see table 8). Column (1) shows that larger systems are more likely to be canceled, and (2) shows the same for those with a lower average cost. However, the small coefficients on the interaction terms suggest that these characteristics do not systematically affect response to weather. Column (3) examines third-party ownership (TPO). TPO is a financing option that allows the household to acquire the system without paying a large up-front cost. The household signs a power-purchase agreement (PPA) or a lease with the solar company, who owns the system.[23] About half of all applications in the data are under TPO. It is unclear whether customers with TPO systems should be more or less sensitive to weather. On one hand, TPO partially insures the risk from weather variations by compensating the customer when production is particularly low. On the other hand, selection into TPO is not random, as the option is created to appeal to customers who are more liquidity-constrained, risk-averse, or hesitant. The estimates suggest that they are much more sensitive to weather: the weather effect on them is more than twice of that on self-owned systems.

Column (4) examines heterogeneous effects based on whether the application is submitted near the next incentive step change. Step changes occur when the total solar capacity reserved in the program exceeds certain thresholds. Although the exact timing of the change is unknown, solar companies might still be able to anticipate an imminent step change. They might even take advantage of it and push marginal customers to sign up in order to secure a higher rebate.[24] Therefore, we might expect higher propensity to cancel and a stronger weather effect among these customers. I calculate the number of days to the next step change for each application based on the realized schedule.[25] The sample is split by whether this duration is less than 47 days, the 25th percentile of this variable (see figure C5 in the online appendix for the distribution of days to next step change).[26] The result suggests that these customers did have a substantially higher probability of cancellation. They are also slightly more sensitive to weather but this effect is insignificant.

Lastly, I examine whether there are heterogeneous responses across zip codes based on aggregate characteristics (as household-level characteristics are not available). These zip-code-level characteristics include median income, the fraction white population, the fraction with a bachelor degree, the fraction of urban households, average household size, average housing cost, whether the zip code is in Northern California, and whether it is in non-coastal

[23]PPA allows the household to pay a low fixed rate for the electricity generated from the system, while the lease is similar to a loan which the household need to pay back.

[24]According to a CSI staff, "... there was always a flurry of reservations and overall activity around step changes".

[25]Source: https://www.californiasolarstatistics.ca.gov/reports/budget_forecast/#Step.

[26]The results are robust to other thresholds in a similar range. There is no special reason to focus on the 25th percentile. The threshold should be far enough from the step change to have sufficient identifying variations while close enough to reflect imminent change.

counties. The results are reported in table D12 in the online appendix. The only notable pattern is that the responses to weather are entirely driven by zip codes that have more urban households than rural ones. The rest of the zip code characteristics do not seem to split the sample in meaningful ways. This might be because these aggregate characteristics are not informative about the highly-selected group of solar customers (Harding and Rapson 2013; Borenstein and Davis 2015).

6 Mechanisms

In this section, I propose a number of psychological and rational mechanisms that might have driven the effects of weather in the solar market. Each mechanism is evaluated against the empirical evidence.

Projection bias. Projection bias is the tendency for people to exaggerate the degree to which their future preference will resemble the present one. This leads solar customers to overestimate the total utility of going solar when weather conditions bring higher marginal return and regret when conditions worsen. The empirical evidence on both cancellation and sign-up are consistent with the theoretical implications of projection bias in section 2.

Misprediction of utility from a durable good comes in two types: mispredicting either the path of states or the utilities generated in those states. They generate the same predictions and have not been disentangled empirically in the literature (DellaVigna 2009). Since solar panels are largely instrumental, I argue that mispredicting the objective states is more plausible.[27]

Salience effects. Bordalo et al. (2013) propose a theory of context-dependent choice, where a consumer's attention is drawn to salient attributes of goods. I find empirical evidence that the solar customers respond most strongly to weather patterns associated with solar productivity. One explanation is that sunny weather makes the benefit from solar panels more salient relative to the cost, leading to more sign-ups. Likewise, gloomier weather reduces the benefit-cost ratio by reducing the salience of the benefit, thus leading to regrets. Therefore, this mechanism also generates implications that are consistent with the sign-up and cancellation results.

One way to distinguish projection bias from salience effects is to test if the consumers are responding to the absolute "levels" of weather rather than deviations from the norm (Busse

[27] This interpretation might not be valid for cases where there is a large subjective component in the utility of going solar, such as a customer with strong green preferences.

et al. 2015). However, the test conflicts with our need to control for non-weather seasonal confounders. For this reason, the literature, including this study, has not been able to tell them apart.

Bayesian learning. Another possibility is that customers are in fact extracting useful information from weather fluctuations to guide their decision. Below, I discuss two main scenarios regarding how learning might occur.

First, the customers might need to learn about existing weather patterns because they were previously inattentive. However, the fixed effects in the regressions preclude the identification of this type of learning. Once the zip code fixed effects absorb most of the meaningful correlations between the short- and long-run weather, the former is no longer predictive of the latter (see table D13 in the online appendix). It is also hard to reconcile with the cancellation results. By the time of the sign-up, the customers should have been informed by the installers on the estimated benefits. In particular, these estimates are based on simulations of local climate, thereby making it unnecessary for the customers to directly observe the weather.

Second, the customers might be learning about climate change, which might affect the desirability of solar power. Previous studies find that exposure to unusual heat increases people's belief in climate change (Li et al. 2011; Deryugina 2013). However, the temporal fixed effects should have absorbed the gradual movements driven by climate change, again precluding identification of meaningful learning. Furthermore, the effect of solar insolation remains highly significant after controlling for extreme heat or cold (table 2). Therefore, learning about climate change is also not likely to be the main mechanism.

Supply-side factors. Supply-side factors might also drive adoption. If solar companies ramp up their promotion efforts to take advantage of the good weather, it would directly increase sign-ups and might also help reduce cancellations if the latter is affected by visible marketing efforts. The anecdotal evidence has been mixed. For instance, one salesman notes that it is easier to make a sale in the summer than the winter. On the other hand, my interview with a major company finds no high-level awareness or effort to incorporate current weather into the marketing plan. While the importance of this mechanism is difficult to gauge, it ultimately relies on people's psychological reactions to the weather. To the extent it is important the results should be interpreted as the overall effect, including the demand-side response together with any amplification effect from marketing practices.

Another possibility is that the solar workers are more productive in better weather. Solar salesmen and installers are outdoor workers whose labor supply and productivity are subject to weather fluctuations (Graff Zivin and Neidell 2014). This could potentially change the entire interpretation because it is independent of any demand-side effect. In particular, one concern is that the installation might be repeatedly re-scheduled due to bad weather, eventually leading to cancellation. I partially test the veracity of this scenario by examining, among the completed projects, whether the completion duration is longer when the average weather (as measured by production index) is worse. There is no evidence of a significant relationship therein (see table D14 in the online appendix). Moreover, I find that extreme heat *reduces* cancellation in table 2. This is also inconsistent with the existing literature on weather and productivity where the biggest impact is often associated with extreme heat. Overall, there is little evidence in support of worker productivity being the main mechanism.

7 Conclusion

This paper studies the effect of short-run weather fluctuations on cancellations of signed solar contracts in California. This decision appears to respond most strongly to short-run weather patterns associated with solar panel productivity. The main result shows that customers whose sign-up for solar panels is followed by unfavorable weather are more likely to cancel their contract. Furthermore, they are more likely to cancel when they experience worse weather after signing up relative to before, suggesting that the negative update changes their valuation of solar panels. I also find suggestive evidence that non-residential customers do not respond to weather, and the marginal ones respond more. The most plausible mechanism consistent with all these results is psychological, such as projection bias or a salience effect. Under such mechanisms, the customers rely too heavily on current weather to predict their future utility from solar panels.

These results suggest that projection bias or salience effects can substantially affect the demand for a big-ticket item such as a solar PV system. Compared to the other contexts studied in the literature, such biases in solar PV adoption create new welfare implications and policy challenges. Solar panels are under-adopted socially since electricity generated from fossil fuels is not correctly priced. They might also be under-adopted privately for reasons similar to the case of energy efficiency. Therefore, the undervaluation caused by these biases creates a much larger welfare loss than over-valuation. This highlights the importance of continual communications with the customers who are waiting for their systems to be installed, especially during periods of unfavorable weather. Given that cancellation reflects a wasteful cycle of over- and undervaluation created by weather fluctuations, it is also important to “de-bias” the customers and mitigate these cycles.

Finally, there are several limitations in this paper. For each I propose a direction for future work. The first limitation is not having the precise cancellation dates. While solar PV adoption is an interesting context where projections are continually updated within the relevant period, the lack of precise dates prevent me from exploring how the weather dynamics *within* this period are incorporated into decisions. Second, without demographic information of the customers, I am not able to fully examine whether the effects differ by education background, income, or household structure. Documenting these heterogeneities can help researchers deepen the current understanding of behavioral biases, and it can also help solar companies and behavioral programs to better target vulnerable populations. Third, this paper finds that information provided by the solar companies or peer effects by neighbors do not eliminate such bias. Future studies can explore more effective strategies. Answering these questions are important for both economic modeling and policy design.

References

Allcott, Hunt. 2011. "Consumers' perceptions and misperceptions of energy costs". *American Economic Review* 101 (3): 98–104.

— . 2016. "Paternalism and energy efficiency: an overview". *Annual Review of Economics* 8:145–176.

— . 2013. "The welfare effects of misperceived product costs: Data and calibrations from the automobile market". *American Economic Journal: Economic Policy* 5 (3): 30–66.

Allcott, Hunt, and Sendhil Mullainathan. 2010. "Behavior and energy policy". *Science* 327 (5970): 1204–1205.

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky. 2014. "Energy policy with externalities and internalities". *Journal of Public Economics* 112:72–88.

Attari, Shahzeen Z, Michael L DeKay, Cliff I Davidson, and Wändi Bruine De Bruin. 2010. "Public perceptions of energy consumption and savings". *Proceedings of the National Academy of Sciences* 107 (37): 16054–16059.

Auffhammer, Maximilian, Solomon M Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. "Using weather data and climate model output in economic analyses of climate change". *Review of Environmental Economics and Policy* 7 (2): 181–198.

Baker, Erin, Meredith Fowlie, Derek Lemoine, and Stanley S Reynolds. 2013. "The Economics of Solar Electricity". *Annual Review of Resource Economics* 5 (1): 387–426.

Bollinger, Bryan, and Kenneth Gillingham. 2012. "Peer effects in the diffusion of solar photovoltaic panels". *Marketing Science* 31 (6): 900–912.

Bollinger, Bryan K, and Kenneth Gillingham. 2014. "Learning-by-doing in solar photovoltaic installations". *Available at SSRN 2342406.*

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and consumer choice". *Journal of Political Economy* 121 (5): 803–843.

Borenstein, Severin. 2012. "The redistributional impact of nonlinear electricity pricing". *American Economic Journal: Economic Policy* 4 (3): 56–90.

Borenstein, Severin, and Lucas W Davis. 2015. "The distributional effects of US clean energy tax credits". In *Tax Policy and the Economy, Volume 30*. University of Chicago Press.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer. 2013. "Are consumers myopic? Evidence from new and used car purchases". *American Economic Review* 103 (1): 220–56.

Busse, Meghan R, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. 2015. "The psychological effect of weather on car purchases". *The Quarterly Journal of Economics* 130 (1): 371–414.

Chang, Tom Y, Wei Huang, and Yongxiang Wang. 2018. "Something in the air: Pollution and the demand for health insurance". *The Review of Economic Studies* 85 (3): 1609–1634.

Conlin, Michael, Ted O'Donoghue, and Timothy J Vogelsang. 2007. "Projection bias in catalog orders". *American Economic Review* 97 (4): 1217–1249.

Crago, Christine Lasco, and Ilya Chernyakhovskiy. 2017. "Are policy incentives for solar power effective? Evidence from residential installations in the Northeast". *Journal of Environmental Economics and Management* 81:132–151.

Davidson, Russell, James G MacKinnon, et al. 1993. *Estimation and inference in econometrics.* JSTOR.

Dell, Melissa, Benjamin F Jones, and Benjamin A Olken. 2014. "What do we learn from the weather? The new climate-economy literature". *Journal of Economic Literature* 52 (3): 740–98.

DellaVigna, Stefano. 2009. "Psychology and economics: Evidence from the field". *Journal of Economic Literature* 47 (2): 315–372.

Deryugina, Tatyana. 2013. "How do people update? The effects of local weather fluctuations on beliefs about global warming". *Climatic Change* 118 (2): 397–416.

Deschenes, Olivier, and Michael Greenstone. 2007. "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather". *American Economic Review* 97 (1): 354–385.

EIA. 2016. "Sources of Heating for California Homes". Visited on 04/28/2016. http://apps1.eere.energy.gov/states/residential.cfm/state=CA.

Gillingham, Kenneth, Amelia Keyes, and Karen Palmer. 2018. "Advances in evaluating energy efficiency policies and programs". *Annual Review of Resource Economics* 10:511–532.

Graff Zivin, Joshua, and Matthew Neidell. 2014. "Temperature and the allocation of time: Implications for climate change". *Journal of Labor Economics* 32 (1): 1–26.

Greene, William. 2004. "The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects". *The Econometrics Journal* 7 (1): 98–119.

Harding, Matthew, and David Rapson. 2013. "Do voluntary carbon offsets induce energy rebound? A conservationist's dilemma". *Working Paper*.

Hastings, Justine S, and Jesse M Shapiro. 2013. "Fungibility and consumer choice: Evidence from commodity price shocks". *The Quarterly Journal of Economics* 128 (4): 1449–1498.

Hausman, Jerry A. 1979. "Individual discount rates and the purchase and utilization of energy-using durables". *The Bell Journal of Economics*: 33–54.

Hughes, Jonathan E, and Molly Podolefsky. 2015. "Getting Green with Solar Subsidies: Evidence from the California Solar Initiative". *Journal of the Association of Environmental and Resource Economists* 2 (2): 235–275.

Ito, Koichiro. 2014. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing". *American Economic Review* 104 (2): 537–563.

Kleissl, Jan. 2013. *Improving Economics of Solar Power through Resource Analysis, Forecasting, and Dynamic System Modeling*. Tech. rep.

Kőszegi, Botond, and Adam Szeidl. 2012. "A model of focusing in economic choice". *The Quarterly Journal of Economics* 128 (1): 53–104.

Lamp, Stefan. 2018. "Sunspots that matter: the effect of weather on solar technology adoption". *TSE Working Paper*.

Li, Ye, Eric J Johnson, and Lisa Zaval. 2011. "Local warming: daily variation in temperature affects beliefs and concern about climate change". *Psychol Sci* 22:454–459.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility". *The Quarterly Journal of Economics*: 1209–1248.

Magistris, Tiziana de, and Azucena Gracia. 2016. "Assessing Projection Bias in Consumers' Food Preferences". *PloS one* 11 (2): e0146308.

Pless, Jacquelyn, and Arthur A van Benthem. Forthcoming. "Pass-Through as a Test for Market Power: An Application to Solar Subsidies". *American Economic Journal: Applied Economics*.

Sallee, James M. 2014. “Rational inattention and energy efficiency”. *The Journal of Law and Economics* 57 (3): 781–820.

Simonsohn, Uri. 2010. “Weather to go to college”. *The Economic Journal* 120 (543): 270–280.

Stackhouse Jr, Paul W, David Westberg, James M Hoell, William S Chandler, and Taiping Zhang. 2015. *Prediction Of Worldwide Energy Resource (POWER) Agroclimatology Methodology*. Tech. rep. NASA Langley Research Center and SSAI/NASA Langley Research Center.

Appendix

Table 9: Variable Descriptions

Variable	Description	Source
A. Application Characteristics		
CSI rating	AC output of photovoltaic module	CSI
total cost ($)	Total cost of eligible portion of the project	CSI
incentive amount ($)	CSI incentive amount for the system	CSI
net unit cost ($)	Net out-of-pocket cost per kW of capacity	CSI
third-party ownership	Indicates whether the system owner is a different entity than the host customer	CSI
year	Application begin year	CSI
B. Weather		
(The term "post-period" below refers to the 113 days following application submission.		
solar insolation	Solar insolation, post-period average	POWER/NASA
max temp (F)	Maximum temperature, post-period average	GSOD/NOAA
#days(max $\leq$ 40)	Number of days with max temp below 40F in the post-period	GSOD/NOAA
#days(max $>$ 100)	Number of days with max temp above 100F in the post-period	GSOD/NOAA
wind speed (knots)	Wind speed, post-period average	GSOD/NOAA
#(precipitation$>$ 0)	Number of days with positive precipitation in the post-period	GSOD/NOAA
production index	Index of solar panel productivity, calculated using solar insolation, max temp, and wind speed, post-period average	Calculated
cooling degree days	Total cooling degree days in the post-period	Calculated
heating degree days	Total heating degree days in the post-period	Calculated
total degree days	Sum of cooling and heating degree days in the post-period	Calculated
C. Economic Conditions		
leading index	The leading index for California, predictor of six-month growth	FRED
unemployment	The unemployment rate in California	FRED
prime interest rate	The U.S. (Fed) prime rate	
index of consumer sentiment	Index of consumer sentiment in the West	Survey of Consumers

Continued on next page

Table 9 – continued from previous page

Variable	Description	Source
index of buying condition	Index of buying conditions in the West	Survey of Consumers
D. Zip Code Demographics and Interconnection Characteristics		
white (%)	Percent of white population in the zip code	2011 ACS
bachelor degree (%)	Percent of population with bachelor's degree in the zip code	2011 ACS
median income ($)	Median income in the zip code	2011 ACS
monthly housing cost ($)	Average monthly housing cost in the zip code	2011 ACS
mean household size	Average household size	2010 Census
urban (%)	Percent of urban households	2010 Census
current installed base	Number of interconnected solar systems in the zip code	NEM Interconnection
installed penetration	Number of interconnected solar systems per thousand households in the zip code	NEM Interconnection