Spatially Differentiated Renewable Energy Subsidies: What Is There to Gain?

Anatolii Kokoza*

University of Arizona

October 31, 2016

Abstract

I investigate the benefits and costs of geographically differentiating subsidies for rooftop solar photovoltaic systems. I estimate the intermittency costs of solar and subsidy spending required to achieve a given target by combining a model of the residential solar installation decision with a model of a system operator managing the region's electric grid. I find that offering the subsidy to a broader range of regions, rather than concentrating subsidies in the sunniest regions, reduces both intermittency costs and subsidy spending. I also find that differentiating subsidies to target the sunniest sites can reduce the subsidy spending needed to meet a solar generation target by over 5%. These savings are greater than the additional intermittency costs imposed by concentrating installations in the sunniest areas. Finally, I find there there is more to gain from differentiating subsidies in climatically and demographically diverse areas.

*I thank Gautam Gowrisankaran, Ashley Langer, Derek Lemoine, and Stanley Reynolds for their guidance during this project. For helpful feedback I thank Severen Borenstein, Jonathan Hughes, and participants at the 2016 AERE Summer Conference, 2016 WEAI Annual Conference and the 18th Annual CU Environmental and Resource Economics Workshop. I also thank Alexander Cronin, William Holmgren and the University of Arizona Photovoltaics Research Lab for solar PV generation data. An allocation of computer time from the UA Research Computing High Performance Computing and High Throughput Computing at the University of Arizona is gratefully acknowledged. Any errors are of course my own.

1 Introduction

Renewable energy sources, such as wind and solar, are subsidized throughout the world to reduce carbon emissions from electricity generation. These subsidies have two main goals: to drive technological innovation, reducing their costs in the future, and as a second-best policy to account for the positive externality of reduced emissions. This paper focuses on the latter; in particular, how the social benefits of solar vary spatially. I consider whether geographically differentiated subsidies for photovoltaic solar installations could incentive a more efficient geographic distribution of installations. The efficiency of a distribution of installations is determined not only by the timing of generation and generation potential of its sites, but also by the covariance of generation across sites. While the former is properly incentivized via, for example, net metering under real time pricing with internalized emission costs, the latter is not. The large sum of public funds set aside for renewable energy subsidies makes an investigation of the efficiency of subsidies pertinent. I show how the benefits of geographically differentiating subsidies vary across regions based on each region's climatic variation and demographics such as income and home value.

To determine how households respond to subsidies, I estimate the parameters influencing households' solar installation decisions. I model this as a discrete choice model of the installation decision together a reduced form choice of system size, conditional on installation. The estimated parameters then determine how the geographic distribution of residential solar installations changes under counterfactual subsidy policies. Previous work by Burr (2014), Hughes and Podolefsky (2015) and Gillingham and Tsvetanov (2016) consider the effects of subsidies on encouraging solar installations. I contribute to this literature by considering how residential households' response to subsidies affects social welfare via the intermittency costs of solar in the electric grid.

To measure social welfare under counterfactual subsidy policies, I combine the model of solar installations with a structural model of an expected-welfare-maximizing electricity market system operator from Gowrisankaran et al. (2016). The system operator takes as given the existing conventional generators and the distribution of solar installations to make an initial investment decision and hourly scheduling of generators and demand management to meet electricity demand. First she chooses investment in new fossil fuel generating capacity and a price at which she curtails electricity demand. Observing the weather forecast in each hour, she dispatches generators and chooses a quantity of demand to curtail in order to meet the expected demand net of non-dispatchable generation. The optimal choices of investment, demand management and scheduling of generators maximize expected welfare from electricity consumption and generation. I compute intermittency costs of solar as the difference in this measure of welfare under a baseline model and one in which solar is perfectly dispatchable.

The location of residential solar installations affects welfare via offset generation and emissions. In particular, the location of individual sites matters through three channels. First, the correlation between solar generation at a site and electricity demand in the region affects how much conventional generation and emissions are offset. Second, the correlation in solar generation across sites affects how intermittent the aggregate solar generation is. Lastly, the correlation between household demographics and solar generation influences where solar is installed, which affects welfare via the first two channels.

I compare a differentiated and undifferentiated subsidy as they are offered to an increasing number of zip codes. I find that offering subsidies to more zip codes reduces subsidy spending as well as intermittency costs. I also find the differentiated subsidy further reduces subsidy spending much more than it affects intermittency costs. In larger regions, such as California, a differentiated subsidy may decrease intermittency costs relative to a non-differentiated one by targeting the sunnier zip codes. When spread over a sufficiently large area, these sites are not highly correlated therefore there is not such a large increase in intermittency costs. Finally, I find that there is more to gain from differentiated subsidies in climatically and demographically heterogeneous regions. The heterogeneity contributes to more variation in households' response to differentiated subsidies, as well as more benefits to spreading installations in the region.

2 Background on Renewable Energy and the Electrical Grid

Renewable energy such as solar and wind generation, are intermittent, non-dispatchable energy sources. In the electricity market, the electricity generation must exactly meet electricity demand, or an outage may occur. This implies any drops in renewable generation must be made up for with other generation sources. In the case is solar, these drops are generally attributed to 1) the day-night cycle and 2) weather events such as cloud cover and environmental factors such as shading of solar panels by trees. The former is mitigated by simply ramping conventional generators up and down, during dusk and dawn. The latter is not perfectly forecastable, necessitating some reserve generation capacity to be available to ramp up when required.

Solar production generates more electricity during midday, when demand for electricity tends to be high. Baker et al. (2013) note that this implies the generators offset by solar tend to have higher operating costs. Households which do not face real-time prices for electricity and whose price of electricity does not include the cost of emissions[1] do not internalize these benefits of solar. Subsidies are therefore intended to account for these positive externalities.

The location of installations affects both the expected level of solar generation (insolation) as well as how intermittent the generation is. Imagine choosing where to install some level of solar capacity in a

[1] With the exception of California, this is true for carbon dioxide emissions in most of the United States.

metropolitan area. Installing it all in the sunniest region maximizes expected generation, but when a cloud passes over all of its generation severely drops. Because generation must equal demand in the electrical grid, this would require a substantial amount of reserve generation. Figure 1 plots the correlation of 15-, 30- and 60-minute changes in output for the three regions. Each point is a pair of sites with the distance between them on the x-axis and the ordinate is the correlation of their changes in output over the course of a year. Naturally, the correlation across sites decreases with distance. Therefore installing the capacity across several sites would reduce the overall variability of generation. This would then reduce the amount reserve generation needed to compensate for such drops in generation.

Several previous studies consider the impact of the location where renewable generators are installed on different aspects of the electrical grid, but to my knowledge there is only one that explicitly considers the effect of differentiated subsidies on where installations occur. Wibulpolprasert (2016) uses a dynamic structural model to consider the impact of time- and location-differentiated subsidies on the location of large-scale wind installations and the subsequent effect on emissions offset. She finds there is little to gain from differentiated subsidies compared to the status quo. A carbon tax has greater benefits, however, by incentivizing other emissions-reducing channels such as demand reduction and fuel substitution. My focus on residential solar allows me to explicitly model the households' installation decision to also account for variation in willingness to install solar across zip codes. Other studies by Cullen (2013) and Novan (2014) consider the impact of renewable wind generators on emissions offset. Using actual wind and conventional generation, they estimate the impact of different existing wind sites on emissions offset. Green and Vasilakos (2010) consider the effect of the location of wind installation on electricity price volatility and generator revenues in the United Kingdom. Wolak (2016) considers the impact of the siting of wind and solar installations on the mean and variance renewable generation, and subsequent impact on locational marginal prices. With the exception of Wibulpolprasert (2016), none of these studies consider the impact of subsidies on where installations occur.

Borenstein (2015) considers how households' electricity rates affect their decision to install solar. He finds that increasing block rate structures, in conjunction with net metering, increase the household's private benefit of solar installation for households which consume more electricity. Without household level electricity consumption data, however, puts such considerations outside the scope of this paper.

There have been substantial subsidies for solar photovoltaics (PV) over the last ten years resulting in increasing levels of solar PV adoption during this period. While some of these subsidies vary according to insolation, no existing subsidies consider the correlation in generation across sites. Existing subsidies for residential solar in the United States include a 30% tax credit from the federal government as well as incentives at the state and utility level. The federal tax credit depends solely on the installation cost, so a \$25,000 installation in Alaska receives the same incentive as a \$25,000 installation in Arizona. This is clearly inefficient.

The California Solar Initiative (CSI) is an incentive program offered by the three main investor-owned utilities in California. The CSI offered a performance-based incentive, proportional to the actual energy generated. Relatively few households, however, chose this subsidy rather than the capacity-based incentive. Furthermore, unless households are on real-time prices, even the performance-based incentive does not account for the timing of the generation nor the emissions offset. The CSI is used in previous studies to consider how subsidies affect the installation decision.[2] Hughes and Podolefsky (2015) examine how the CSI affected overall adoption in California. They find that without the CSI there would have been 53% fewer installations in California, and using a back-of-the-envelope calculation that this would have resulted in 2-3 million tons of carbon dioxide emitted over a period of 20 years. Burr (2014) models the installation decision as a dynamic discrete choice to consider how different types of subsidies affect installation decisions. She finds that cost-based incentives are at least as effective as production-based incentives for incentivizing solar adoption. She finds production-based incentives, however, are more efficient by encouraging solar installations in locations more suited for solar production.

Gowrisankaran et al. (2016) develop a structural model of a system operator managing an electric grid with an intermittent energy source. They find that overall intermittency costs are a substantial share of the cost of solar, but unforecastable intermittency is relatively insignificant. The main contributor to the cost, however, is the installation cost. With substantially lower prices for solar ($1.52) and a $39 tax on carbon emissions, however, they find solar to be welfare neutral. I combine the household's solar installation decision with the model from Gowrisankaran et al. (2016) to estimate welfare and intermittency cost under different subsidy policies.

3 Data

In this study I use two datasets which I compile from a variety of sources. I use one to estimate the solar installation decision of households and the other to estimate the welfare from the electricity market and solar's intermittency costs under counterfactual policies.

3.1 Solar Installation Decision Data

I obtain residential solar photovoltaic (PV) installations data for Arizona, California and Massachusetts from Arizona Goes Solar, California Solar Initiative (CSI), and Massachusetts Clean Energy Center (MassCEC).

[2] Gillingham and Tsvetanov (2016) models households' response to solar subsidies as a Poisson-hurdle model using residential solar installation data from Connecticut.

For each installation I observe the zip code where the PV system is installed, installation cost, subsidy received, system size installed, the date the household applied for subsidies (date of application) and the installation date. For Massachusetts I also observe whether an installation qualified for additional subsidies for having income or home value below a certain threshold. I exclude third-party owned systems because the factors influencing their installation differs substantially from household-owned systems. For California, I also observe the electric utility serving the household and the subsidy type. California offered two types of subsidies to residential households: an expected-production capacity-based subsidy and a production-based one. My analysis focuses only on the capacity-based subsidies. The timing of subsidies in California varies for the three major utilities: Pacific Gas and Electric (PGE), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E). These utilities correspond roughly to the San Francisco, Los Angeles and San Diego metropolitan areas, respectively. A summary of system size, cost per watt and incentive per watt of capacity installed and other demographics for Boston, California and Tucson is presented in Tables 1, 2, and 3, respectively. Note that Tucson has significantly less variation in solar radiation, income, home values than California and Massachusetts.

Because the finest geographic unit in the installation data is zip code, I aggregate installations to the month-zip code level. For each region I aggregate the average installation cost per watt of installed capacity across all zip codes for each month. Figure 2 shows the average (line) and standard deviation (shaded region) of the installation cost per watt of capacity in the regions; it shows a gradual decline in the installed cost per watt over time. In the estimation I use the average cost for each time period, so the variation is only across time and not zip codes. Figure 3 shows the capacity installed over time for each region; it reveals new installed capacity has increased over time in most of the regions. The large spikes in installations correspond to anticipated decreases in subsidies.[3] Figure 4 shows the subsidy per watt in the regions. I use the date of application to determine the per watt level of incentive received according to each program's subsidy schedule.[4] The variation in the subsidy within a time period in California is because the timing of decreases in the subsidy varied across the three utilities. For Boston the variation within a time period (as well as within zip codes, which is not shown) is due to additional subsidies offered for households below an income and/or home value threshold.

I also collect zip code demographics from the 2008-2012 American Community Survey (ACS) 5-Year Summary File. I use the number of bedrooms, household size, median age of homes, and the distribution of household income by home value in the zip code. Thus the variation in demographics is solely across

[3]This paper, however, focuses solely on the spatial variation so I only consider the household installation decision in a static setting.

[4]I received the subsidy schedules from representatives at Tucson Electric Power (TEP) for Tucson and MassCEC for Massachusetts.

zip codes, not over time nor within zip codes. There is an exception for Massachusetts because the policy includes an additional subsidy per watt for lower income or home value households. Therefore I split each zip code in the sample by eligibility for additional subsidy based on the joint distribution of income and home value within each zip code from the ACS.

I add each zip code's Democratic Party vote share in the 2012 presidential general election from each state's Secretary of State to these household demographics. I also use a measure of each zip code's solar insolation from the National Renewable Energy Laboratory (NREL). This provides a measure of solar irradiation even in zip codes for which I do not observe solar generation. For zip codes for which I do observe solar generation, I also compute their solar capacity factor.

3.2 Solar Generation and the Electric Grid Data

I also have a year of sub-hourly solar generation for residential and commercial sites in Arizona, California and Massachusetts. The solar generation data for Tucson, Arizona are from the the University of Arizona (UA) Physics department. The data for California are from the California Independent System Operator (CA ISO) and Yaskawa-Solectria Solar, a producer of PV inverters and monitoring software based in Massachusetts. The solar generation data for Massachusetts are from Yaskawa-Solectria Solar. The data from Yaskawa-Solectria contain the capacity of panels installed, zip code of installation, and whether each installation's generation is recorded as alternating (AC) or direct current (DC). These are minute-level power data for each inverter at a site, which are aggregated for each site and converted to AC. The solar generation data from the CA ISO and UA Physics are fifteen-minute AC power data and contain both the capacity installed and the zip code of installation for each site. All of the CA ISO data are from residential installations while the data from Yaskawa-Solectria and UA Physics include commercial and residential sites.

The site-level power data are divided by the installed capacity to normalize it to be a measure of output as a fraction of its rated capacity between zero and one. To get hourly zip-code level output, I average the hourly output across all sites in each zip code. The generation data are aggregated to the hourly zip code level because the zip code is the finest geographic unit in the solar installation data and the hour is the finest temporal unit for the load (electricity demand) data.

In addition to the residential solar generation, I have hourly load for the CA ISO, NE ISO and TEP from the Federal Energy Regulatory Commission (FERC) Form 714. I also have hourly generation for utility scale solar and wind generation for California and Massachusetts from the CA ISO Renewables Watch and the New England Independent System Operator (NE ISO) Operations Report, respectively. These data also include other (arguably) non-dispatchable generation including run-of-river hydroelectric, geothermal, nuclear for California, and biofuel. I subtract the hourly generation from these sources from the hourly load. In Arizona and Massachusetts, where I do not observe hourly nuclear generation, I treat nuclear generating capacity as always on at full capacity – subtracting its rated capacity from the hourly load for all hours.

For each region I compile a list of dispatchable generators from the United States Environmental Protection Agency (EPA) eGRID2012 data file. These data include the generators' rated capacity, heat rate, and emissions. To calculate each plant's marginal cost of operation I use the emissions and heat rate, together with carbon dioxide (CO_2), sulfur dioxide (SO_2), and nitrogen oxides (NO_x) permit futures prices from Bloomberg and fuel costs from the United States Energy Information Administration (EIA) Form EIA-923. Each generator's fuel costs are based on either the actual costs for that plant, based on its plant identification number (if available), or the average cost for that fuel type in the region.

The probability of unscheduled outages and maintenance for each fuel type is from the North American Electric Reliability Corporation (NERC) Generating Availability Data System (GADS). I also obtain day ahead forecasts from the National Oceanic and Atmospheric Administration (NOAA) FOUS5 forecasts for Boston, Los Angeles, San Diego, San Francisco and Tucson. Sunrise and sunset data are from Astronomical Applications Department of the U.S. Naval Observatory for the same cities. These forecasts are used by the system operator to predict hourly load and solar generation.

4 Estimation and Methodology

In order to determine the costs and benefits of geographically-differentiated subsidies I first estimate how subsidies affect the decision to install a solar PV system. This is modeled as two independent choices: a discrete choice of whether or not to install solar and a choice of the size of solar panels to install conditional on installation. The estimation procedure and identification are described in section 4.1. I use these parameters to determine the geographic distribution of installations under counterfactual subsidies. I use these counterfactual policies as an input to the structural model of an electricity system operator from Gowrisankaran et al. (2016). Section 4.2 describes the model and how altering the distribution of installations affects the estimated welfare. I use this model to estimate electricity market welfare and intermittency costs under the different counterfactual subsidy policies.

4.1 Solar Installation Decision

Let u_{ijt} be the indirect utility of a household in zip code i choosing choice j in period t. If the household installs solar ($j = 1$) it receives indirect utility

$$u_{i1t} = \beta^p(CPW_t - IPW_{it}) + x_{it}\beta + \varepsilon_{i1t}$$

and

$$u_{i0t} = \varepsilon_{i0t}$$

if it does not ($j = 0$). The price coefficient, β^p, captures the effect of the net price per watt: the price per watt, CPW_t, which varies over time, minus the incentive per watt, IPW_{it}, which varies over time and, in some cases, across zip codes as well.[5] The parameter vector β captures the effect of the remaining covariates denoted x_{it}. These include the time-invariant demographics for the zip code such as the average household income, home value, number of bedrooms, household size, democrat vote share and irradiation. It may also include year fixed effects and polynomial time trends. Let $X_{it} \equiv (CPW_t, IPW_{it}, x_{it})$ denote all covariates. The household's unobservable components of utility, ε_{ijt}, is assumed to be drawn from a type 1 extreme value (EV1) distribution and is independently and identically distributed across households, choices and time periods.

A household installs solar panels if its indirect utility from installing is greater than that of not installing ($u_{i0t} < u_{i1t}$). The EV1 errors imply that the probability household i decides to install solar panels in period t ($d_{it} = 1$) is

$$P(d_{it} = 1|X_{it}) = \frac{\exp\Big(\beta^p(CPW_t - IPW_{it}) + x_{it}\beta\Big)}{1 + \exp\Big(\beta^p(CPW_t - IPW_{it}) + x_{it}\beta\Big)} \equiv P_{i1t}, \tag{1}$$

and the probability household i does not install solar in time period t is simply $P_{i0t} = 1 - P_{i1t}$.

I estimate the model using a maximum-likelihood estimator which maximizes the log-likelihood function

$$L(\theta) = \sum_{i}^{N} \sum_{t}^{T} \Big[N_{i1t} \log\Big(P_{i1t}\Big) + N_{i0t} \log\Big(P_{i0t}\Big)\Big],$$

where i denotes zip codes, N is the number of zip codes in the sample, t denotes time periods, T is the number of time periods. N_{i1t} is the number of households who install solar panels in zip code i in period t, and N_{i0t} is the number of households who do not. The probability of installing solar, P_{i1t}, is as defined in (1) and $P_{i0t} = 1 - P_{i1t}$.

[5] In a slight abuse of notation, for California and Massachusetts i may also denote subsets of a zip code which receive different levels of subsidies. In California this is the case when a zip code is served by multiple utilities. In Massachusetts this is due to different levels of subsidy for households in different income and home value brackets.

To estimate the households' choice of size, I regress the average system size installed in zip code i, month t, S_{it}, on the net price per watt and demographics including the year fixed effects and time trend:

$$S_{it} = \gamma^p (CPW_t - IPW_{it}) + x_{it}\gamma + \eta_{it}. \tag{2}$$

This is estimated using ordinary least squares. The parameter γ^p dictates how the counterfactual choice of size of panels, $\widetilde{S}_{it}$, responds to a counterfactual subsidy policy, $\widetilde{IPW}_i$.

Combining the choice of installation and size in the predictions assumes the two choices are independent – a fairly strong assumption. I check the sensitivity of this assumption by modeling the decision as a standard censored Tobit model. This model implies perfect correlation in the error terms for the two decisions, in addition to the normal rather than EV1 distributional assumption on the error terms.[6] The total capacity installed in zip code i in month t is

$$S^*_{it} = \delta^p (CPW_t - IPW_{it}) + x_{it}\delta^T + \nu_{it},$$

where $\nu_{it}|X_{it} \sim N(0, \sigma^2)$. As in (1) and (2), x_{it} includes the various zip code demographics, year fixed effects and a polynomial time trend. In the data, I only observe the size in zip codes where the choice of capacity to install is positive:

$$\breve{S}_{it} = \begin{cases} S^*_{it} & \text{if } S^*_{it} \geq 0 \\ 0 & \text{if } S^*_{it} < 0. \end{cases}$$

4.1.1 Identification of the Installation Decision Models

My goal is to determine whether geographically differentiated solar subsidies could improve welfare, so it is paramount to identify and have an unbiased estimate for the coefficients on net price: β^p and γ^p. Because the cost per watt, CPW_t, is an average across zip codes for each time period, it varies solely across time periods. The incentive per watt, IPW, is the scheduled subsidy for each region. Therefore the variation in subsidy per watt and net cost per watt is only over time for Tucson. Thus the coefficient on net price for Tucson is identified solely off the variation in the price and subsidies over time and how the number of installations or system size installed varies in response to these changes.

When I estimate the model for California, I have variation in the subsidy within time periods because the three major utilities' subsidy policies are not identical. Massachusetts has similar variation within time periods, as well as within zip codes, because the subsidy is greater for households with lower income and/or

[6] It's also possible to correct for the correlation in the choices using a Heckman correction or to estimate the correlation using the model outlined in Jeffrey A. Dubin (1984).

home values. Thus in both California and Massachusetts I use the variation in subsidies across these groups and over time together with the variation in price across time periods to identify the net price coefficients.

To address concerns of endogeneity through reverse causality – e.g. a drop in price due to technological change driven by increased installations in the region – I include year fixed effects and a polynomial time trend. These capture the variation in installations and choice of system size driven by technological changes and changes in demand for solar panels which would arguably affect all zip codes in a time period equally. Because I focus on residential installations, which tend to be much smaller than commercial or utility-scale solar installations, the general equilibrium effect on panel prices should be limited.

4.2 Welfare from the Electricity Market

To calculate welfare from the electricity, I combine the previously described models of the installation decision with a structural model of an electrical grid system operator described adapted from Gowrisankaran et al. (2016). In the model, a system operator schedules generators and curtails demand to meet hourly load, given a distribution of solar installations (and their generation). The system operator makes her decision in two stages. She initially chooses the number of combined cycle natural gas generators to install and a curtailment price, p_c, to offer to customers, who may have their demand curtailed. The model has a finite time horizon, so both new and existing generators, as well as the solar installations, last a fixed number of years. In the second stage, in each period the system operator decides the level of electricity demand curtailment and scheduling of generators after observing the available generators and a forecast for demand and solar generation.

4.2.1 The Impact of the Geographic Distribution of Solar on the Electricity Market

The location of residential solar installations affects social welfare via offset generation and emissions. In particular, the location of individual sites matters through three channels: 1) correlation between solar generation at a site and demand for electricity in the region, 2) correlation in solar generation across sites, and 3) how household demographics, which influence the solar installation decision, correlate with solar generation. Due to lack of data, I am unable to consider geographic heterogeneity in the local distribution grid, which is potentially an additional source of geographic heterogeneity in the social benefit of solar.[7]

First, the solar generation at different locations may differentially correlate with the aggregate demand for electricity. This may be due to local geography and microclimates. For example, installations on a hillside may reduce generation during dusk when demand is high, or a marine layer of cloud cover may reduce solar

[7] Cohen et al. (2016) finds considerable heterogeneity in the value of PV for deferring investment in the distribution network in Pacific Gas & Electric's territory in California.

generation during summer for installations in coastal California. Vegetation or the azimuth and tilt of a system on a roof at individual sites, which reduce generation or alter its timing, are outside the scope of this paper. Such factors are only pertinent in the current setting if they are representative of the area, as is the case in a heavily forested region. The timing of the solar generation and its correlation with demand affect social welfare via the conventional generation and corresponding emissions it offsets.

The second factor which affects the social benefit of solar at a site is the correlation of its solar generation with other sites. For instance consider two zip codes that are situated on the east and west side of a mountain. The decrease in generation during dusk and dawn at one is offset by increased generation at the other, reducing the variability of the total solar generation.

The third factor is the demographics of the households installing solar. When a certain region is more price sensitive and all else is equal, a price-discriminating subsidy favors such a region because each dollar of subsidy results in a greater amount of capacity installed.

The benefit of differentiating subsidies also depends on the regions' characteristics including the portfolio of generators and their characteristics, demand for electricity and weather. This is in a similar vein to Holland et al. (2015) who compare the benefits of electric vehicles to the cost of their subsidies across regions. However, in the case of solar it is insolation across installation sites *within a region*, not only differences in generation portfolios across regions, which drives the spatial variation in the marginal benefits of solar installations. The variation in generator characteristics across regions matters only when I consider differences in the gains from spatially differentiating subsidies across regions.

4.2.2 Overview of the System Operator's Problem

The model assumes a constant retail price of electricity, $\bar{p}$, and demand for electricity is modeled by a constant elasticity of demand function for prices below a reservation value, v,

$$Q^D(\bar{p}, D_h) = \begin{cases} 0 \text{ for } \bar{p} > v, \text{ and} \\ D_h p^{-\eta} \text{ for } \bar{p} \leq v, \end{cases} \tag{3}$$

where η is the elasticity of demand and the scale of forecasted demand in hour h, D_h, is a random variable which depends on the weather forecast, $\vec{w}_h$, and is distributed according to $F^D(D_h|\vec{w}_h)$. The system operator's choice of curtailment price, p_c, in the first stage determines the maximum amount of demand curtailment, z_h, she can implement in any period h because only customers whose valuations are below p_c sign up for interruptible power contracts.[8]

[8] As noted in Gowrisankaran et al. (2016), the effect of an increase in the curtailment price on welfare is ambiguous. An increase in p_c decreases the probability of a system outage, by allowing more demand to be curtailed, but also increases the

The operator has an amount of solar PV capacity, K^{SL}, distributed across N zip codes which depend on the subsidies offered and the installation decision model in section 4.1. The total distributed residential solar generation in a period is the capacity multiplied by the draw for forecasted solar output: $K^{SL}S_h^{SL}$. The draw of forecasted solar output is S_h^{SL} and is distributed as $F^S(S_h^{SL}|\vec{w}_h)$. More precisely, the draw of residential solar output, S_h, for hour h is a weighted average of the solar generation in each zip code for that hour, S_{ih}^{SL}, weighted by the share of capacity installed in each zip code, $\frac{K_i^{SL}}{K^{SL}}$:

$$S_h^{SL} = \frac{1}{K^{SL}} \sum_{i=1}^{N} \left(K_i^{SL} \times S_{ih}^{SL} \right).$$

Therefore changing the capacity installed in the zip codes, $\{K_i^{SL}\}_{i=1}^N$ via counterfactual subsidies changes the distribution of forecasted solar generation conditional on the forecast, $F^S(S_h^{SL}|\vec{w}_h)$. Changing the system operator's forecast of solar impacts her scheduling of conventional generation.

There are G existing generators. Each generator g has probabilities of failure and maintenance, P_{gh}^{fail} and P_{gh}^{maint}, which are identically and independently distributed across generators, g and periods, h. The maintenance status of generator g in period h, denoted $m_{gh} \in \{0,1\}$, determines whether the generator is available for the system operator to schedule for generation. The system operator's decision to schedule generator g in hour h is $on_{gh} \in \{0,1\}$. Thus, if a generator is scheduled for maintenance, $m_{gh} = 1$, then it is not scheduled for production, $on_{gh} = 0$, for that period. All generators may be used for generation and operating reserves. Each has a constant marginal cost, c_g, and capacity, K_g. The marginal cost of operating reserves is a fraction c^r of the cost of generation for the portion of the generator's capacity that is used for operating reserves. The system operator chooses the number of new combined cycle gas generators to install, n^{FF}. Each of these new units has the same maintenance and failure probabilities as existing gas generators. Each is assumed to have the same generating capacity, k^{FF}, cost of investment per MW of capacity, AFC^{FF}, and marginal cost of operation, c^{FF}.

As residential solar generates electricity during peak demand, this reduces the discounted present value of investment in new transmission lines and their maintenance according to,

$$TFC(K^{SL}) = AFC^T \max_{\vec{w}_h} \mathbb{E}[D_h(\vec{w}_h)\bar{p}^{-\eta} - K^{SL}S_h^{SL}(\vec{w}_h)], \tag{4}$$

where AFC^T is the average fixed cost of transmission. Solar and demand curtailment in period h reduce

average valuation of a curtailed user.

line losses in transmission lines, according to

$$LL(Q_h) = \frac{1 - 2\alpha Q_h - \sqrt{1 - 4\alpha Q_h}}{2\alpha}, \tag{5}$$

where α is a constant in the specification of line losses[9] and Q_h is the electricity demand in a period net of distributed solar generation and demand curtailment:

$$Q_h \equiv D_h \bar{p}^{-\eta} - z_h - K^{SL} S_h^{SL}.$$

Given a fixed retail electricity price, $\bar{p}$, capacity of solar installations, K^{SL}, the system operator chooses the number of new fossil fuel plants to build, n^{FF}, and a curtailment price, p_c, in the first stage. In each period h of the second stage, upon observing the weather forecast and maintenance statuses of generators, she schedules generators, $\vec{on}_h$, and the level of demand curtailment, z_h. The period's realizations of generator failures, P_{gh}^{fail}, and the draws of forecasted solar generation and scale of demand for electricity, S_h^{SL} and D_h, are realized. Generator g's available capacity in hour h, x_{gh}, after generator failures are realized is equal to the generator's capacity, $x_{gh} = K_g$, if the generator did not fail, and is equal to 0 otherwise. If the available generators' capacity plus the realization of solar generation is *less than* the realization of demand minus demand curtailment plus line losses in a period h, an outage occurs:

$$outage(\vec{on}_h, z_h, \vec{w}_h) = \mathbb{1}\left\{ \sum_{g=1}^{G+n^{FF}} x_{gh} + K^{SL} S_h^{SL} < D_h \bar{p}^{-\eta} - z_h + LL(Q_h) \right\}. \tag{6}$$

In each period h, the system operator's second stage payoff is

$$\begin{aligned} W(\vec{w}_h, \vec{m}_h | n^{FF}, p_c) = \max_{\vec{on}_h, z_h} \mathbb{E}\Big[& \left(1 - d^{outage} outage(\vec{on}_h, z_h, \vec{w}_h)\right)\left(D_h \bar{p} VOLL - WLC(z_h, p_c)\right) \\ & - C\left(Q_h + LL(Q_h), \vec{x}_h\right) | \vec{w}_h, \vec{m}_h \Big], \end{aligned} \tag{7}$$

with the expectation being over the random variables D_h and S_h^{SL} and the probabilities of failure at each plant, P_{gh}^{fail}. The expectation of the fraction of customers who lose power times the number of periods for which they lose power, conditional on there being an outage, is is d^{outage}. $C(\cdot)$ denotes the production costs of generation and operating reserves when generators are operated in the order of increasing marginal costs. The highest marginal cost generators which were scheduled to be on but are not needed are left as reserves. The cost of this is a fraction, c^r, of their operating cost, c_g, for the portion of their capacity which

[9] Line losses are assumed to satisfy the relationship, $LL = \alpha(Q + LL)$, and (5) is the root at which welfare is maximized. See Gowrisankaran et al. (2016) for more details.

is not required. The value of lost load (VOLL) is the mean value of a unit of electricity for consumers. Gowrisankaran et al. (2016) show that there exists a closed-form relationship between the reservation value, v, and the value of lost load (VOLL). They also show the welfare loss from curtailment, $WLC(z_h, p_c)$, in any second stage period from a level of curtailment, z_h, given the curtailment price, p_c, has a closed form solution and is independent of D_h.

The system operator, given n^{FF} new generators, chooses a curtailment price to maximize the expectation of $W(\cdot)$ in (7) over all periods in a year, resulting in the value function

$$V(n^{FF}) = \max_{p_c} \mathbb{E}\left[\sum_{h=1}^{H} W\left(\vec{w}_h, \vec{m}_h | n^{FF}, p_c\right)\right]. \tag{8}$$

The value from the optimal choice of investment in new generators is

$$V^* = \max_{n^{FF}} \left\{V\left(n^{FF}\right) - K^{SL} AFC^{SL} - n^{FF} AFC^{FF} - TFC\left(K^{SL}\right)\right\}, \tag{9}$$

where AFC^{SL} is the average installation cost of residential solar per unit of capacity installed.

4.2.3 Intermittency Costs

To measure the benefit of differentiated subsidies I consider how counterfactual subsidies affect subsidy spending and the intermittency costs. Following Gowrisankaran et al. (2016), I compute the intermittency cost as the difference in total welfare, given by (9), between the baseline model described in the previous section and one in which solar is perfectly dispatchable. When solar is perfectly dispatchable, the system operator allocates solar generation from the hour in which solar generation is greatest to the hour with the highest electricity demand, and so on in decreasing order. This is analogous to using an ideal battery to store the solar energy for the period in which demand is greatest and therefore production costs highest. Because generators are dispatched in order of increasing marginal costs, dispatching solar in this manner reduces generation costs the most. This measure of intermittency costs includes both the forecastable component of intermittency – which is due to the day-night cycle and expected weather conditions – as well as the its unforecastable component – due to short-term, unforecasted fluctuations in weather and other conditions which may affect solar generation. I consider total intermittency costs because Gowrisankaran et al. (2016) find the unforecastable component of intermittency to be quite small.

5 Results and Discussion

I first present the results of the demand for solar installations described in Section 4.1. Next, I present simulations showing the impact of the correlation across sites and the geographic distribution of installations on intermittency costs using the model of the electricity market. Finally I combine the installation decision with the model of the system operator managing the regional electric grid to consider the impact of different types of subsidy policies on subsidy spending and intermittency costs.

5.1 Estimation of the Solar Installation Decision

I model the installation decision as two independence choices for households: a decision of whether or not to install solar and a choice of system size to install, conditional on installation. Table 4 presents the estimation results for the simple logit model of whether or not to install solar. The net cost per watt is the average cost per watt (of capacity installed) minus the subsidy per watt. The cost per watt is an average across all zip codes for a given time period; thus it only varies over time. The subsidy per watt varies over time for all three regions and for Tucson it varies solely over time. For California the subsidy per watt also varies across zip codes based on electric utility. For Massachusetts, it varies within zip codes according to the distribution of income and home values within a zip code. Consistent with intuition, the coefficient on net cost per watt is negative so an that increase in subsidy leads to more installations.

Installations are increasing in the Democratic Party vote share, which serves as a proxy for green preferences. Wealthier households are more likely to install solar panels. The effect of household size varies across the three regions. This is because household size acts as a proxy for electricity demand, but also greater household size reduces income per member of household.

The effect of the homes age also varies for the three regions. Installations are more likely in newer homes in Boston and California, while the reverse is true for Tucson. The former may be the case because individuals who are likely to install solar are also concerned about energy efficiency, and hence tend to live in newer homes. For Tucson, the sign of the coefficient for age of home and home value may be due to retirees tending to live in older, more expensive homes. They are perhaps less likely to install solar because their remaining life expectancy is less than the life span of the PV system, thus they would not expect to receive the full return of the investment. Additionally, many retirees leave Tucson during summer months when the benefit of avoided electricity expenditures from solar are greatest. Meshing with intuition, households in more expensive homes are more likely to install solar in California and Boston.

I also control for the price of electricity, which is most important in California because the three major utilities there changed the prices in their block rates several times during the sample period.[10] While both California and Arizona have increasing block pricing for electricity in which households who consume more electricity pay a higher marginal price, the difference in prices across blocks is much more extreme in California. I also use NREL's annual median measure of solar radiation to control for the sunniness of different zip codes.

Table 5 presents the estimation results for the choice of PV system size conditional on installation for the three regions. The results do not differ drastically from the results for the simple logit model. Higher system costs reduce the system size installed, while higher income households intuitively install larger panels. The system size, however, is decreasing in Democratic Party vote share. This is likely due to the higher share of Democrats residing in denser city centers, with restricted rooftop space. Older households tend to install smaller panels, which may be due to the size of homes increasing over time or standardization of roof designs facilitating larger installation of contiguous panels on newer homes. The effect of bedrooms is positive for Boston – dovetailing with bedrooms being a proxy for electricity use – and the sign of the coefficient on household size varies. Controlling for electricity prices is important in California as previously discussed, but not so for Boston and Tucson.

Figure 5 shows the predicted and actual capacity installed, where each point is a zip code. I sum the capacity across all time periods for each zip code, given the results of the installation decision and choice of size conditional on installation presented in Tables 4 and 5. Figures **??**, **??**, and **??** show maps of how actual and predicted installations are dispersed across zip codes in Boston, California and Tucson, respectively. While the predictions do not exactly match the actual installations, they are not too far off. Another aspect worth noting is that installations in California are predominantly in coastal southern California. In the simulations I show how these particular sites are highly correlated, which results in fewer benefits from spreading solar across more sites.

For the results presented in Tables 4 and 5, I used the full sample of installation data. This includes zip codes for which I do not observe actual solar generation. Without observing solar generation, I am unable to consider these zip codes' contribution to intermittency costs in the simulations. Tables A2 through A7 in the appendix present the results for the subsample of zip codes for which I observe generation, using both the actual generation data to compute capacity factors (column 1) and NREL's measure of solar radiation (column 2). The third column again presents the results for the full sample. The results do not change very much across the three samples. One exception is that when using the capacity factor calculated using the actual generation data, I find that sunnier zip codes are more likely to install and more likely to install larger panels. This may suggest that NREL's measure of insolation is not quite as precise on such a fine level.

[10] I thank Jonathan Hughes for this suggestion.

I also consider a standard censored Tobit specification, presented in Table A8. In contrast to my assumption of independence of the decision of choice of size and whether or not to install, the Tobit model implicitly assumes that these decisions are governed by the same equation. This combines the two previous models into one, making direct comparison of coefficients is not straightforward. The results for the main coefficient of interest, net cost per watt, however are qualitatively similar.

5.2 Welfare Impacts of Differentiated Subsidies

I compare the subsidy and intermittency costs of differentiated and non-differentiated subsidies by combining the installation decision results with the model of an electric grid system operator from Gowrisankaran et al. (2016), as discussed in section 4.2. For a given distribution of solar, I use this model to estimate total welfare from the electricity market. To estimate the intermittency cost, I compute the difference in estimated welfare between perfectly dispatchable solar and the baseline model.

5.2.1 Correlation, Spreading Installations, and Intermittency Costs

Recall Figure 1 which shows how the correlation in solar generation drops with the distance between sites. Intuitively, decreasing the correlation across sites should reduce the intermittency costs, but to what extent this is true is an empirical question. To address this question, consider two extreme cases: one in which the solar generation is perfectly correlated (i.e. it is all installed in one site) and another in which solar generation takes the form of i.i.d. draws. To generate such (clearly nonexistent) i.i.d. solar generation data, I first estimate the mean and standard deviation of a normal distribution for each daylight time period[11] in a year using a kernel density estimator. I then take either one draw to use for all N zip codes (for perfect correlation, $\rho = 1$) or N i.i.d. draws from each time period's estimated normal distribution for the entire year.

I set the total capacity for each region so that annual solar generation corresponds to 10% of the demand for electricity, i.e. a 10% renewable portfolio standard (RPS), and $1/N$ of this capacity is installed at each zip code. Table 6 shows the total intermittency costs (for the entire twenty five years of the system operator's model) and the intermittency cost per megawatt-hour of solar generation, using a discount rate of 6%. By having N i.i.d. draws rather than a single draw, intermittency costs decrease by approximately \$9.6, \$7.5 and \$4.5 per MWh for Boston, California and Tucson, respectively. The smaller reduction in intermittency costs in Tucson, Arizona are a result of little climatic variation within the region. This suggests intermittency costs are not trivial.

[11] Recall the solar generation for Boston is at the minute level and for California and Tucson it is at fifteen minute intervals.

Before considering the role of subsidies and their impact on the installation decision, it is also worth considering to what extent spreading installations reduces intermittency costs. As seen in Figure 1, spreading installations reduces correlation, which should reduce intermittency costs. Consider taking the same 10% RPS capacity and either installing it in the sunniest zip code or spreading it equally across the N zip codes. Using the observed solar generation (rather than the i.i.d. generation) data, I estimate intermittency costs using the system operator's model. Because installing solar in the sunniest site significantly increases the solar generation, I compute intermittency costs per watt of solar capacity installed. Table 7 shows that spreading installations across the N sites rather than installing all the capacity in the sunniest location decreases intermittency costs by \$0.36, \$0.42, and \$0.24 per watt in Boston, California and Tucson, respectively. Again there is more to gain from spreading installations to reduce intermittency costs in Boston and California, than in climatically-homogeneous Tucson. An important caveat, however, is the cost of achieving these distributions of installations is not accounted for because I have not yet incorporated the installation decision.

5.2.2 Subsidy Policies and Intermittency Costs

The previous simulations only considered the gains from reducing intermittency without considering the subsidy costs needed to incentivize the installations. I now consider how both subsidy costs and intermittency costs vary when I account for the installation decision. I consider two types of subsidies: one which is the same for all zip codes to which subsidies are offered and another which offers a different subsidy to each zip code. I consider how intermittency costs and subsidy spending vary as the number of zip codes to which subsidies are offered, $\breve{N}$, increases. When deciding to which zip codes to offer subsidies, I assume the regulator prefers to incentivize installations in the sunniest zip codes. Furthermore, I assume the regulator has a target for solar generation, which is based on a 10% RPS as in the previous simulations. For the non-differentiated subsidy, the regulator first determines how much capacity is installed in the $N - \breve{N}$ least sunny zip codes. She then computes the subsidy, $\widetilde{IPW}$, to offer the $\breve{N}$ sunniest zip codes which meets the 10% RPS target. For the differentiated subsidy the first step is identical: the regulator determines the level of generation under a subsidy of zero in the $N - \breve{N}$ least sunny zip codes. She then minimizes total subsidy spending by choosing the subsidy, $\widetilde{IPW}_i$, to offer each zip code i of the $\breve{N}$ sunniest zip codes, which attains the 10% RPS.

I abstract away from the time dimension by using the a reference time period for each city,[12] and using the solar prices in that month. As a baseline, similar to Table 7, I estimate the intermittency costs under a policy which installs the same capacity in the $\breve{N}$ zip codes to satisfy the 10% target. My main simulations of interest, however, are computing subsidy spending and estimated intermittency costs for different values

[12]The baseline month is January 2014 for California, and June 2014 for Boston and Tucson.

of $\breve{N}$ under a differentiated and non-differentiated subsidy.

It is important to note that as I install solar in less sunny zip codes (as $\breve{N}$ increases), I need to install more solar capacity because I constrain solar generation to meet a 10% RPS. Therefore I present the costs in Figure 6 as the intermittency and subsidy costs per megawatt-hour ($/MWh) of solar generation.

Figure 6 shows how intermittency costs (left panel) and subsidy spending (right panel) vary with $\breve{N}$ under the different policies. The solid (red) line with circles is the uniform, non-differentiated subsidy and the dotted (green) line with triangles is the differentiated subsidy. The dashed (blue) line with squares is the policy which spreads the solar capacity equally across the $\breve{N}$ zip codes, ignoring the installation decision.

First consider how intermittency costs vary with $\breve{N}$, shown in the left panel of Figure 6. Based on the previous simulations, we expect the policy which spreads solar across the $\breve{N}$ zip codes to have the lowest intermittency costs, and these intermittency costs should decrease with $\breve{N}$. Such is the case in Boston and Tucson, but not in California. For Boston and Tucson, there is a sharp initial decline in intermittency costs, after which they tend to level off. In contrast, there is little benefit to intermittency costs from spreading installations in California because of the high correlation in the sites in coastal southern California. What differentiates California is both its size and climatic variation. Figure 7 shows how installations are distributed across zip codes in California under the differentiated subsidy for several values of $\breve{N}$. Increasing $\breve{N}$ from 1 to 16 places installations mostly throughout inland southern California. Because the sites are spread sufficiently far apart, intermittency costs from clustering are not evident – therefore we see intermittency costs fall significantly. However, as we go from $\breve{N} = 16$ to $\breve{N} = 32$, we add more installations along coastal southern California. These areas experience low cloud cover caused by marine layer clouds from May to September. Furthermore, the marine layer clouds often affect all of coastal southern California, resulting in a high level of correlation even across distant sites. This is evident in the large spread in correlation for any given distance for California Figure 1. As a result, after the initial gains from increasing $\breve{N}$ from 1 to 16, spreading installations across more zip codes in California does not significantly reduce correlation. Eventually, as installations are spread to northern California ($\breve{N} = 48$ and $\breve{N} = 64$), this reduces the correlation across sites, slightly reducing intermittency costs.

Tables A9 through A17 in the appendix show how capacity, solar generation, investment in new NGCC plants, intermittency costs and subsidy costs vary with $\breve{N}$ under the different policies for the three regions. Here the intermittency costs is annualized, so the intermittency cost per megawatt-hour is simply this cost divided by the annual solar generation. To calculate the subsidy costs per megawatt-hour, one must discount solar generation over the life time of the system. Considering the reduction in total costs, there are substantial reductions in intermittency costs from increasing $\breve{N}$ from 1 to N in all three regions. The reduction is about 4-6% in Boston, 2% in California and 5% in Tucson. These reductions in intermittency costs, however, are dwarfed by the reduction in subsidy spending which range from 15-45%.

Now consider the subsidy costs in the right panels of Figure 6. As with intermittency costs, the solid (red) line with circles denotes the non-differentiated subsidy costs and the dotted (green) line with triangles is the differentiated subsidy costs. Offering the subsidy to more zip codes decreases subsidy spending because it captures more of the households who place a high value on installing solar. These savings dominate the increased cost of installing more capacity in the less sunny zip codes to meet the 10% RPS. The reduction in subsidy with $\breve{N}$ is most pronounced in Boston and Tucson.

The non-differentiated subsidy spending line is always below than the line for the differentiated subsidy cost because the latter takes advantage of both the differences in the generation potential and the households' value for solar across zip codes. Thus the differentiated subsidy is able to target both sunnier zip codes and those which contain more high value households who will install in response to a lesser subsidy. There is very little difference between the differentiated and undifferentiated subsidy costs for Tucson. This is because Tucson is fairly homogeneous in insolation and demographics – its standard deviations for insolation, income and home value in Table 3 are significantly less than those for Boston and California.

Next, consider the difference in intermittency costs between differentiated and non-differentiated subsidies. Due to the lack of climatic variation, the difference between the two policies is insignificant for Tucson. In Boston the differentiated subsidy has unequivocally higher intermittency costs than the undifferentiated subsidy. This is likely due to the concentration of sites in a particularly sunny area, resulting in high correlation across these sites and exacerbating the intermittency costs. In contrast, California's differentiated subsidy has slightly lower intermittency costs than the non-differentiated one. This is due to the location of the inframarginal households responding to the non-differentiated subsidy. These households tend to be in the affluent coastal areas, which have more intermittent solar generation caused by the marine layer clouds. By targeting zip codes with high insolation, the differentiated subsidies offer less incentive to these areas with low insolation but a high propensity to install. Furthermore, because installations are spread across a larger area, the increased intermittency costs due to high correlation across the sunny sites, as seen in Boston, is avoided.

Figure 6 and Tables A9, A10, A12, A13, A15, and A16 clearly show that the decrease in subsidy spending dwarfs any changes in the intermittency costs. The difference in intermittency cost between differentiated and undifferentiated subsidies is several orders of magnitude less than the difference in subsidy costs for all three regions.[13] Additionally, the decrease in subsidy costs from offering subsidies to more zip codes is much greater in magnitude than the change in intermittency costs. Thus by targeting the sunnier zip codes, the

[13]Note however that in the tables the intermittency cost is an annualized figure while the subsidy costs are for the PV systems' 25 year lifespan.

lower cost of the differentiated subsidy more than compensates any increase in intermittency costs.

6 Conclusion

I find that the correlation of solar across sites impacts intermittency costs, and that these costs may be mitigated by spreading installations across more sites. Not accounting for the installation decision, there is a 5-20% decrease in intermittency costs from spreading installations across all sites rather than installing all of the capacity in the most productive site.

Accounting for the installation decision, there are gains from offering subsidies to a broader area via reduced subsidy spending and to a lesser extent decreased intermittency. The decrease in intermittency costs from offering subsidies to all zip codes rather than the sunniest zip code is roughly 5% in Boston and Tucson and 1% in California. In contrast, the decrease in subsidy spending is between 15-45%. Interestingly, intermittency costs in California do not monotonically decrease as subsidies are offered to more zip codes – as they do in Boston and Tucson. After an initial decrease in intermittency costs from spreading installations to the sunniest sites across inland California, adding more of the less sunny, coastal sites increases the intermittency costs. This is due to clustering of installations in coastal regions in southern California. These areas have more intermittent generation, which is highly correlated due to marine layer clouds during summer. The propensity for these clouds to form during summer is significant because this corresponds to peak electricity demand in California.

By targeting areas with more sun, the differentiated subsidy makes it possible to meet a 10% RPS with less installed capacity, driving down the subsidy costs. However, by targeting sunnier areas, the differentiated subsidy also results in higher intermittency costs than the non-differentiated subsidy in Boston and Tucson. In California the reverse is true because the inframarginal, coastal households who receive a higher subsidy under the non-differentiated policy have more intermittent and highly correlated solar generation. This suggests that by allowing households to choose capacity- rather than production-based subsidies, the California Solar Initiative may have over-incentivized installations in the coastal regions.

References

Baker, E., Fowlie, M., Lemoine, D., and Reynolds, S. S. (2013). The economics of solar electricity. *Annual Review of Resource Economics*, 5(1):387–426.

Borenstein, S. (2015). The private net benefits of residential solar pv: The role of electricity tariffs, tax incentives and rebates. Working Paper 21342, National Bureau of Economic Research.

Burr, C. (2014). Subsidies, tariffs and investments in the solar power market. Working paper.

Cohen, M., Kauzmann, P., and Callaway, D. (2016). Effects of distributed {PV} generation on californias distribution system, part 2: Economic analysis. *Solar Energy*, 128:139 – 152. Special issue: Progress in Solar Energy.

Cullen, J. (2013). Measuring the environmental benefits of wind-generated electricity. *American Economic Journal: Economic Policy*, 5(4):107–33.

Gillingham, K. and Tsvetanov, T. (2016). Hurdles and steps: Estimating demand for solar photovoltaics. Working paper.

Gowrisankaran, G., Reynolds, S. S., and Samano, M. (2016). Intermittency and the value of renewable energy. *Journal of Political Economy*, 124(4):1187–1234.

Green, R. and Vasilakos, N. (2010). Market behaviour with large amounts of intermittent generation. *Energy Policy*, 38(7):3211 – 3220. Large-scale wind power in electricity markets with Regular Papers.

Holland, S. P., Mansur, E. T., Muller, N. Z., and Yates, A. J. (2015). Environmental benefits from driving electric vehicles? Working Paper 21291, National Bureau of Economic Research.

Hughes, J. E. and Podolefsky, M. (2015). Getting green with solar subsidies: Evidence from the California solar initiative. Working paper.

Jeffrey A. Dubin, D. L. M. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52(2):345–362.

Novan, K. (2014). Valuing the wind: Renewable energy policies and air pollution avoided. Working paper.

Wibulpolprasert, W. (2016). Optimal environmental policies and renewable energy investment in electricity markets. Forthcoming in Climate Change Economics.

Wolak, F. A. (2016). Level versus variability trade-offs in wind and solar generation investments: The case of california. Working Paper 22494, National Bureau of Economic Research.

7 Tables and Figures

Table 1: Boston Residential Installations (Full Sample)

Statistic	N	Mean	St. Dev.	Min	Max
Cost Per Watt (Net of ITC)	56,050	5.21	1.89	2.96	9.14
Incentive per Watt	56,050	1.86	1.48	0.25	5.25
Net Mean CPW	56,050	3.35	1.11	0.75	7.14
Median Income (10,000)	56,050	9.73	3.88	2.76	15.00
Household Size	56,050	2.67	0.42	1.48	6.73
Age of Homes (Decades)	56,050	5.55	1.40	3.10	7.70
Bedrooms	56,050	2.46	0.27	1.60	3.08
Median House Value (100,000)	56,050	3.81	1.43	1.28	7.43
Capacity Factor	14,440	0.14	0.02	0.10	0.19
NREL Annual Mean Radiation (kWh/m2/day)	56,050	4.52	0.06	4.25	4.61
NREL Annual Median Radiation (kWh/m2/day)	56,050	4.52	0.06	4.25	4.61

Table 2: California Residential Installations (Full Sample)

Statistic	N	Mean	St. Dev.	Min	Max
Cost Per Watt (Net of ITC)	103,684	5.05	1.70	2.57	8.30
Incentive per Watt	103,684	0.83	0.68	0.16	2.04
Net Mean CPW	103,684	4.22	1.07	2.41	6.62
Median Income (10,000)	103,684	6.60	2.70	1.19	21.82
Household Size	103,684	2.86	0.65	1.25	5.74
Age of Homes (Decades)	103,684	4.23	1.30	1.10	7.70
Bedrooms	103,684	2.46	0.27	1.20	3.78
Median House Value (100,000)	103,684	4.22	2.38	0.46	10.00
Capacity Factor	11,868	0.17	0.03	0.10	0.24
NREL Annual Mean Radiation (kWh/m2/day)	103,684	5.91	0.35	4.52	6.93
NREL Annual Median Radiation (kWh/m2/day)	103,684	5.92	0.35	4.53	6.96

Table 3: Tucson Residential Installations (Full Sample)

Statistic	N	Mean	St. Dev.	Min	Max
Cost Per Watt (Net of ITC)	2,754	4.13	1.34	2.72	9.33
Incentive per Watt	2,754	5.95	3.69	1	10
Net Mean CPW	2,754	−1.82	2.80	−5.99	2.16
Median Income (10,000)	2,754	5.36	1.81	2.45	8.28
Household Size	2,754	2.54	0.36	2.05	3.69
Age of Homes (Decades)	2,754	3.15	1.30	1.30	7.10
Bedrooms	2,754	2.45	0.30	1.85	3.00
Median House Value (100,000)	2,754	2.10	0.91	0.84	4.69
Capacity Factor	1,701	0.22	0.01	0.21	0.23
NREL Annual Mean Radiation (kWh/m2/day)	2,754	6.66	0.03	6.57	6.70
NREL Annual Median Radiation (kWh/m2/day)	2,754	6.66	0.02	6.61	6.70

Table 4: Estimating Discrete Choice Model of the Installation Decision

	Adoption		
	Boston	California	Tucson
	(1)	(2)	(3)
Net Cost Per Watt	−0.488***	−0.961***	−0.263***
	(0.038)	(0.025)	(0.023)
Democrat Share	2.856***	1.924***	1.821***
	(0.168)	(0.043)	(0.202)
Median Income	0.056***	0.004	0.372***
	(0.008)	(0.003)	(0.021)
Household Size	−0.222***	−0.215***	0.331***
	(0.042)	(0.008)	(0.036)
Age of Homes	−0.316***	−0.043***	0.130***
	(0.015)	(0.005)	(0.023)
Bedrooms	−0.148**	−0.213***	0.014
	(0.071)	(0.022)	(0.090)
Median House Value	0.311***	0.177***	−0.134***
	(0.013)	(0.004)	(0.027)
Electricity Prices	−8.942***	2.845***	−33.695**
	(0.954)	(0.065)	(15.480)
NREL Annual Median Radiation	−6.332***	0.450***	−0.494
	(0.255)	(0.018)	(0.479)
Linear Time Trend	19.954***	16.281***	12.435***
	(1.539)	(0.329)	(1.375)
Quadratic Time Trend	−14.776***	−17.577***	−5.256***
	(1.374)	(0.361)	(1.158)
Constant	13.164***	−12.334***	−9.857***
	(1.330)	(0.225)	(3.692)
Observations	56,050	103,684	2,754
Year FE	Yes	Yes	Yes

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5: Estimating Capacity Installed Conditional on Installation Using OLS

	Capacity Installed		
	Boston	California	Tucson
	(1)	(2)	(3)
Net Cost Per Watt	−0.292***	−0.333***	−0.129
	(0.085)	(0.091)	(0.089)
Democrat Share	−2.911***	−4.172***	−1.546
	(0.446)	(0.165)	(0.957)
Median Income	0.194***	0.010	0.086
	(0.019)	(0.012)	(0.096)
Household Size	0.343***	−0.021	−0.443***
	(0.093)	(0.028)	(0.166)
Age of Homes	−0.225***	−0.101***	−0.227**
	(0.038)	(0.018)	(0.096)
Bedrooms	0.116	−0.023	1.065**
	(0.197)	(0.082)	(0.448)
Median House Value	−0.074**	0.147***	0.676***
	(0.034)	(0.014)	(0.121)
Electricity Prices	3.708	7.565***	−50.207
	(2.481)	(0.257)	(35.017)
NREL Annual Median Radiation	−3.779***	1.541***	−0.375
	(0.752)	(0.070)	(2.443)
Linear Time Trend	0.426	−0.002	6.599
	(0.812)	(0.297)	(4.383)
Quadratic Time Trend	−1.302	−0.164	−7.116
	(1.193)	(0.374)	(5.028)
Constant	23.781***	−3.431***	13.359
	(3.565)	(0.660)	(16.515)
Observations	3,287	28,376	1,684
Year FE	Yes	Yes	Yes
Adjusted R^2	0.235	0.092	0.212
Note:			*p<0.1; **p<0.05; ***p<0.01

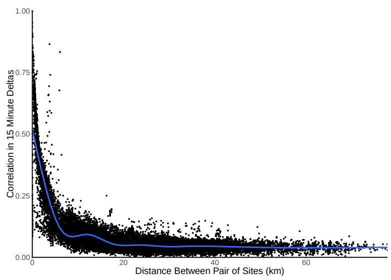
Table 6: Intermittency Cost with I.I.D. and Perfectly Correlated Solar Generation

Region (Distribution)	Solar Capacity (MW)	Solar Generation (MWh)	Total Intermittency Cost	Intermittency Cost Per MWh of Generation
Boston ($\rho = 1$)	10,234.4	13,225,600	17,701,798,900	102.03
Boston (IID)		12,770,900	15,481,929,346	92.41
California ($\rho = 1$)	12,056.3	18,731,700	28,196,080,197	114.75
California (IID)		18,368,300	25,833,355,206	107.21
Tucson ($\rho = 1$)	822.4	1,614,820	2,511,789,160	118.57
Tucson (IID)		1,611,450	2,411,325,225	114.07

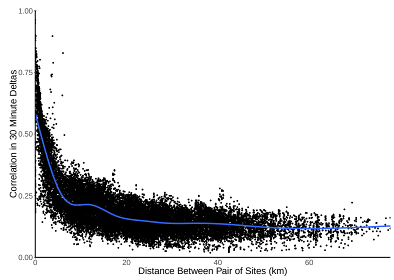
Table 7: Intermittency Cost with Concentrated or Spread Solar Installations

Region (Distribution)	Solar Capacity (MW)	Solar Generation (MWh)	Total Intermittency Cost	Intermittency Cost Per MWh of Generation
Boston (Concentrated)	10,234.4	17,760,300	19,007,108,057	1.86
Boston (Spread)		12,763,100	15,364,589,364	1.50
California (Concentrated)	12,056.3	26,345,100	30,914,918,409	2.56
California (Spread)		18,365,600	25,822,391,449	2.14
Tucson (Concentrated)	822.4	1,688,070	2,587,949,227	3.15
Tucson (Spread)		1,610,480	2,395,469,858	2.91

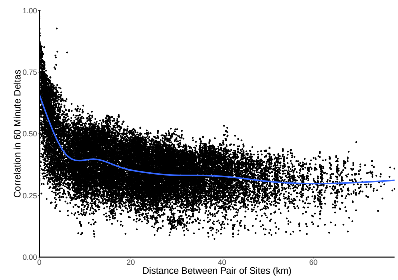
Figure 1: Correlation of changes in output as a function of distance for each pair of sites in generation sample



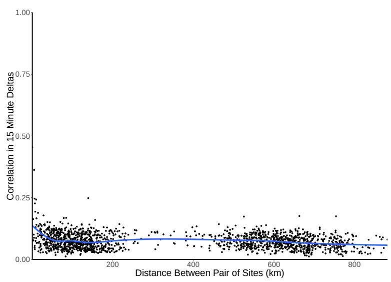
15 Minute Changes Boston



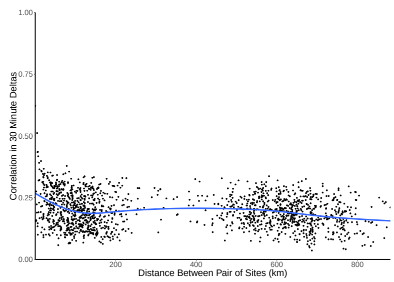
30 Minute Changes Boston



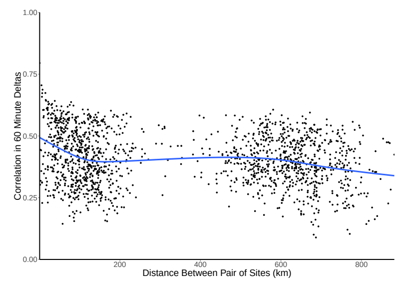
60 Minute Changes Boston



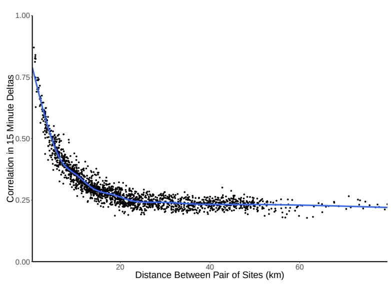
15 Minute Changes California



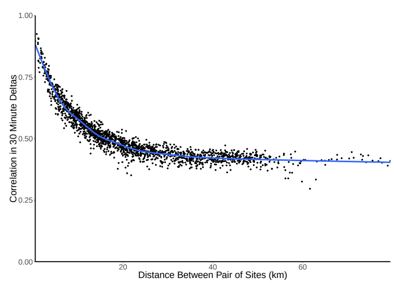
30 Minute Changes California



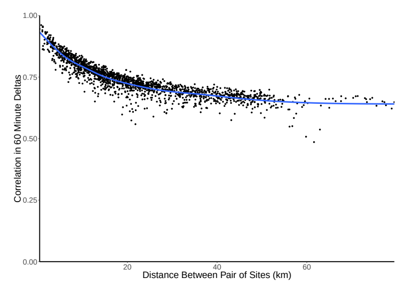
60 Minute Changes California



15 Minute Changes Tucson

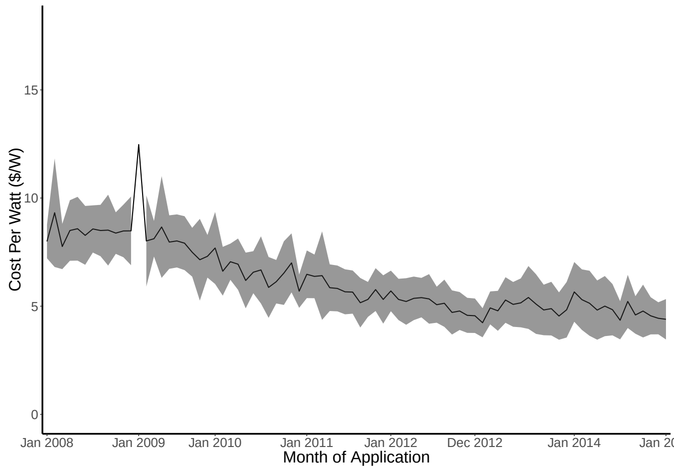


30 Minute Changes Tucson

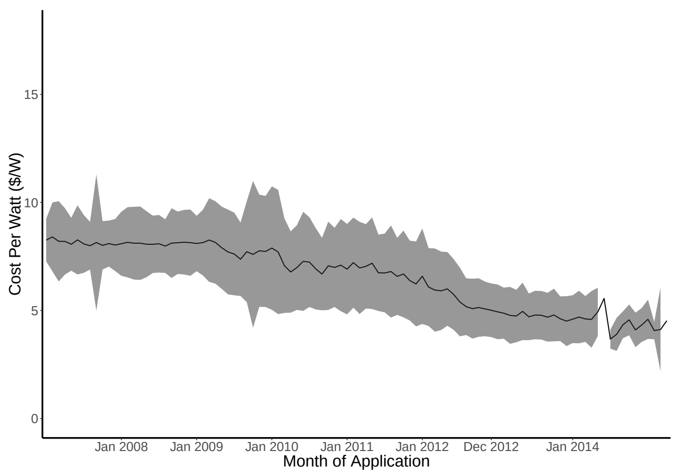


60 Minute Changes Tucson

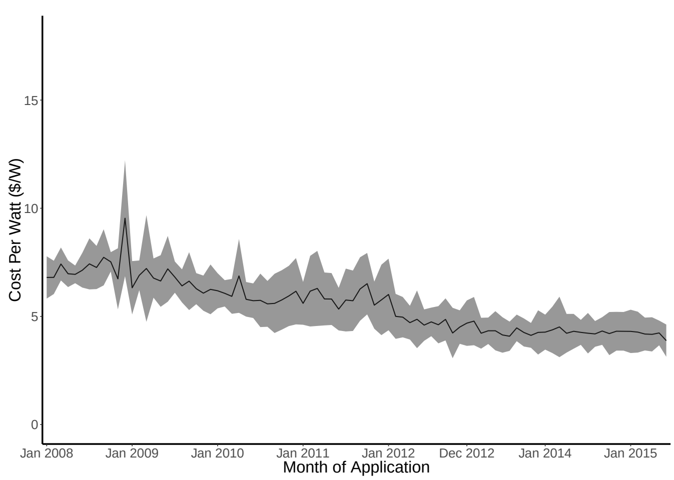
Figure 2: Cost per watt for different cities



(a) Cost per watt in Boston

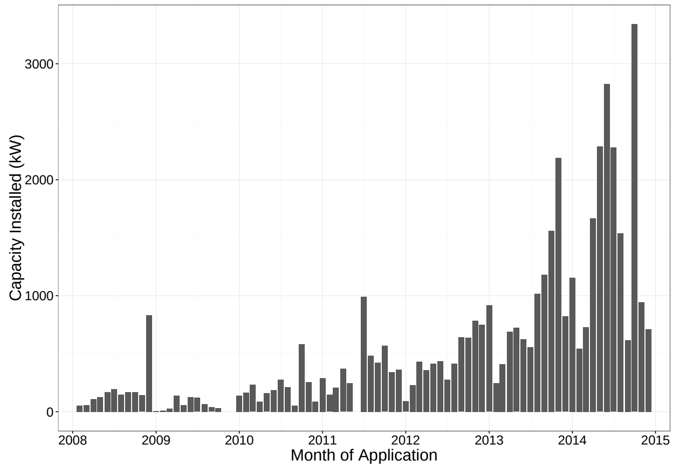


(b) Cost per watt in California

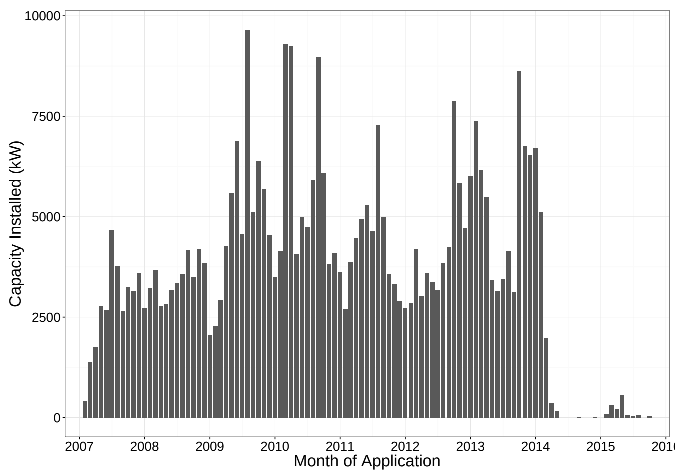


(c) Cost per watt in Tuscon

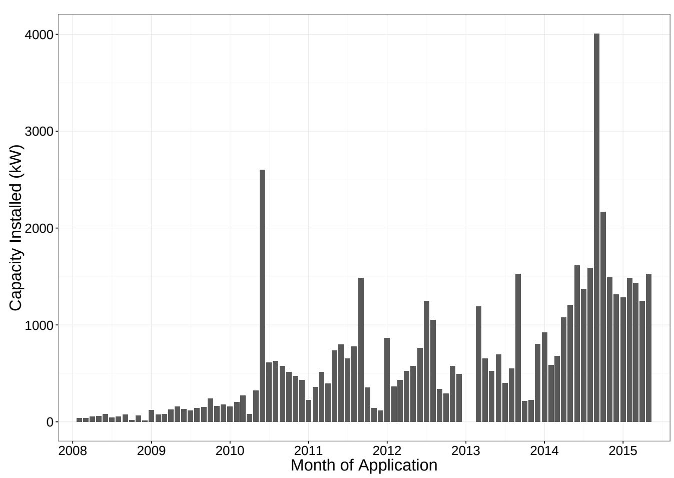
Figure 3: Capacity Installed by Month



(a) Capacity Installed in Each Month in Boston

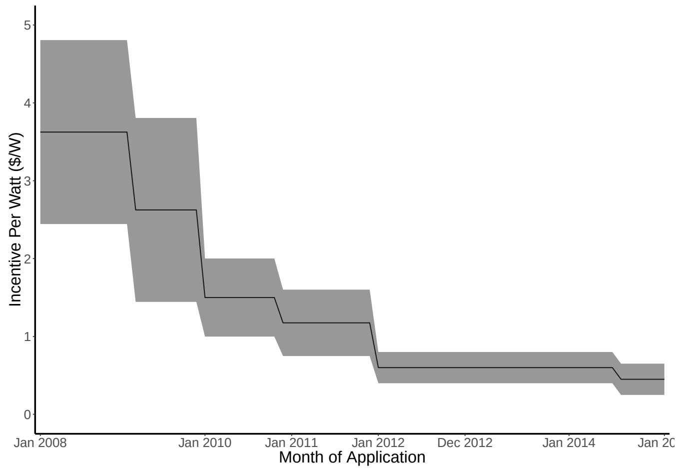


(b) Capacity Installed in Each Month in California

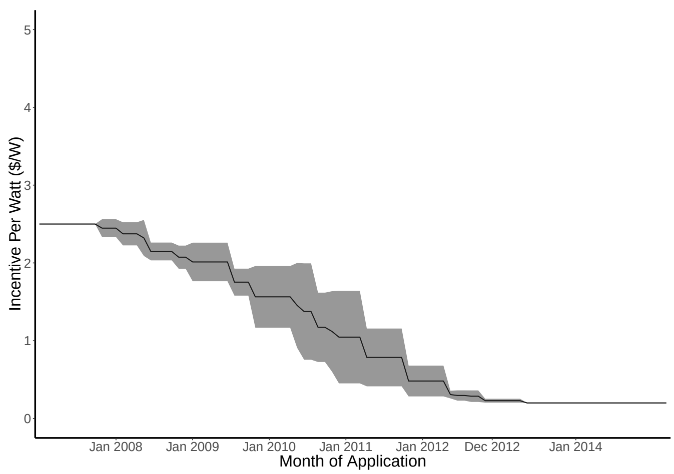


(c) Capacity Installed in Each Month in Tuscon

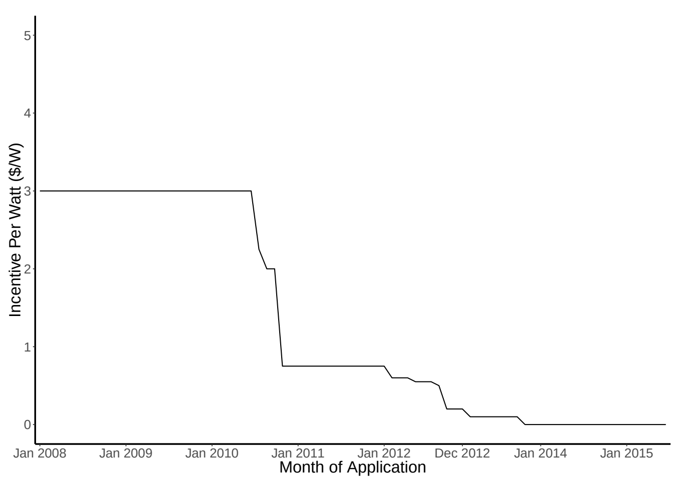
Figure 4: Incentive per watt for different cities



(a) Incentive per watt in Boston

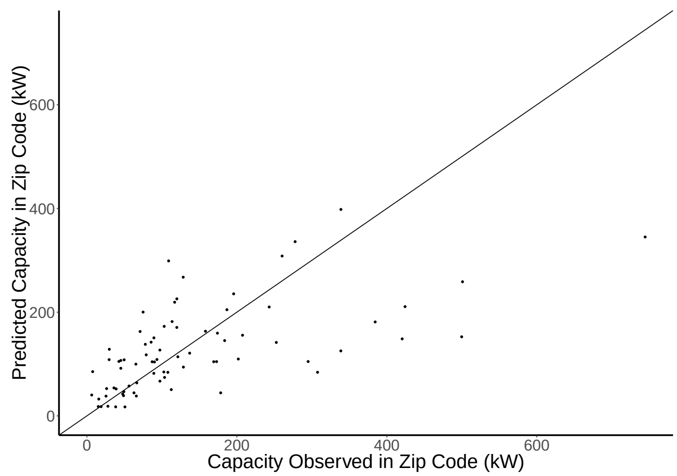


(b) Incentive per watt in California

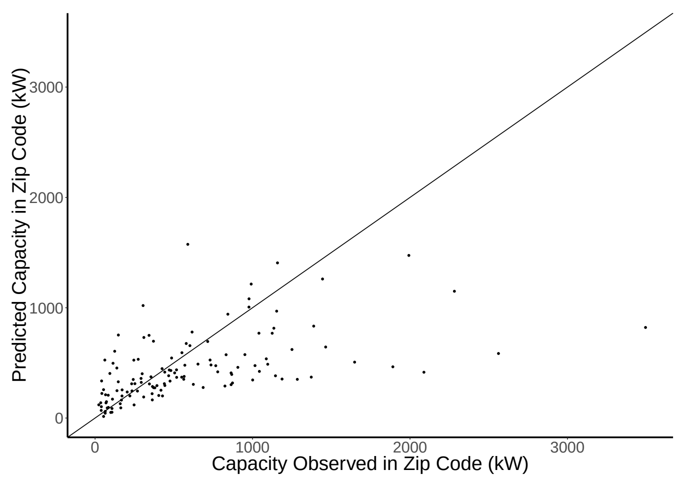


(c) Incentive per watt in Tuscon

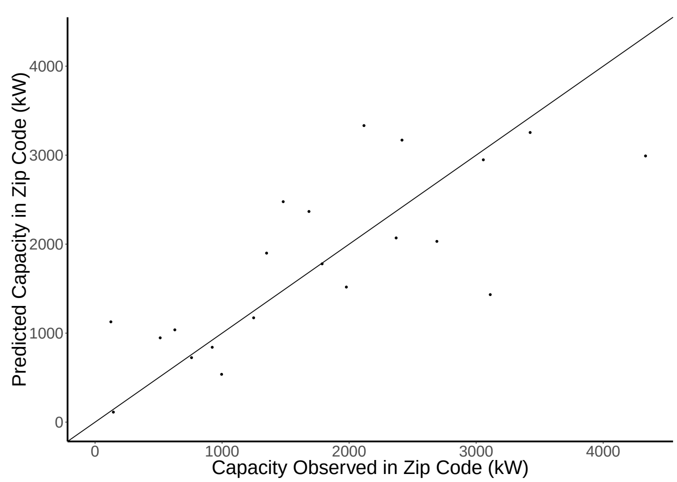
Figure 5: Predicted and Actual Capacity Installed By Zip Code



(a) Incentive per watt in Boston

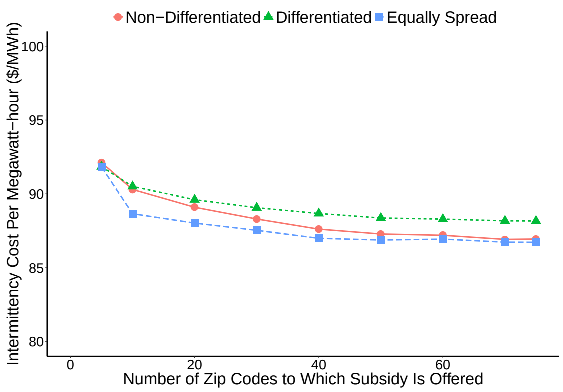


(b) Incentive per watt in California

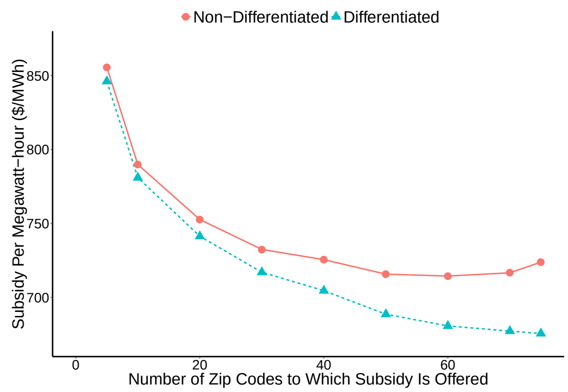


(c) Capacity in Tuscon

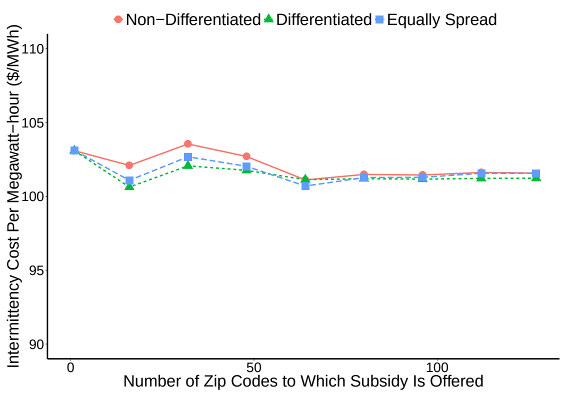
Figure 6: Intermittency Costs and Subsidy Costs as Installations Occur in More Zip Codes



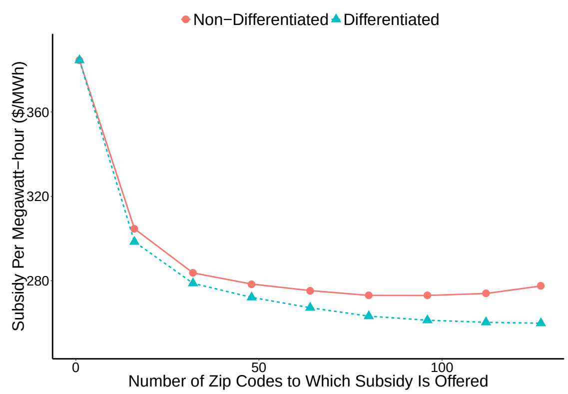
Intermittency Costs for Boston



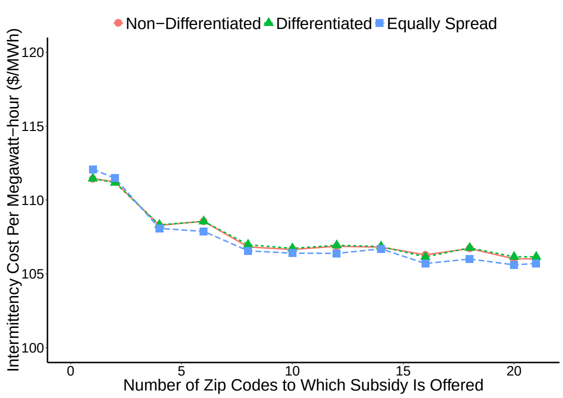
Subsidy Costs for Boston



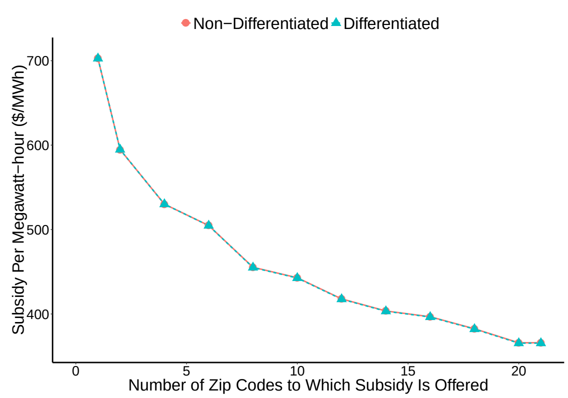
Intermittency Costs for California



Subsidy Costs for California

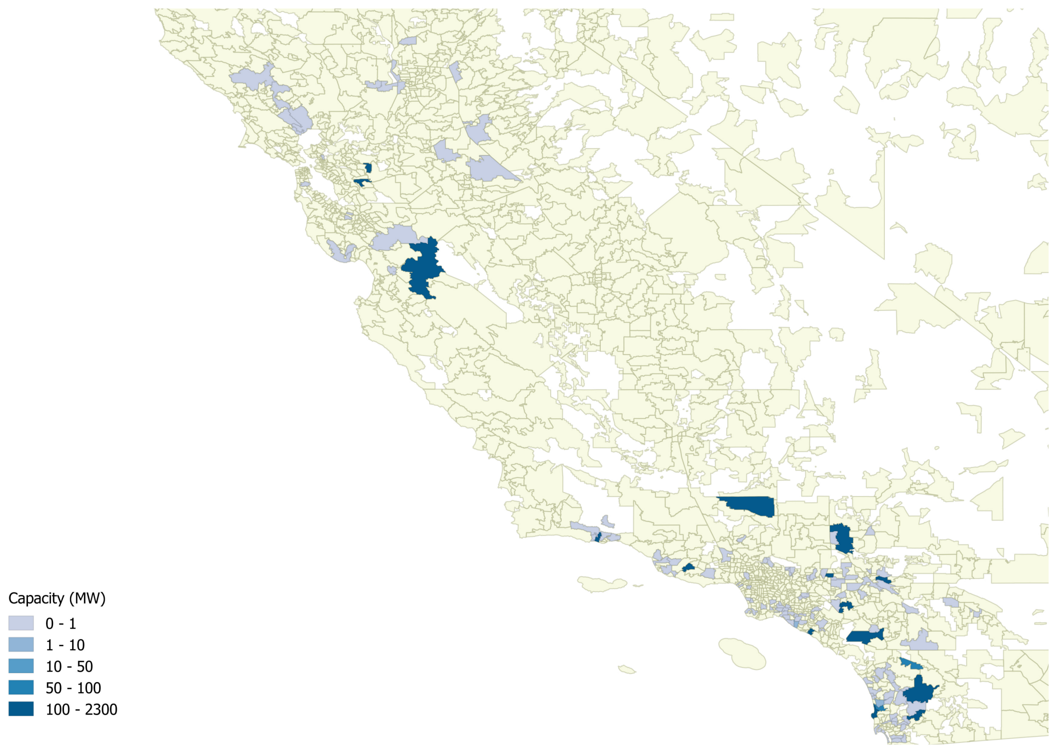


Intermittency Costs for Tucson



Subsidy Costs for Tucson

Figure 7: Capacity Installed Under Differentiated Subsidy in California



$\mathcal{N} = 16$

A Table Appendix

Table A1: Notational Glossary

α	constant fraction of the square of non-distributed generation resulting in line losses from transmission; see equation (5)
β^p, β	price coefficient and other parameters influencing household's installation decision, respectively in the discrete choice model
γ^p, γ	price coefficient and other parameters influencing household's choice of size, respectively in the reduced form model
δ^p, δ	price coefficient and other parameters influencing zip code's choice of size, respectively in the Tobit model
ε_{ijt}	unobserved component of utility for a household in zip code i choosing option j in period t in the Tobit model
η_{ijt}	unobservable component of the choice of size, conditional on installation, for a household in zip code i in period t
ν_{it}	unobservable component of the choice of size, conditional on installation, for zip code i in period t
σ	standard deviation of the normal distribution of ν_{it}
η	elasticity of electricity demand
c^{FF}	constant marginal cost of operating each new fossil fuel (combined cycle natural gas) generator
c_g	constant marginal cost of generator g
c^r	marginal cost of operating a generator as reserve capacity as a fraction of its constant marignal cost of operation
d_{it}	decision of a household in zip code i, period t to install (d_{it}=1) or not ($d_{it} = 0$)
d^{outage}	fraction of customers who lose power times the number of periods they lose power, conditional on there being an outage
g	index of conventional generators, $g \in \{1, ..., G\}$
h	index of time periods in a year (hour or sub-hour) for solar generation and in the system operator's problem, $i \in \{1, ..., H\}$

i	index of zip codes, $i \in \{1, ..., N\}$
j	decision of whether ($j = 1$) or not ($j = 0$) to install solar
m_{gh}	indicator denoting maintenance status of generator g in period h
$\vec{m}_h$	vector of generators' maintenance statuses in period h
n^{FF}	system operator's choice of number of new fossil fuel (combined cycle natural gas) generators to install
on_{gh}	indicator denoting whether generator g is available to be scheduled for generation in period h
$\vec{on}_h$	vector of indicators of generators' availability for generation in period h
$outage(\vec{on}_h, z_h, \vec{w}_h)$	indicator function for whether there is an outage in period h; see equation (6)
$\bar{p}$	retail price of electricity
p_c	curtailment price of electricity
t	index for time periods for installation data (month of sample)
u_{ijt}	indirect utility of household in zip code i choosing option j in period t
v	reservation value of electricity consumers below which they are willing to have their electricity demand curtailed
$\vec{w}_h$	weather forecast in period h
x_{it}	observables other than price and subsidy in zip code i, period t; see sections 3 and 4.1 for details
x_{gh}	available capacity of generator g in period h; if $on_{gh} = 0$, $x_{gh} = 0$; if $on_{gh} = 1$ $x_{gh} = 0$ with probability P^{fail}_{gh} and $x_{gh} = k_g$ with probability $1 - P^{fail}_{gh}$
$\vec{x}_h$	vector of available capacities for conventional generators in period h
z_h	level of electricity demand curtailed in period h
AFC^{FF}	average fixed cost of installing a new fossil fuel (combined cycle natural gas) generator per MW of capacity
AFC^{SL}	average fixed cost of installing solar generation per MW of capacity installed; varies based on the zip codes in which solar is installed
AFC^{T}	average fixed cost of transmission
$C(\cdot)$	cost of generation and operating reserves

CPW_t	average cost per watt of solar capacity across all zip codes in period t
D_h	forecasted scale of electricity demand in period h
$F^D(D_h \vert \vec{w}_h)$	distribution of D_h conditional on weather forecast
$F^S(S_h^{SL} \vert \vec{w}_h)$	distribution of S_h conditional on weather forecast
G	total number of existing conventional generations
IPW_{it}	scheduled incentive per watt of solar capacity for zip code i, in period t. In Massachusetts the variation across i within a time period is due to different level incentives for low income or low home value households. For California this is due to a few zip codes in which more than one utility (with different subsidy time paths) is present. For Los Angeles, San Diego, San Francisco and Tucson there is no variation across zip codes and this is in fact IPW_t
$\widetilde{IPW}_i$	counterfactual-geographically differentiated subsidy policy
K^{FF}	capacity of each new fossil fuel (combined cycle natural gas) generator
K_g	generating capacity of generator g
K_i^{SL}	capacity of distributed solar installed in zip code i
K^{SL}	total distributed, residential solar capacity installed
$LL(Q_h)$	line losses given demand for electricity, net of solar and demand curtailment in period h , Q_h; see equation (5)
N	total number of zip codes in a city
$\breve{N}$	number of zip codes which receive a non-zero subsidy or have non-zero solar capacity installed
N_{ijt}	number of households in zip code i choosing option j in period t
$\widetilde{N}_{ijt}$	number of households in zip code i choosing option j in period t under the counterfactual subsidy
P_{ijt}	probability (or share) of households in zip code i choosing option j in period t
P_{gh}^{fail}	probability of failure for generator g in period h
P_{gh}^{maint}	probability of maintenance for generator g in period h
Q_h	electricity demand net of solar and demand curtailment in period h; $Q_h = D_h \bar{p}^{-\eta} - z_h - K^{SL} S_h^{SL}$
$Q^D(\bar{p}, D_h)$	constant elasticity of scale specification of electricity demand; see equation (3)
S_h^{SL}	forecasted solar output for the entire region in hour h

S_{ih}^{SL}	forecasted solar output for hour h in zip code i
S_{it}	average system size installed in zip code i, period t conditional on installation ($d_{it} = 1$)
S_{it}^{*}	zip code i's choice of capacity to install in period t
$\breve{S}_{it}$	zip code i's choice of capacity to install in period t
$\widetilde{S}_{it}$	average system size installed in zip code i, period t conditional on $d_{it} = 1$ under the counterfactual subsidy
T	number of time periods over which installations occur
$TFC(K^{SL})$	discounted present value of investment in new transmission lines
$V(n^{FF})$	value function for system operator given some choice of new fossil fuel investment, n^{FF}; see equation (8)
V^{*}	total welfare from electricity market at the system operator's optimal choices of new fossil fuel investment, n^{FF}, and demand curtailment, p_c; see equation (9)
$VOLL$	value of lost load
$W(\vec{w}_h, \vec{m}_h \| n^{FF}, p_c)$	system operator's second stage payoff in a period h, see equation (7)
$WLC(z_h, p_c)$	welfare loss in a period h given demand curtailment z_h and curtailment price p_c
X_{it}	all observables including price and subsidy in zip code i, period t

Table A2: Estimating Installation Decision for Boston using Net Mean CPW (NREL)

	Adoption		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	-0.781^{***}	-0.778^{***}	-0.488^{***}
	(0.056)	(0.056)	(0.038)
Democrat Share	2.606^{***}	2.342^{***}	2.856^{***}
	(0.331)	(0.326)	(0.168)
Median Income	0.151^{***}	0.153^{***}	0.056^{***}
	(0.015)	(0.015)	(0.008)
Household Size	-0.239^{**}	-0.241^{**}	-0.222^{***}
	(0.095)	(0.100)	(0.042)
Age of Homes	0.043	0.044	-0.316^{***}
	(0.034)	(0.033)	(0.015)
Bedrooms	-0.947^{***}	-0.929^{***}	-0.148^{**}
	(0.122)	(0.123)	(0.071)
Median House Value	0.232^{***}	0.221^{***}	0.311^{***}
	(0.025)	(0.026)	(0.013)
Electricity Prices	-7.864^{***}	-7.582^{***}	-8.942^{***}
	(1.571)	(1.569)	(0.954)
Capacity Factor	5.725^{***}		
	(1.190)		
NREL Annual Median Radiation		-2.685^{***}	-6.332^{***}
		(0.998)	(0.255)
Linear Time Trend	23.518^{***}	25.416^{***}	19.954^{***}
	(2.935)	(2.982)	(1.539)
Quadratic Time Trend	-18.917^{***}	-20.552^{***}	-14.776^{***}
	(2.648)	(2.685)	(1.374)
Constant	-16.429^{***}	-3.907	13.164^{***}
	(1.118)	(4.719)	(1.330)
Observations	13,984	13,984	56,050
Log Likelihood	17,613	17,620	48,778
Actual Installations (Subsample)	1,770	1,770	1,770
Pred. Installations (Subsample)	1,777	1,777	1,918
Actual Capacity Installed (Subsample)	0	0	0
Pred. Capacity Installed (Subsample)	8.75	8.74	9.66
Year FE	Yes	Yes	Yes
Akaike Inf. Crit.	13,668.480	13,673.660	43,653.120
Bayesian Inf. Crit.	13,819.390	13,824.580	43,831.800

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A3: Estimating Installation Decision for California using Net Mean CPW (NREL)

	Adoption		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	−0.522***	−0.525***	−0.961***
	(0.046)	(0.046)	(0.025)
Democrat Share	−2.137***	−2.138***	−1.924***
	(0.101)	(0.103)	(0.043)
Median Income	−0.045***	−0.053***	0.004
	(0.007)	(0.007)	(0.003)
Household Size	−0.212***	−0.209***	−0.215***
	(0.015)	(0.015)	(0.008)
Age of Homes	−0.096***	−0.078***	−0.043***
	(0.010)	(0.010)	(0.005)
Bedrooms	−0.295***	−0.357***	−0.213***
	(0.050)	(0.050)	(0.022)
Median House Value	0.174***	0.219***	0.177***
	(0.008)	(0.008)	(0.004)
Electricity Prices	3.649***	4.827***	2.845***
	(0.138)	(0.155)	(0.065)
Capacity Factor	0.470		
	(0.372)		
NREL Annual Median Radiation		0.747***	0.450***
		(0.043)	(0.018)
Linear Time Trend	13.297***	13.440***	16.281***
	(0.739)	(0.738)	(0.329)
Quadratic Time Trend	−13.039***	−13.110***	−17.577***
	(0.797)	(0.796)	(0.361)
Constant	−10.268***	−15.211***	−12.334***
	(0.357)	(0.452)	(0.225)
Observations	11,739	11,739	103,684
Log Likelihood	130,230	130,082	570,428
Actual Installations (Subsample)	13,588	13,588	13,588
Pred. Installations (Subsample)	13,705	13,707	10,343
Actual Capacity Installed (Subsample)	0	0	0
Pred. Capacity Installed (Subsample)	77.66	77.81	56.8
Year FE	Yes	Yes	Yes
Akaike Inf. Crit.	55,100.700	55,026.390	302,529.400
Bayesian Inf. Crit.	55,255.480	55,181.180	302,729.900
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A4: Estimating Installation Decision for Tucson using Net Mean CPW (NREL)

	Adoption		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	−0.272***	−0.274***	−0.263***
	(0.025)	(0.025)	(0.023)
Democrat Share	2.866***	2.955***	1.821***
	(0.239)	(0.238)	(0.202)
Median Income	0.299***	0.206***	0.372***
	(0.029)	(0.028)	(0.021)
Household Size	0.333***	0.661***	0.331***
	(0.065)	(0.061)	(0.036)
Age of Homes	0.101***	0.121***	0.130***
	(0.026)	(0.026)	(0.023)
Bedrooms	−0.585***	−0.726***	0.014
	(0.122)	(0.131)	(0.090)
Median House Value	−0.157***	0.028	−0.134***
	(0.038)	(0.038)	(0.027)
Electricity Prices	−56.510***	−60.053***	−33.695**
	(20.385)	(21.265)	(15.480)
Capacity Factor	−41.548***		
	(3.013)		
NREL Annual Median Radiation		−0.360	−0.494
		(0.511)	(0.479)
Linear Time Trend	12.926***	13.253***	12.435***
	(1.523)	(1.524)	(1.375)
Quadratic Time Trend	−6.294***	−6.587***	−5.256***
	(1.311)	(1.312)	(1.158)
Constant	0.336	−6.525	−9.857***
	(2.635)	(4.251)	(3.692)
Observations	1,701	1,701	2,754
Log Likelihood	45,484	45,578	61,562
Actual Installations (Subsample)	5,395	5,395	5,395
Pred. Installations (Subsample)	5,514	5,509	5,246
Actual Capacity Installed (Subsample)	0	0	0
Pred. Capacity Installed (Subsample)	38.36	38.43	37.72
Year FE	Yes	Yes	Yes
Akaike Inf. Crit.	12,628.010	12,643.470	18,366.410
Bayesian Inf. Crit.	12,736.790	12,752.250	18,484.830

Note: *p<0.1; **p<0.05; ***p<0.01

Table A5: Estimating Capacity Installed via OLS for Boston using Net Mean CPW (NREL)

	Capacity Installed		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	-0.239^{*}	-0.240^{*}	-0.292^{***}
	(0.124)	(0.124)	(0.085)
Democrat Share	-3.484^{***}	-3.490^{***}	-2.911^{***}
	(0.885)	(0.873)	(0.446)
Median Income	0.139^{***}	0.139^{***}	0.194^{***}
	(0.038)	(0.038)	(0.019)
Household Size	0.319	0.318	0.343^{***}
	(0.248)	(0.247)	(0.093)
Age of Homes	-0.288^{***}	-0.291^{***}	-0.225^{***}
	(0.087)	(0.087)	(0.038)
Bedrooms	0.256	0.289	0.116
	(0.322)	(0.335)	(0.197)
Median House Value	0.200^{***}	0.202^{***}	-0.074^{**}
	(0.068)	(0.068)	(0.034)
Electricity Prices	8.957^{**}	8.953^{**}	3.708
	(4.155)	(4.155)	(2.481)
Capacity Factor	0.097		
	(3.454)		
NREL Annual Median Radiation		0.893	-3.779^{***}
		(2.900)	(0.752)
Linear Time Trend	1.447	1.436	0.426
	(1.654)	(1.653)	(0.812)
Quadratic Time Trend	−2.944	−2.931	−1.302
	(2.494)	(2.492)	(1.193)
Constant	4.748^{***}	0.653	23.781^{***}
	(1.618)	(13.420)	(3.565)
Observations	1,051	1,051	3,287
Actual Capacity Installed	0	0	0
Pred. Capacity Installed	6.24	6.24	20.06
Year FE	Yes	Yes	Yes
Adjusted R^2	0.267	0.267	0.235

Note: *p<0.1; **p<0.05; ***p<0.01

Table A6: Estimating Capacity Installed via OLS for California using Net Mean CPW (NREL)

	Capacity Installed		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	-0.359^{**}	-0.369^{**}	-0.333^{***}
	(0.159)	(0.158)	(0.091)
Democrat Share	-3.363^{***}	-3.196^{***}	-4.172^{***}
	(0.431)	(0.429)	(0.165)
Median Income	0.102^{***}	0.081^{***}	0.010
	(0.025)	(0.025)	(0.012)
Household Size	-0.035	-0.026	-0.021
	(0.065)	(0.064)	(0.028)
Age of Homes	-0.104^{**}	-0.101^{**}	-0.101^{***}
	(0.041)	(0.040)	(0.018)
Bedrooms	0.780^{***}	0.613^{***}	-0.023
	(0.200)	(0.200)	(0.082)
Median House Value	-0.021	0.069^{**}	0.147^{***}
	(0.029)	(0.031)	(0.014)
Electricity Prices	4.771^{***}	6.940^{***}	7.565^{***}
	(0.619)	(0.683)	(0.257)
Capacity Factor	2.266		
	(1.448)		
NREL Annual Median Radiation		1.455^{***}	1.541^{***}
		(0.194)	(0.070)
Linear Time Trend	-0.850	-0.773	-0.002
	(0.856)	(0.852)	(0.297)
Quadratic Time Trend	1.089	1.018	-0.164
	(1.000)	(0.995)	(0.374)
Constant	4.697^{***}	-4.626^{***}	-3.431^{***}
	(0.900)	(1.555)	(0.660)
Observations	5,466	5,466	28,376
Actual Capacity Installed	0	0	0
Pred. Capacity Installed	30.3	30.3	156.08
Year FE	Yes	Yes	Yes
Adjusted R^2	0.051	0.060	0.092
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table A7: Estimating Capacity Installed via OLS for Tucson using Net Mean CPW (NREL)

	Capacity Installed		
	Subsample		Full Sample
	(1)	(2)	(3)
Net Cost Per Watt	−0.091	−0.094	−0.129
	(0.103)	(0.103)	(0.089)
Democrat Share	−3.111***	−3.283***	−1.546
	(1.155)	(1.170)	(0.957)
Median Income	0.052	0.108	0.086
	(0.132)	(0.130)	(0.096)
Household Size	−0.095	−0.240	−0.443***
	(0.314)	(0.303)	(0.166)
Age of Homes	−0.234**	−0.224*	−0.227**
	(0.115)	(0.115)	(0.096)
Bedrooms	1.698***	1.784***	1.065**
	(0.616)	(0.644)	(0.448)
Median House Value	0.917***	0.814***	0.676***
	(0.173)	(0.173)	(0.121)
Electricity Prices	−63.745	−63.194	−50.207
	(44.232)	(44.271)	(35.017)
Capacity Factor	20.544		
	(13.933)		
NREL Annual Median Radiation		−0.435	−0.375
		(2.655)	(2.443)
Linear Time Trend	7.569	7.654	6.599
	(5.119)	(5.123)	(4.383)
Quadratic Time Trend	−9.476	−9.507	−7.116
	(5.891)	(5.897)	(5.028)
Constant	6.535	13.971	13.359
	(6.584)	(18.115)	(16.515)
Observations	1,187	1,187	1,684
Actual Capacity Installed	0	0	0
Pred. Capacity Installed	7.89	7.89	11.1
Year FE	Yes	Yes	Yes
Adjusted R^2	0.261	0.260	0.212

Note: *p<0.1; **p<0.05; ***p<0.01

Table A8: Estimating Tobit Model of the Installation Decision

	Capacity Installed		
	Boston	California	Tucson
	(1)	(2)	(3)
Net Cost Per Watt	−1.808***	−11.611***	−4.993***
	(0.430)	(0.455)	(1.390)
Democrat Share	8.951***	−13.887***	57.660***
	(2.265)	(0.742)	(12.853)
Median Income	−0.072	1.418***	8.852***
	(0.083)	(0.050)	(1.273)
Household Size	0.370	−1.346***	3.381
	(0.497)	(0.130)	(2.346)
Age of Homes	−1.458***	−0.724***	3.149**
	(0.203)	(0.082)	(1.285)
Bedrooms	−1.846*	−8.218***	−10.652**
	(0.965)	(0.344)	(5.258)
Median House Value	2.538***	1.189***	2.076
	(0.167)	(0.061)	(1.651)
Electricity Prices	−80.483***	−1.266	−534.699
	(13.916)	(1.258)	(368.781)
NREL Annual Median Radiation	−10.643***	4.310***	46.794
	(3.580)	(0.321)	(33.270)
Linear Time Trend	162.806***	207.073***	212.210***
	(16.104)	(5.661)	(63.729)
Quadratic Time Trend	−116.773***	−239.110***	−37.709
	(15.926)	(6.326)	(64.512)
Constant	−33.603*	−12.146***	−426.248*
	(17.886)	(3.916)	(227.709)
Observations	56,050	103,684	2,754
Log Likelihood	-21,806.6	-151,243.7	-9,163.2
Akaike Inf. Crit.	43,653.1	302,529.4	18,366.4
Bayesian Inf. Crit.	43,831.8	302,729.9	18,484.8
Note:			*p<0.1; **p<0.05; ***p<0.01

Table A9: Differentiated Subsidies Simulations for Boston

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
5	8,310	13,365,900	20 / 7	1,227,581,791	148,351,028,657
10	8,490	13,354,300	20 / 7	1,208,693,813	136,816,895,082
20	8,741	13,343,100	20 / 7	1,195,634,516	129,759,678,900
30	8,949	13,349,200	20 / 7	1,188,923,616	125,560,679,162
40	9,099	13,344,700	20 / 7	1,183,340,339	123,337,270,949
50	9,302	13,350,900	20 / 7	1,179,796,589	120,608,513,568
60	9,437	13,352,700	20 / 7	1,178,944,413	119,239,479,865
70	9,512	13,353,200	20 / 7	1,177,437,713	118,630,594,479
75	9,559	13,353,600	20 / 7	1,177,341,473	118,350,342,724

Note: Baseline / Perfectly Dispatchable Solar

Table A10: Non-Differentiated Subsidy Simulation for Boston

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
5	8,392	13,382,600	21 / 7	1,232,857,833	150,202,907,259
10	8,615	13,359,700	20 / 7	1,206,374,734	138,428,713,280
20	8,959	13,334,700	20 / 8	1,188,058,442	131,657,472,131
30	9,281	13,346,900	20 / 7	1,178,532,723	128,244,373,561
40	9,557	13,335,400	20 / 8	1,168,376,043	126,919,099,248
50	9,928	13,350,200	20 / 7	1,165,284,486	125,344,875,886
60	10,239	13,355,500	20 / 7	1,164,578,182	125,170,229,058
70	10,462	13,355,900	21 / 7	1,160,916,236	125,561,748,170
75	10,700	13,357,600	21 / 7	1,161,393,923	126,841,123,107

Note: Baseline / Perfectly Dispatchable Solar

Table A11: Install Equal Capacity Across N Zip Codes Simulation in Boston

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost
5	7,716	13,390,600	21 / 7	1,229,619,246
10	8,625	13,312,100	20 / 7	1,180,213,278
20	9,063	13,310,800	20 / 8	1,171,660,522
30	9,377	13,322,400	20 / 7	1,166,178,221
40	9,654	13,318,600	20 / 7	1,158,629,394
50	9,909	13,327,400	20 / 7	1,157,869,019
60	10,180	13,333,300	20 / 8	1,159,175,143
70	10,478	13,334,000	20 / 8	1,156,508,156
75	10,656	13,336,600	20 / 8	1,156,621,504

Note: Baseline / Perfectly Dispatchable Solar

Table A12: Differentiated Subsidies Simulations for California

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
1	8,796	19,195,700	37 / 4	1,978,904,199	96,824,677,346
16	9,761	19,038,100	34 / 6	1,915,745,623	74,537,013,872
32	10,123	19,048,600	33 / 6	1,944,161,415	69,663,343,696
48	10,326	19,042,900	32 / 6	1,937,770,110	67,975,903,983
64	10,512	19,040,800	33 / 6	1,925,522,632	66,742,217,764
80	10,685	19,042,800	33 / 6	1,927,044,077	65,753,155,378
96	10,790	19,042,300	33 / 6	1,926,378,898	65,269,490,025
112	10,855	19,042,000	33 / 6	1,927,309,644	65,023,600,008
127	10,896	19,041,200	33 / 6	1,927,466,078	64,910,357,645

Note: Baseline / Perfectly Dispatchable Solar

Table A13: Non-Differentiated Subsidy Simulation for California

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
1	8,796	19,195,700	37 / 4	1,978,904,199	96,824,677,346
16	10,117	19,030,900	34 / 6	1,942,866,041	76,055,044,067
32	10,467	19,053,100	37 / 6	1,972,870,287	70,908,330,014
48	10,776	19,043,700	37 / 6	1,955,690,392	69,556,972,882
64	11,101	19,040,400	32 / 6	1,925,172,681	68,771,814,733
80	11,419	19,048,000	33 / 6	1,933,032,293	68,229,753,224
96	11,666	19,046,500	33 / 6	1,932,143,891	68,217,734,754
112	11,865	19,045,500	33 / 6	1,935,296,335	68,446,967,976
127	12,141	19,039,900	33 / 6	1,933,920,512	69,315,827,089

Note: Baseline / Perfectly Dispatchable Solar

Table A14: Install Equal Capacity Across N Zip Codes Simulation in California

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost
1	8,785	19,196,300	37 / 4	1,979,353,715
16	10,075	19,031,000	33 / 6	1,923,685,028
32	10,444	19,052,900	33 / 6	1,956,296,883
48	10,786	19,043,000	33 / 6	1,942,971,555
64	11,112	19,038,700	32 / 6	1,917,071,499
80	11,407	19,045,500	33 / 6	1,928,732,617
96	11,711	19,039,900	33 / 6	1,928,460,755
112	12,019	19,040,100	33 / 6	1,933,715,060
127	12,447	19,032,200	33 / 6	1,932,804,893

Note: Baseline / Perfectly Dispatchable Solar

Table A15: Differentiated Subsidies Simulations for Tucson

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
1	842	1,721,610	4 / 0	191,878,565	15,869,072,573
2	843	1,719,640	4 / 0	191,170,584	13,416,263,841
4	846	1,716,140	4 / 0	185,887,749	11,933,047,489
6	849	1,715,820	4 / 0	186,258,879	11,358,935,524
8	858	1,716,030	4 / 0	183,570,728	10,242,366,620
10	860	1,716,240	4 / 0	183,173,264	9,964,016,619
12	863	1,716,560	4 / 0	183,555,962	9,402,035,150
14	866	1,716,450	4 / 0	183,408,054	9,079,433,899
16	867	1,716,110	4 / 0	182,165,142	8,923,100,682
18	870	1,716,090	4 / 0	183,211,791	8,601,771,742
20	874	1,716,340	4 / 0	182,198,134	8,227,215,780
21	874	1,716,340	4 / 0	182,198,134	8,227,215,780

Note: Baseline / Perfectly Dispatchable Solar

Table A16: Non-Differentiated Subsidy Simulation for Tucson

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost	Subsidy Spending
1	842	1,721,610	4 / 0	191,878,565	15,869,072,573
2	843	1,719,680	4 / 0	191,261,184	13,416,491,208
4	846	1,716,110	4 / 0	185,832,034	11,934,014,069
6	849	1,715,800	4 / 0	186,271,998	11,361,222,351
8	859	1,716,080	4 / 0	183,318,531	10,250,353,694
10	861	1,716,320	4 / 0	183,042,698	9,971,803,480
12	865	1,716,660	4 / 0	183,463,509	9,410,234,628
14	867	1,716,530	4 / 0	183,335,277	9,087,174,025
16	868	1,716,090	4 / 0	182,397,708	8,930,886,025
18	871	1,716,140	4 / 0	183,140,255	8,610,163,028
20	876	1,716,420	4 / 0	181,970,581	8,238,580,820
21	876	1,716,420	4 / 0	181,970,581	8,238,580,820

Note: Baseline / Perfectly Dispatchable Solar

Table A17: Install Equal Capacity Across N Zip Codes Simulation in Tucson

N	Capacity (MW)	Solar Generation (MWh)	New NGCC	Intermittency Cost
1	839	1,722,310	4 / 0	193,033,100
2	841	1,719,890	4 / 0	191,722,292
4	846	1,715,820	4 / 0	185,438,652
6	851	1,715,650	4 / 0	185,074,653
8	856	1,715,910	4 / 0	182,848,565
10	860	1,716,500	4 / 0	182,633,056
12	863	1,716,700	4 / 0	182,638,765
14	866	1,716,720	4 / 0	183,156,384
16	869	1,715,890	4 / 0	181,369,890
18	871	1,715,870	4 / 0	181,892,896
20	874	1,716,070	4 / 0	181,234,074
21	877	1,716,380	4 / 0	181,427,511

Note: Baseline / Perfectly Dispatchable Solar