Solar Photovoltaic Installation in California: Understanding the Likelihood of Adoption Given Incentives, Electricity Pricing and Consumer Characteristics

Emily Rothfield

Professor Christopher Timmins, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2010

Table of Contents

Abstract

This research aims to empirically examine the effect of previous installations in a given location on the rate of residential solar photovoltaic installation in California; controlling for electricity prices, varying levels of policy incentive, and personal characteristics. Using a combined probit and negative binomial regression to describe the number of installations observed each year in each zip code, parameter coefficients are estimated and marginal effects of installation are understood. Potential government policy is explored and a marginal cost of emissions abatement in California is recovered.

Acknowledgements

I would like to thank Professor Christopher Timmins, my faculty adviser, for his enthusiasm, patience and support. His commitment to this work was crucial and I am grateful for all the effort and time he devoted to this study. Professor Kent Kimbrough, and my classmates in our honors seminar, provided indispensable suggestions that improved my work immensely. My family, for their interest and endless support.

1. Introduction

i. Motivation

In the wake of growing concerns about global warming and fossil fuel depletion, public focus is shifting towards the implementation of renewable energy sources. In 2008, renewable energy sources such as wind energy, biomass, geothermal energy, hydroelectricity and solar energy constituted only 7 percent of U.S. energy supply (EIA 2009). Of that 7 percent, solar energy comprised 1 percent of total renewable energy supply, with 65 percent of such solar energy being produced in California. The purpose of this study is to understand the potential for increasing solar usage, in order to ultimately decrease our reliance on fossil fuels and mitigate our carbon footprint. Around the world, political leaders are facing greater pressures from the public who have concerns about the means of acquisition, costs and availability of fossil fuel sources. Thus, the need for renewable energies has never been greater.

The adoption of renewable energy technologies is pertinent in today's society because of the wealth of evidence detailing the effects of climate change. If the atmospheric concentration of carbon dioxide (CO_2) were to double, the average surface temperature is predicted to increase by three degrees Celsius (IPCC 2007). Such warming has major repercussions for industrial agriculture and ecology as a whole. The burning of fossil fuels for energy and electricity generation is the major source of CO_2 in the atmosphere. Nitrous oxides are also emitted when fossil fuels are burnt (Maslin 2009). The rapid industrialization of society in the last two centuries has seen the demands for energy sky rocket, causing a 30 percent increase in CO_2 parts per million by volume. It is widely understood that this is causing the Earth to warm faster than any other period of time in the last 2000 years. This could result in the accelerated melting of

ice-caps in Antarctica and Greenland, and an increase in extreme climate events (such as hurricanes, heat waves and floods).

In an effort to combat climate change, consumers are presented with a wide range of choices of how to fulfill their energy needs; whether this is through regular fossil fuel based technologies or newer, renewable energy sources such as hydro power, wind generation or solar power. Because of the rapidly increasing rate of technological innovation, solar equipment is more readily available to the individual consumer than ever before.

The aim of this research is to understand the best policy options that the government can adopt in order to increase solar panel installation rates at the residential level. A model is created to understand the relationship between both fluctuating electricity prices and installation incentives for a given household, and the likelihood of that household to purchase solar equipment today. Demographics of each location are controlled in this model such as income level, age, race, and education. This gives an understanding of how specific variables impact the likelihood that any given individual will adopt. Then we can understand the marginal increase in incentive levels (or any other variable) that is required in order to increase the installation rate by a certain amount of households. From this information, this research seeks a relevant policy that can be applied to the incentive variables. Specifically, the effect of past installations in a given location is observed as a key influence on future installations.

I have chosen to examine solar photovoltaic electricity and not other sources of renewable energy because solar energy can be generated on the same cost schedule on almost any scale at any location with sufficient solar radiation, ranging from small solar powered calculators in urban areas to expansive rural solar panel fields. Solar photovoltaic panels are the dominant technology for residential renewable electricity generation. It is comparatively easy to

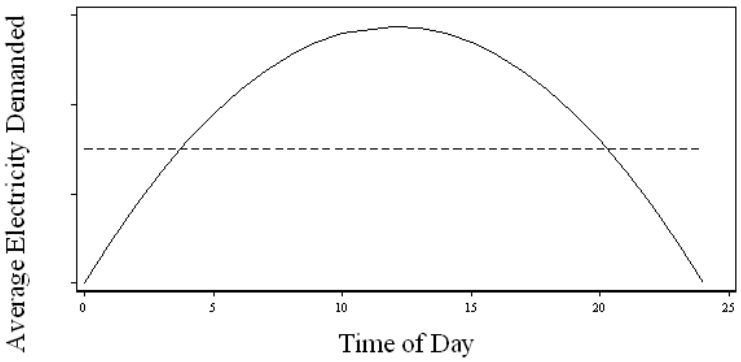
distribute, and sunlight is the most reliable source of renewable energy nature has to offer. If solar PV becomes common, economies of scale will develop, allowing extensive dissemination of the technology and countless numbers of households switching to renewable energy. This will decrease the amount of CO_2 emitted into the atmosphere and reduce our overall carbon footprint.

ii. Background

So, what is solar renewable energy? The most prevalent system of solar electricity generation is the photovoltaic panel (PV). For residences, these panels are mounted either on the rooftop or adjoining land. An inverter system converts direct current into alternating current at the site. By producing and converting electricity at the intended site of use, PV significantly reduces transmission costs that are incurred by remotely located power plants that feed electricity onto the main electricity grid (which connects all residences, commercial and industrial entities to their central shared source of electrical power). If electricity is generated by PV in excess of a household's needs, the excess energy is transferred back onto the grid. In many cases, the PV owner's electricity bill is credited for the power they add back onto the grid. Grid-connected systems do not require a battery as the grid is effectively used as a large base storage system. The bulk of the solar market is in such grid-connected systems, which range from as small as a few hundred watts, to tens of megawatts at peak capacity (Bradford 2006). Such systems are the major driving force behind the growth of the solar PV market. Grid-connected systems are favored over off-grid systems because solar power production peaks during the day when electricity usage is typically highest, and provides no electricity at night unless the owner purchases a (rather expensive) battery. Silverman (2008) contends that on average PV systems generate one-fifth of their peak capacity daily (known as their average capacity). As shown in the

Figure 1, this is because average electricity demand follows a peak-curved path, which is directly correlated with normal hours of daylight through a 24-hour cycle.

Figure 1. Hourly Demand for Electricity.



There is great potential for future developments in solar PV technology that will markedly increase its efficiency and capabilities. The conversion efficiency of PV technology is presently around 17 percent, because energy is lost when the direct current electricity is changed to alternating current for use in the home (Maslin 2009). This shows that there is plenty of room for technological innovation in the means of converting sunlight to electrical voltage. The industry is rapidly expanding, making PV a necessary area to conduct research into the dynamics of its installation rate.

Solar equipment and electricity provision is different from typical power sources (such as grid-electricity and gas) because it incurs a high initial fixed cost, with little to no future input cost of subsequently generating electricity. Some advantages of PV that the consumer considers are its flexibility (the consumer can choose the size of system to fit their available rooftop, or land, space), it has a lifetime that averages 25 years, and has typically low maintenance costs (Borenstein 2008).

Currently, it is widely acknowledged that the costs associated with solar PV are much greater than the financial benefits, even after adjusting for gains due to avoided transmission

losses (see Austin et al, 2005; Bradford 2006; Roaf and Gupta 2007; Borenstein 2008; Heal 2009). This is clearly a disincentive to purchase. However, as of late 2008 California alone has 29,628 grid-connected systems installed.[1] The current rate of installation despite the documented lack of pure financial incentives shows that there must be other factors that drive the installation rate. Such factors may be environmental consciousness, age of householder, reflection of an independent lifestyle, and prestige (Durham et al. 1988). Such social ideas will be further discussed in the following literature review. This is the motivation for this research, and the reason behind looking at previous installations, electricity prices, incentives and demographic characteristics as explanatory variables in a count model regression.

The literature in solar energy, and specifically solar photovoltaics, has so far been focused on two different topics; the cost-benefit analysis of installation to the consumer, and the socio-economic factors that motivate installation. While these areas are relevant in understanding the propensity to install PV panels, this research hopes to isolate the specific impact that previous installations in a given location may have on the decision for residential users. I will use the observed rate of solar PV installation to understand how these certain factors influence the installation rate of PV in California. To the best of my knowledge, no study of this kind has been conducted thus far.

2. Literature Review

i. Social Spillover

Jacobsson and Johnson (2000) highlight the importance of social networks that aid in the diffusion of new renewable technologies. Once a technology is present in the consciousness of consumers it becomes more viable for installation than an unknown source of energy. This is known as a social spillover mechanism. At present, this mechanism has not been empirically

[1] http://www.energyalmanac.ca.gov/renewables/solar/pre-1998.html

examined in the case of solar PV panels. Baerenklau (2005) contends that individuals tend to be more willing to adopt an innovative technology when they observe their peers doing so first, and because individuals in the same group tend to behave similarly because they have comparable characteristics and face like constraints. However, it is difficult to decisively recover such social spillover effects. This is because it is unclear whether the individual adopts because their neighbor has adopted, or whether they adopt because they face similar environmental constraints to the larger group at hand. The following research seeks to address this gap in the literature by quantitatively looking at the cumulative count of installations in a given location, and observing how it affects the current rate of installation.

ii. Cost-Benefit Analyses

It is widely understood that presently, the financial returns to solar PV do not outweigh the costs. Much of the past literature discusses the consumer decision process, and whether or not purchasing solar PV will result in cost benefits. Heal (2009) identifies that most renewable energy sources have a large fixed cost, with little or no variable cost. Renewable energy sources are almost entirely upfront capital costs. Heal contends that the user is essentially paying for a flow of services from the given investment. Borenstein (2008) also discusses the market value of solar PV, incorporating factors such as solar radiation available at each panel location, and the costs of installation. He shows that because the wholesale electricity market has somewhat stabilized flat-rate electricity prices, and the costs of solar technology are still relatively high, that at their current price solar PV panels will not financially benefit the residential investor. Borenstein's research incorporates no social characteristics, and concludes that no individual should adopt at present, because costs outweigh benefits. Neither Borenstein nor Heal consider that other factors may be driving adoption, and they do not discuss why there are so many

installations at present despite the fact that they have shown installing to be unprofitable at this time. This is a gap in their analysis, one that this research aims to address.

Fry (1986) concludes that solar hot water heating will only be more cost effective than gas or electricity in Hawaii and Arizona. Using a finance model Fry looks at net savings and net cost from installation. He concludes that the optimal type and size of solar panels vary by location throughout the country. Fry finds that the inconsistency of solar radiation availability across the U.S. presents too much variation for solar hot water heating in 1986 to be financially beneficial to a majority of households. Fry's work does look at characteristics such as household income and level of education; however his examination of solar hot water technology does not allow a full crossover examination into the adoption of solar PV panels. Fry finds that higher levels of income and education for any household increase the propensity to install solar hot water heating. This work indicates that there may be similar effects in adoption of PV panels.

Borenstein (2007) calculates whether installing solar PV will be financially beneficial to the consumer. He shows that the time value of money is central to the evaluation of solar PV decisions, because it is a long-lived investment with large initial capital costs. Borenstein calculates the lifetime cost per kilowatt hour (kWh) that PV panels will produce (inflation adjusted) and uses different interest rates to find a net present cost of the system, which is then amortized over the lifetime of power being provided to the household. To find the cost of PV power to the consumer Borenstein then compares this constant real cost per kWh each year to the present value of money saved on electricity bills due to reduced demand for grid-provided electricity. Here, Borenstein concludes that there are conditions that exist in which solar PV can be financially beneficial to the consumer. This is for some Pacific Gas and Electricity (PG&E) customers, as long as their electricity prices stay as high as they were in 2007. Borenstein also

shows quite clearly that it is the large initial cost of PV that causes installation to be not financially viable for most customers, and that the electricity itself that is generated by PV is markedly cheaper than electricity from the grid. Without this substantial initial cost, PV would be optimal in all locations. In this work, Borenstein identifies that high electricity price may be driving installation even though it is not financial sustainable for all consumers.

iii. Like Borenstein, Fujii and Mak (1984) demonstrate that, on average, higher electricity rates increase the likelihood of installing solar hot water equipment. Their model addresses how different households aim to conserve energy in the face of rising electricity costs. The authors look at efficiency methods such as air-drying clothes, turning off the heating or lights, and installing restricted water showerheads. The authors anticipate that the benefits of such conservation will vary by household income and the marginal price of electricity. Their examination of household income as a variable shows that, in general, those with a higher income are more likely to adopt expenditure-intensive solutions to their conservation problems; such as installing solar hot water equipment or purchasing low-flow showerheads. Fujii and Mak also indicate that the adoption of solar increases with the marginal price of energy. ***Pricing Solar Power***

Wegner and Herig (1997) look at the benefits of net metered electricity and the California legislative reforms that relate to solar energy generation. They illustrate how net metering allows PV owners to "spin the meter backwards" as they feed electricity back onto the grid, thus obtaining the full retail rate for electricity (Wegner and Herig 1997). So, when a residential electricity bill is calculated for a home with solar PV, their net electricity usage will be total usage minus kilowatt hours the residence fed back onto the grid. The authors also conclude that market-based government policies can change the viability of grid-connected PV

into a competitive source of energy in the U.S. market through rebates, low-interest loans and the continuation of net metering. As an extension of this work, Borenstein (2005) shows that with the combination of net-metering and switching to a time-of-use (TOU) electricity rate after the installation of solar panels, the electricity price faced by the installed household will almost always be less than that of the household without PV panels. This reiterates the idea that solar power is much cheaper than grid power.

iv. Social Characteristics of Adoption

The propensity to install solar PV panels also has many social influences. Roaf and Gupta (2007) argue that solar energy is the most favorable renewable technology amongst members of the community because it is clean, quiet, and can be used as a favorable aesthetic design feature on many forms of architecture. Durham et al. (1988) demonstrate that the likelihood of solar hot water heater installation increases with some level of college education and household size. Using survey data, the authors conclude that higher levels of education, larger household size, higher incentive levels and reduced costs of installation significantly affected the probability of installing solar hot water heaters. While their conclusions are indicative of likely consumer behavior in regards to solar PV installation, these empirical findings given by Durham et al. cannot be directly applied to solar PV panels because the costs and uses of technologies differ so greatly. Also, their results regarding income were insignificant. The following research will seek to improve this variable as I believe it is likely to have a great impact on PV installation rates.

Fuchs and Arensten (2002) illustrate that consumers have little experience purchasing alternative forms of energy because monopoly providers have dominated the market in all history of utility provision. So, consumers find the breadth of choice overwhelming, and often

opt to remain with their current provider instead of switching to solar PV power. Clearly this is a barrier to installation, and it cannot be easily quantified. The authors also demonstrate the importance of early solar adopters on the market, as these individuals ease the transition for subsequent users who have a basic lack of knowledge regarding the supply side of electricity provision.

v. Technological Innovation

Technological innovation is another barrier to PV installation. Goldman et al. (2005) discuss the tendency for investors to shy away from opportunities because of potential future innovations that are likely to be more cost efficient for the consumer than current PV technologies. They show that investors worry about "technology risk"; the idea that the given technology will not perform optimally over its lifetime, or that the technology becomes prematurely obsolete (Goldman et al. 2005). However, Bubenzer and Luther (2003) contend that all technology follows a learning curve that keeps pace with the market, guaranteeing steady progress and productivity. For solar panels, this means that future technology advancement is more than likely – that is, technology development is unlikely to stagnate. At any point in time the technology improvement will be increasing at a constant rate. Bradford (2006) argues that this implied "learning rate" is roughly 18 percent; meaning that every time the installed volume of PV doubles, the per-unit cost falls an additional 18 percent (Bradford 2006). Bradford asserts that an 18 percent learning rate combined with a 20 percent annual market growth rate results in a cost decline of 5 to 6 percent each year. Austin et al. (2005) develop a model that includes an uncertainty factor, showing that the consumer can predict technology improvement with confidence for fifteen years into the future, but after that there is great uncertainty as to how any given technology will improve.

Clearly, this has implications for adopters who are holding off purchasing PV panels for hopes of technological innovations. From the work done by Austin et al. it can be assumed that the consumer will make no definite decisions about adopting after fifteen years into the future. However, as this uncertainty period is in a far distant future time period, it will be assumed for this research that the consumer understands this general approach to technological innovation, and that the innovation of solar PV is improving steadily and consistently, and therefore the "technology risk" present today will be exist tomorrow at a very similar rate.

vi. Other Barriers to Entry

Apart from the previously discussed cost-benefit analyses, some of the other barriers to installation that have been identified are the difficulties surrounding installation and the fact that PV does not add sufficient property value to merit installation (Faiers and Neame 2006). The authors conclude this by a sample drawn from 100 solar power adopters in the United Kingdom during 2005. Dymond (2002) also shows that the initial upfront cost – and the necessary leveraged financing required – is a significant barrier to solar PV installation. Those individuals who would like to install PV, but cannot because they aren't credit worthy, ultimately do not install because of this financial hurdle that they cannot surpass. In my research I will be assuming that individuals do not face this kind of constraint, as data is unavailable for this purpose in the models I hope to use.

In summary, while the literature is extensive regarding financial benefits to consumers, there is a gap in empirically examining the social spillover effects of solar panel installation, as well as a lack of research on electricity pricing affecting installation rates, which this paper hopes to examine. This paper outlines the proposed research and beginnings of a theoretical model for regression.

3. Theoretical Framework

Past models that look at solar installation rates have only predicted the installation rate of solar hot water heating systems, and not photovoltaics. Durham et al. (1988) develop a probit model that predicts the probability of installing solar hot water heating:

$$SOL = \beta_0 + \beta_1 LTC + \beta_2 EP + \beta_3 SRA + \beta_4 INC + \beta_5 EDU + \beta_6 SEP + \beta_7 NR + \varepsilon \quad (1)$$

where the variables used are listed below in Table 1.

Table 1. Variables Used in Durham et. al Analysis.

Variable	Description
SOL	Dummy variable: 1 for install, 0 for not install;
EP	Energy price faced by the household;
LTC	Level of state tax credit;
SRA	Solar radiation available to the household;
INC	Household income
EDU	Level of educational attainment (3 dummy groups);
SEP	Perceived seriousness of global environmental problems;
NR	Number of residents in a household.

The authors cite consumer maximization theory as the reason for including economic variables such as tax credits and energy pricing, and extend their model into social variables such as education to observe consumer responsiveness to changes in these economic variables. I use their model to draw ideas on the basic variables that this model uses, as I believe they can be extended to solar PV panels. This is because solar PV panels are comparable to solar hot water heating systems, as they are both similar technologies that are capital intensive at the purchase point, with future returns extending over the lifetime of the unit and low maintenance costs.

Using the framework for inputs provided to Durham et al. (1988), the variables I use in my analysis are summarized in Table 2.

Table 2. Variables for Analysis.

Variable	Abbreviation	Description	*Mean*	*Std. Dev*	*Min*	*Max*
Electricity Price	*EP*	Average price of least cost available electricity source, ($/mWh), real; deflated using CPI Energy Index);	85.43	15.0619	38.36872	118.8473
Some College	*somecoll*	Percentage of people in a given location with some college education;	0.5456	0.1899	0	1
Cumulative Installs	*ccount*	All installs in a given location up to but not including the current period;	3.593	12.184	0	444
Rich	*rich*	Household Income (divided by Population)				
		$125,000 or more	0.09275	0.1018	0	1
Electricity Demanded	*edm*	Annual average of electricity demanded, extracted by county and then expanded to zip code level (kWh/day);	0.2660	0.08304	0.17583	0.695
Old	*old*	Percentage of households that have household owner older than 45 years	0.5732	0.1369	0	1
Incentive	*avincent*	Imputed average level of policy incentive, (whether it be in the form of a tax credit or installation rebate, real$ deflated with CPI);	1129.72	2790.876	5753.42	22241.83
Non-White	*nonwhite*	Percentage of people in given location who identify with race other than white;	0.3026	0.2131	0	0.95226
Population	*pop*	Number of people in a given zip code location.	20397.28	20690.57	6	105277

I have chosen to look at education level, because I believe that in many ways it is a proxy for awareness and perceived seriousness of global warming. Those people with a higher level of educational attainment are more likely to understand the dynamic relationship behind atmospheric warming and ecological damage, as well as having a greater understanding of the role they have in affecting global emissions. The level of income is essential for this analysis as PV has such a high initial cost, those individuals with larger incomes have sufficient disposable income to purchase PV panels or have the financial stability to secure a loan that will finance their purchase. The age variable is present because younger people may have a greater concern

for the condition of the global environment, thus they are more likely to install solar PV (if they have the financial capital to do so). Population is included to control for rural versus urban areas.

Variables that have been excluded from the final analysis are solar radiation available, income groups other than *rich*, and cost of installed panels. As the study occurs in California, solar radiation is already higher on average compared to the rest of the continental United States (Roberts 2008). Therefore, it is likely that people who live in California consider themselves as having higher than average sunshine, and the variance of solar radiation that does exist throughout California is trivial in regards to the decision of installing solar PV panels. If this study were applied across a broader geographical area solar radiation would certainly play a larger role. But for present purposes, this variable was observed as statistically insignificant (with a negative sign) and has thus been excluded from further analysis. All income bucket groups other than *rich* proved to be insignificant. These variables were dropped from the regression and had little impact on all other coefficients in the model. I hope that the *rich* will capture the notion that PV is a large sunk cost. Instead of including the average installed cost in my regression, the income variable will identify if the individual is wealthy enough to incur the initial sunk cost, and from that point on will only consider the cost of grid electricity versus the markedly cheaper solar generated electricity.

The cost of installing the panels has been excluded because of concerns that the variable is endogenous. If many people in a given location install solar panels the price may increase if the supply curve is upward sloping; conversely, if more panel providers may enter the market in that location and the price will begin to fall. Economies of scale are likely to develop in areas where installations have been quite prevalent in the past. Wiser et al. (2009) conduct a comprehensive study of the costs associated with solar PV panels from 1998 to 2008, and show

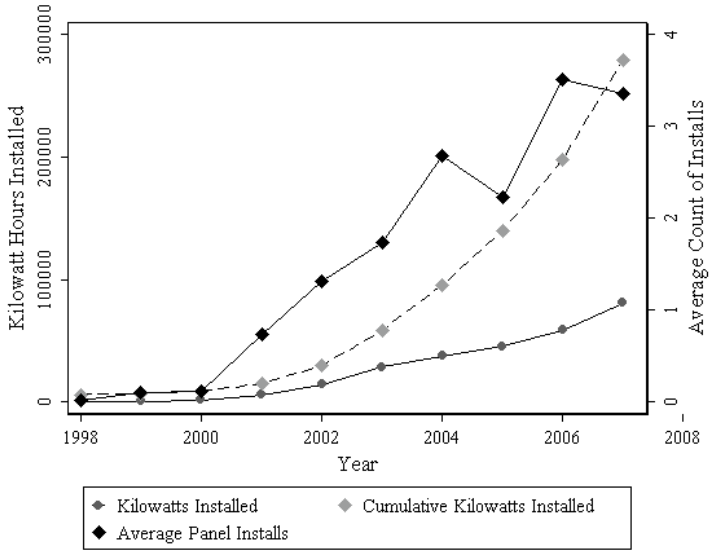
that the more recent cost reductions are likely due to decrease in module costs rather than changes in installation costs, because of the developing economy of scale for PV panels. Therefore there is a relationship between the number of installs and the price of installation. In particular, there are roughly 227 residential installers in California, and many of these companies are situated in and around the larger cities (such as Los Angeles and San Fransisco) that have seen a very large number of installs. Because more companies results in increased competition, it is very likely that these companies in the areas with lots of installs charge less than those installers who have a monopoly over the residential market in sparsely populated areas such as Alpine county (located in the mid-east of the state). Also, particular providers may charge more in traditionally wealthier areas (such as Orange County). As I have suspicions that the price of installs is endogenous, I feel that it is prudent to exclude it from the regression at this time.[2] As I am looking at discrete count data in a given location that is not a dummy variable, I cannot use the same probit model that Durham et al. have developed. However, my model will partially use the probit model as the data contains such a large amount of zeros. I combine a probit probability with the negative binomial model in order to gain results that best fit the observed probability of installation. Traditionally, the Poisson model is used to analyze count data, but the negative binomial distribution is more appropriate in this case as it allows for the variance of the data to be greater than the mean (a prominent feature of the installation data). Negative binomial takes such "overdispersion" of the variance into account by relaxing the Poisson assumption that the condition mean is equal to the variance. Like Durham et al. I will use this model to interpret the signs of the coefficients.

[2] Techniques that deal with endogenous variables (such as instrumental variables) are often quite difficult to implement in the type of count model I am going to estimate (the negative binomial), and are not attempted in this research.

4. Data

Following Germany and Japan, California is the third largest solar PV market in the world (Wiser et al. 2007). Thus, California is a prime location from which to extract data for this research. As previously mentioned, California has 65 percent of all PV in the United States. Some suggested reasons behind this high proportion include the temperate weather in California (which lends itself to high amounts of solar radiation availability), and the California Electricity Crisis of 2001 that saw rolling blackouts and brownouts sweep the state. Some have suggested that because of this crisis, Californians are more risk-averse and have actively sought out a stable method of electricity supply. California has long been a leader in environmental activism because it has many pressures on land use, concerns of acid rain and conservation of its famous coastline (Houriuchi 2006). California has numerous organizations with comprehensive databases that lend themselves to this type of research, such as the California Independent System Operator (CAISO), the California Energy Commission (CEC), California Public Utilities Commission (CPUC), and Go Solar California. Figure 2 shows the average rate of installation in each zip code location, as well as the total yearly and cumulative kilowatt hours installed in California.

Figure 2. Annual Kilowatts, Cumulative Kilowatts, and Average Count Installed in California.



The time period I am analyzing is between 1998 and 2007. I have chosen this period because the first clear increase in residential solar panel installations were in 1998, with growth in the industry continuing to present day. The installation data was extracted from the CEC website. Each installation point had the given incentive and its zip code location, Initially, 2008 was included in the data set. However the data set is incomplete at this point, with less than half of the actual installations reported in the CEC data, and thus was removed from the data set. Each data piece contains the city and zip code where the panels were installed, the size of the unit (in watts), the utility provider's name (for example PG&E), the installed price (in nominal dollars), the incentive amount (in nominal dollars), the date the incentive was approved, and the date the incentive was completed. The incentive amount was deflated using the CPI to obtain real dollars.

Figure 2 shows that the first clear period of installation is between 1998 and 2007 in California; therefore, if I hope to gain the most complete picture of installation behavior, this research will look primarily at this time period. Van Benthem et al. (2008) contend that the

reason behind solar PV growth in California after 1998 resulted from the implementation of two state incentive programs. These are the solar rebate program – when a consumer is compensated per installed Watt – and tax credits, which the state pays for a percentage of the installed cost of the PV system.

Initially, I had hoped to find data that was in a survey format, so that each data piece would pertain to a single installer. If this were available, I would have information on each individual's characteristics as well as the specifications on their PV unit. However, as these data do not currently exist, the average characteristics of individuals in the smallest possible area demarcation in California were used instead, and attach these statistics to known frequencies of PV installation. So, by defining the sample groups by zip code area, I believe that I am able to create the most comprehensive picture of a solar PV installation decision. This is because the inputs range from economic data (income, electricity prices) to social history (education levels), and the only clear means to gather a standardized form of each data set is by local zip code (or conglomerated zip code area) records. Therefore, each of these variables has been obtained for each zip code location. This has allowed me to build a profile of a typical PV installer in California.

To create the average electricity price variable, I obtained the revenue and sales data for all utility providers in California for my specific time period. The data were deflated using the Consumer Price Index for Energy (CPIENGNS), made available through the EIA. I took the given sales revenue of each company and divided this by the number of megawatt hours (mWh) sold in a given month to generate Figure 3 below.

Figure 3. Utility Pricing in California, 1998 to 2007.

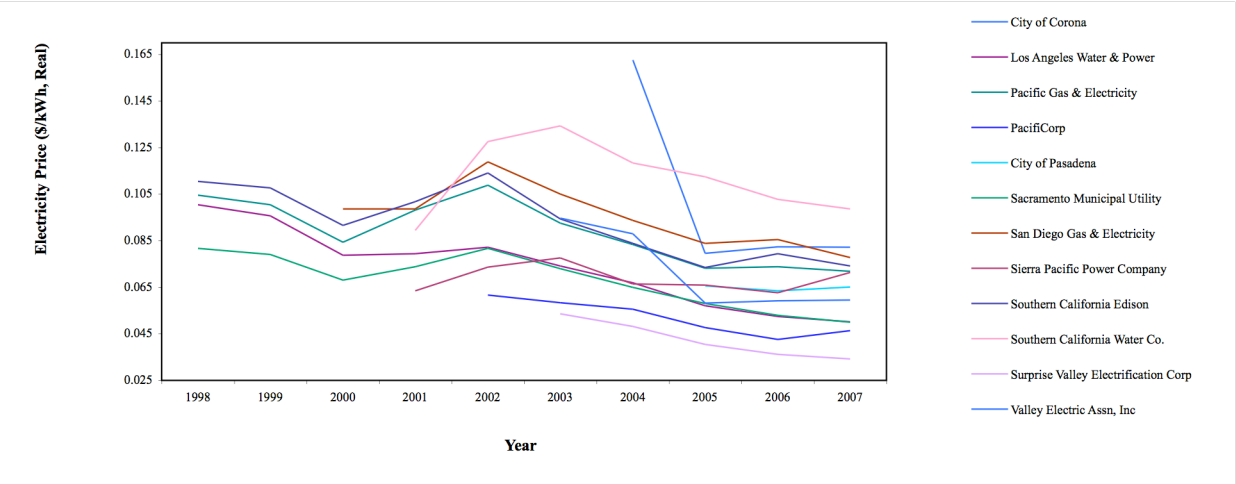


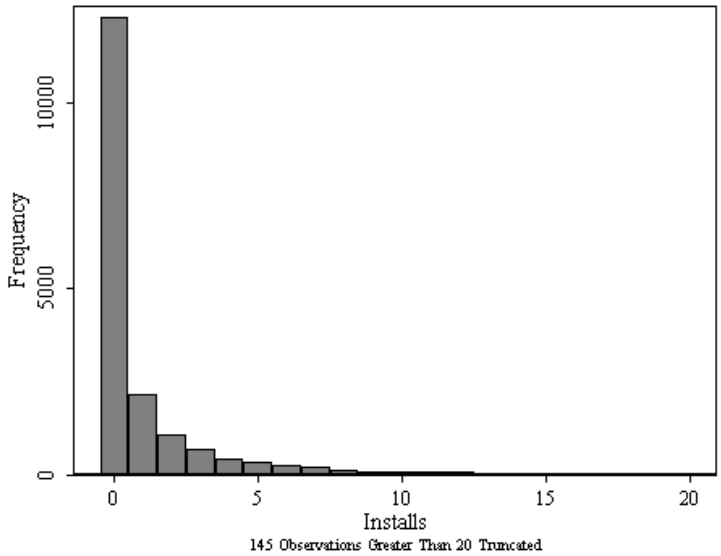
Figure 3 shows all of the pricing schedules in real terms for California over my specified time period. While I understand that electricity companies do not change their rate structures this frequently, this is the only historical data available for my analysis, and it builds a strong picture of how electricity prices in California have been moving in this time period. I will assume that when the electricity companies do change their prices, they create an average yearly value from which my data is created.

In the period of 1998 to 2007, 1206 of the 1660 zip codes that I am working with saw at least one instance of solar PV panel installation. The total number of installations seen is 26,296. For any year period the minimum and maximum number of installations are 0 and 243 respectively. The maximum cumulative installation for any zip code location in 2007 is 444. The variance of yearly installations is 28.21, and the mean is 1.58. Clearly, the variance of installations is much greater than the mean installations in any given year.

Figure 4 shows the frequency of installs in California. So that the data is clear, those zip codes with greater than 20 installations in a year were truncated for this graph. Each of these 145 truncated data points had a frequency of one install. As you can see, the data contains a large amount of zeros. It is for

this reason that a single negative binomial regression is not used. It is more logical to predict the probability that any location will ever install panels, and then to use the negative binomial regression to model the remaining count data for locations that see at least one install.

Figure 4. Frequency of Observed Install Count.



The education, income, age, and race data was drawn from the 2000 Decennial Census. There is no census data available for the different years in my time period, so I have to assume that the demographics of each zip code location do not change dramatically between 1998 and 2007, and that they are in fact constant at 2000 levels.

The education dummy group *somecoll* was created by summing all persons who indicated that they had attended an associate, bachelor, masters, professional or doctorate degree. Then, this value was divided by the total number of people in that zip code so that each parameter would illustrate demographics and the larger zip codes would not be over-weighted. Similarly, the *old* variable was created by summing all people in given location that were older than and including the age of 45, and divided through by population total. *Nonwhite* was calculated by subtracting the number of people in a given location who identified themselves as white on the 2000 census from the total population in that area, then dividing by total population. This

remaining group of people was also equal to the sum of individuals who identified with any race other than white. The income buckets were already defined by $25,000 groups at the higher levels of income (greater than $75,000) by the Census questions. So, the *rich* variable was created by summing all the income groups above $125,000. Again, this is represented as a percentage of total population.

The data set I created represented each zip code in each year, resulting in a total of 16600 rows (10 times 1660 zip codes) of information. Although there are more than 1660 zip codes in California, I removed those from the data set that applied only to businesses (for example, 96XXX), those with no data in the 2000 census, or those with a recorded population of zero. The number of installs for each zip code in each year was entered. The cumulative installs for each zip code was calculated by summing the number of installs seen in that location in all previous years of the examined time period (not including the current year, as cumulative installs looks to explain the impact of previous installs on current installs, and cannot do so if current installs is included).

However, as most of the zip code locations did not see any installations in the earliest years (and some even in the later years) I had no data on the incentive level seen in these locations. The legislation on the level of state and federal tax credit data has been incredibly hard to pin down. Along with the state-wide incentive programs that provide a blanket rebate for solar installation, there are many intricacies of local policy that feed into the total rebate that a given household will receive. For example, there are addition programs that exist in the county of Anaheim, in the Los Angeles metro area, and Sacramento municipal districts, that vary by the size of installation and the metropolitan location of the individual installer (Wiser 2009). Each of these programs adds some factor to the level of incentive that is not clear in the historical data.

Ultimately, I was unable to identify the correct incentive level in each zip code location for the given time period. Instead, I have used the CEC data to impute what the average incentive would be in any given location in any year.

Fitted values were used to impute incentive levels for locations without observed installs (394 locations saw no installations between 1998 and 2007). To generalize the likely incentive level a given zip code location in any year, the zip codes were first grouped into their larger county areas.[†] Dummy variables for each county and year were created and all the observations with zero installs dropped. By regressing each of the dummy variables (created for year and county) against the natural log of the observed incentive variables, coefficients were estimated using an OLS regression. From this, incentive levels for each zip code could be imputed using the regression output. From this output, each zip code now has an imputed incentive level regardless of whether installs were seen in that location.

5. Estimation

The model used is a combination of a probit model and the negative binomial regression. This type of model is commonly used to deal with data that has a large amount of zeros. Because only 5784 data points saw any installs in the period between 1998 and 2007 the data set contains a large amount of zeros. To prevent these zeros from masquerading as overdispersion in a negative binomial model, the data was first analyzed using a probit model to see the probability of a location either installing or not installing PV panels. Then, using the principles of joint probability it was found that

$$\Pr(Y_i = y_i \mid x_i) = \Phi(x_i\beta) \times \Pr(\widetilde{Y}_i = \widetilde{y}_i \mid \widetilde{x}_i) \qquad (2)$$

[†] There are 58 counties in California. Zip codes were attached to county locations using crosswalk data available through the MABLE/Geocorr utility: http://mcdc2.missouri.edu/websas/geocorr2k.html

where the model specifications are given in table 3 below.

Table 3. Model Summary.

Y_i :	Predicted installs given by probit model.
y_i :	Observed installs in full data set (zeros included).
x_i :	Explanatory variables of all locations.
β_i :	Coefficients of explanatory variables for probit regression.
$\Phi(\cdot)$:	Probit probability function for installing in any given location
$\tilde{Y}_i$:	Predicted installs for the negative binomial
$\tilde{y}_i$:	Observed installs, locations with zero installs removed.
$\tilde{x}_i$:	Explanatory variables for truncated data set (conditional upon positive installs).
$\tilde{\beta}_i$:	Coefficients of explanatory variables for negative binomial regression.
$\Pr(\tilde{Y}_i = \tilde{y}_i \mid \tilde{x})$:	Negative binomial probability (conditional upon positive installs).
μ_I :	Conditional mean of the negative binomial.
α :	Overdispersion parameter for negative binomial.

The problem of excess zeros in count data is discussed in detail in the literature, and is a well known problem that affects certain types of count data analysis. In many economic studies, a zero data point can reflect a corner solution in an economic choice model (Winklemann 2008). This is true for the solar installation data set. If a zip code location sees no installations in a given period it may be that their optimal number of installations is negative. However, a negative installation is not possible, and the data point is simply represented as zero. In these cases the process that generates the zeros is different from the process creating the strictly positive count data. Excess zeros affect the precision of inference. This model does not empirically address perceptions on future technological improvement. As the literature suggests, it is assumed that individuals understand the progress of solar technology in years to come.

Using the probit results in tandem with the negative binomial regression prevents the zero

observations from giving bias to the regression results. Intuitively, approaching the data in this way makes sense. If the excess zeros were included in a single negative binomial model, the model would need to compensate for its extreme skewed nature, and thus the coefficients would not estimate a viable number of installs. This single negative binomial was tried, and when the data was reentered into the model roughly 270,000 installs were estimated; almost ten times the observed value. The probit model accounts for the fact that for the most, zip codes are likely to see zero installations. Then, the negative binomial determines the likely count of installs in any location given that installs is not zero.

i. Probit Estimation

The probit model is given as

$$E(Y_i \mid x_i) = P(Y_i = 1 \mid x_i) = \Phi(x_i\beta) \tag{3}$$

where Y_i is a dummy variable created for each location in each year (1 = installations observed, 0 = no installations observed). Using the cumulative normal distribution Φ a probability is found for seeing installs in any given location. All the data points were used in this regression. The results from this probit regression are given below in Table 4. The complete specification of the probit model is given in equation 4 below.

$$\begin{aligned} P(Y_i = 1 \mid x_i) = \Phi(&\beta_0 + \beta_1\text{pop} + \beta_2\text{edm} + \beta_3\text{ccount} + \beta_4\text{ccount}^2 + \beta_5\text{EP} + \beta_6\text{EP}^2 \\ &+ \beta_7\text{EP}^3 + \beta_8\text{avincent} + \beta_9\text{somecoll} + \beta_{10}\text{rich} + \\ &\beta_{11}\text{old} + \beta_{12}\text{nonwhite}) \end{aligned} \tag{4}$$

ii. Negative Binomial

To use the negative binomial regression, all data points with zero installs (10816 of 16600) were removed from the data set. The data that did see observations remained overdispersed (variance is greater than the mean). The negative binomial regression is considered to be the

standard method used to analyze overdispersed count data (when the variance of the count is greater than the mean). Overdispersion of the truncated data set was confirmed by a Z-test, which had a value of 6.645 and a P-value = 0, and the test is significant. Thus, the hypothesis of no overdispersion in the data is rejected. The negative binomial model estimates a conditional mean μ_i through which the covariates $\tilde{x}_i$ are introduced in $\mu_i = e^{\tilde{x}_i \tilde{\beta}}$. This is the expected mean of the data (average expected installs). Count data are discrete values given as $y_i = 0, 1, 2, \ldots$ and the expected value of installs given the set of covariates is

$$E(\tilde{Y}_i = \tilde{y}_i \mid \tilde{x}_i) = \mu_i \tag{5}$$

The variance of a negative binomial model is given as

$$Var(\tilde{Y}_i) = \mu_i(1 + \alpha\mu_i) \tag{6}$$

where α is identified as the "overdispersion parameter". The most commonly seen form of this negative binomial production function is

$$\Pr(\tilde{Y}_i = \tilde{y}_i \mid \tilde{x}_i) = \frac{\Gamma(\tilde{y}_i + 1/\alpha)}{\Gamma(\tilde{y}_i + 1)\Gamma(1/\alpha)} \left(\frac{1}{1+\alpha\mu_i}\right)^{1/\alpha} \left(1 - \frac{1}{1+\alpha\mu_i}\right)^{\tilde{y}_i} \tag{7}$$

which uses observed covariates $\tilde{x}_i$ and the overdispersion parameter α to estimate coefficients using maximum likelihood. A further discussion of how the negative binomial model is derived appears in Appendix A.

To predict the expected count of installs in the truncated data set – and thus the coefficients of the regressors – the following form was used:

$$\begin{aligned} \ln(E(\tilde{Y}_i = \tilde{y}_i \mid \tilde{x}_i)) = \beta_0 &+ \beta_1 \text{pop} + \beta_2 \text{edm} + \beta_3 \text{ccount} + \beta_4 \text{ccount}^2 + \beta_5 \text{EP} + \beta_6 \text{EP}^2 \\ &+ \beta_7 \text{EP}^3 + \beta_8 \text{avincent} + \beta_9 \text{somecoll} + \beta_{10} \text{rich} + \\ &\beta_{11} \text{old} + \beta_{12} \text{nonwhite} \end{aligned} \tag{8}$$

and the results given below were obtained using Stata's *nbreg* command. Also given below are the results from the probit regression.

Table 4. Dependent Variable: Installs.

Variable	Abbreviation	Probit	Negative Binomial
Population	*pop*	1.13×10^{-5} **	8.37×10^{-6} **
		(7.72×10^{-7})	(1.08×10^{-6})
Electricity Demanded	*edm*	-0.53924**	0.66668**
		(0.19924)	(0.22734)
Cumulative Installs	*ccount*	0.26602**	0.04440**
		(0.01083)	(0.00186)
Cumulative Installs Squared	$ccount^2$	-0.00059**	-0.00010**
		(2.42×10^{-5})	(1.02×10^{-5})
Electricity Price	*EP*	1.1865**	0.81019**
		(0.15382)	(0.148)
Electricity Price Squared	EP^2	-0.01309**	-0.00892**
		(0.00170)	(0.00165)
Electricity Price Cubed	EP^3	4.76×10^{-5} **	$3.22 \times 10^{--5}$ **
		(6.26×10^{-6})	(6.06×10^{-6})
Incentives	*avincent*	0.00013**	5.52×10^{-5} **
		(4.33×10^{-6})	(5.99×10^{-6})
Education: Some College	*somecoll*	0.76539**	0.28281*
		(0.09255)	(0.14246)
Income: $125,000 or more	*rich*	0.64292**	0.93277**
		(0.15461)	(0.16806)
Age: Older than 45	*old*	0.25964*	-0.43989*
		(0.10251)	(0.2213)
Non-White	*nonwhite*	-0.50189**	-0.69781**
		(0.09012)	(0.12396)
Constant	β_0	-38.42542**	-24.149**
		(4.5621)	(4.389)
Alpha	α		0.28942**
			(0.0192)
Number of Observations		16600	5784
Pseudo R^2		0.4014	

NOTE: Standard errors are in parentheses. Marked ** if significant at the 99% level, * if significant at the 95% level. Negative Binomial log pseudolikelihood = -12782.336, Prob>Chi^2 = 0, Wald $Chi^2(13)$ = 1961.83.

The statistical significance of the α term again shows us that the installs data is overdispersed and the negative binomial regression is appropriate for this research. All coefficient estimates are of the expected sign, and all are significant at the 99% level (except for Education and Age,

which were significant at the 95% level). The significance of these results indicates that each of the chosen independent variables explain the propensity to install solar panels in California.

All income groups below $125,000 were dropped. These groups were insignificant when regressed separately and in any combination of income buckets, and had illogical signs. Many combinations of these bucket groups were tried and this final group of $125,000 and over was found to be the best fit for the data.

Preliminary regressions included solar radiation and price of installation variables. The solar radiation variable was discarded due to insignificance and the illogical negative sign of the coefficient. As discussed in section 3, the price of installation variable was dropped due to concerns of endogeniety.

Interpreting these coefficients is difficult because the probit and negative binomial must be combined to produce the predicted installs. The most appropriate interpretation comes from looking at the magnitude and size of each pair of coefficients (for the same variable, one from the probit and one from the negative binomial). The results in Table 4 show that the propensity to install panels increases with population.

The cumulative count variable shows that the number of previous installs influences future installation decisions for any given location. This is because of social spillover. If an individual sees their neighbor benefit from panel installation, they are more likely to install panels. The cumulative count variable was included as a squared term to see the rate at which previous installs was affecting current installs. Previous installations in a given location are likely to affect an individual's decision to install panels because they are more aware of PV power as a viable option. Also, there is likely to be a local installer available to them, as well as a neighborhood social network that can help educate individuals on the benefits of PV panels. As

expected, cumulative installs increase the count of current installation at a decreasing rate. This supports the findings in the literature; people are more likely to install when they have seen other people around them do so first.

The conflicting signs for electricity demanded and older than 45 are most interesting. *Old* seems to suggest that if a location has a higher percentage of people over the age of 45 this location is less likely to install, but if other effects are dominant and the location does see at least one install, it will see a large number of installs (not just one). This may be explained by technology uncertainty: older individuals are less comfortable adopting a new technology that they are unfamiliar with, because this older generation is not as adaptable nor are they as used to rapid technology change as younger individuals. Thus, it may be so that once a single installation occurs in this location, the older inhabitants become familiar with the technology and are prepared to invest in solar panels. Similarly, electricity demanded may have conflicting signs because locations with large electricity demands are hesitant to install panels because they are unsure that the technology will support their current needs. However, when they do install and are able see how solar panels provide sufficient electricity that is cheaper than grid-electricity, they install at a greater rate than those locations that don't have high electricity demands.

As expected, the *rich* group has positively signed coefficients. This indicates that as a location becomes wealthier individuals there are more likely to install solar PV panels. This is because more people have greater disposable income and are able to surpass the initial sunk cost incurred when installing panels.

The coefficients on the electricity price variable are both positive and significant. This is as expected – if electricity prices rise, individuals will seek alternative methods of procuring electricity. This supports many of the findings in the literature. The incentive variable also has a

coefficient that is positive and significant. If the government provides a large incentive there will be more installations. This is similar to a reduction in purchase price of panels. Any given amount of additional incentive may push a number of individuals over the marginal threshold and they will purchase solar panels. This result is important for the policy discussion in section 6.

6. Policy Implications

From the results obtained in table 4, a marginal cost of abatement curve for 1998 can be derived; this curve measures at the marginal cost of carbon reductions taking into account the resulting carbon abatement in the ensuing 10 years. To do this, characteristics of solar panel conversion were needed. Silverman (2008) asserts that the typical installed panel system produces one-fifth of its total capacity. The data shows that the average installed size of panel is 3.023kWh (with a standard deviation of 0.59). Therefore, the average panel produces roughly 0.6kWh power daily. Per year, this becomes 5400kWh. Silverman (2008) also shows that the average household uses 7350kWh power a year, meaning that solar panels cover on average 73 percent of a household's electricity needs. So, if each panel installed prevents 5400kWh from being generated using grid-power, then 0.952543kg per kWh of CO_2 is not emitted (EIA 2000).

Using this information in an Excel spreadsheet, the marginal cost of abatement (in tons of CO_2) was generated. The marginal cost of abatement curve shown in Figure 7 was created by implementing a range of incentives (from an additional $100, to $1000) during 1998. The cumulative effects are considered in this graph; that is the ensuing additional installs that result because of the increased installations in 1998 for the next 10 years are included in the calculation of abatement resulting from the given incentive increase. The 10 year period was used because this is the period the data spans. Also, the abatement was calculated in terms of 2007 total abatement. As the panels in 1998 would be in operation for 10 years, their annual abatement

level is multiplied by 10, and so on. A discount rate of 7 percent was used to see the cost in terms of 1998 dollars. This rate of 7 percent was chosen because it guarantees that costs in the future are small, meaning later annual incentives have a lesser cost than one large initial incentive. As the outcomes will show that a large initial incentive is the most cost-effective policy, using a large discount rate prevents the most recent incurred costs from being over-weighted, and shows that using an initial incentive is the best policy in all conditions.

Figure 5. Marginal Cost of Abatement 1998.

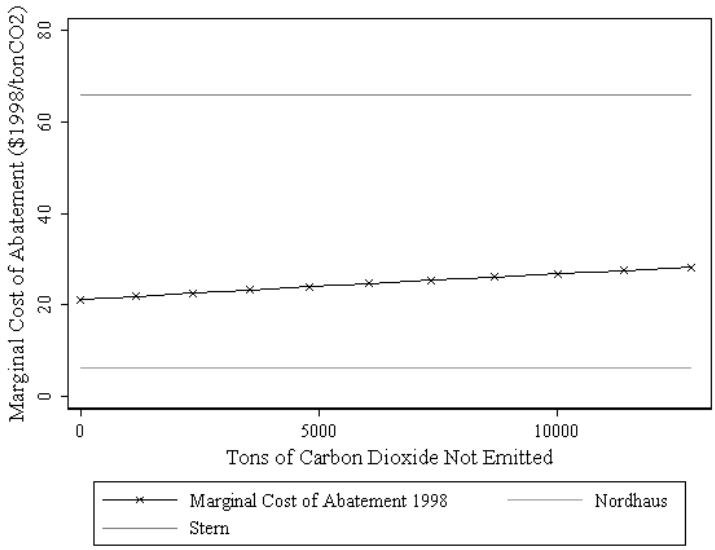
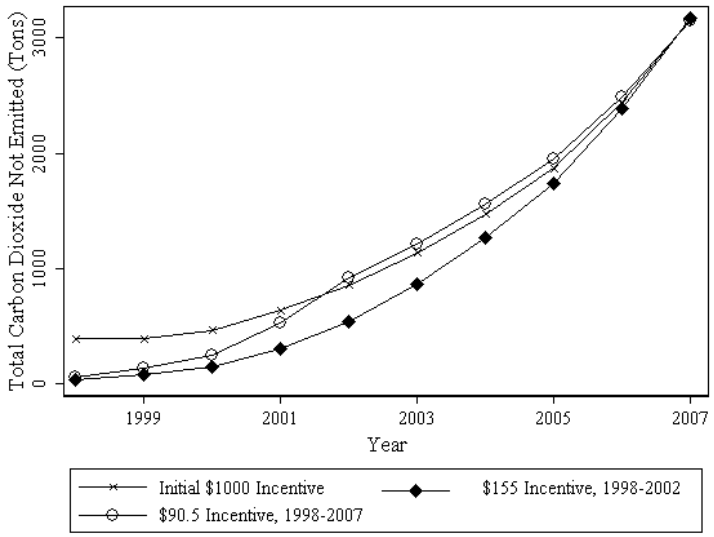


Figure 5 shows that in 1998 the cost of abating the 5000th ton of carbon is roughly $23. Included on this graph are upper and lower bounds for the social cost of carbon emission that are most prominent in the literature. Nordhaus (2008) asserts that the social marginal cost of emitting CO_2 is $6.18 per ton in 1998; this number is generally accepted to be at the lower end of social cost estimates in today's literature. On the other hand, Stern (2007) gives the higher bound as he estimates the social cost of carbon to be nearly $65. It is clear that in 1998, the marginal cost of abatement using solar panels is within the realm of generally accepted social costs. Therefore, it is valid to suggest that the government look at solar panels as a method of abating CO_2 from the

atmosphere.

The most notable policy implication that the regression results show is the effect of previous installs in a given location. Figure 6 shows that if a $90.50 incentive is given each year for 10 years, or $155 for the first 5 years, then these incentive structures will ultimately yield the same additional installations than if a single $1000 incentive were given in the first year. This is because the additional installers in the first period will influence additional installers in the subsequent periods, meaning that the larger the upfront incentive the more installers in the subsequent period all else equal. The coefficients recovered in the estimation of this model show that there is a clear effect on installation through the previous installations seen in a given location. Looking at Figure 6, we can see this effect in action. The one-time $1000 incentive leads to 170 additional installs in 2001 that would not have occurred without this incentive. This data is summarized in Appendix B.

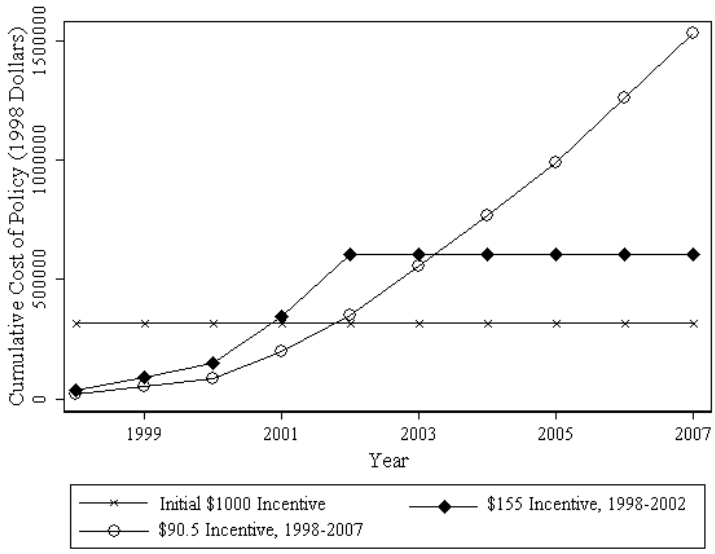
Figure 6. Differences in Incentive Timing and its Effect on Installs.



However, even though the $90.50 incentive every year yields the same absolute installations, it also costs significantly more than the $1000 initial incentive. This is because the government

must pay $90.50 to each installer in every period, including the additional installers. If the government were to only pay the initial $1000 incentive, this would cost $317,923. If the government implemented the annual incentive, they would face an overall cost of $2,573,036. For the same number of panel installations, the government would have to pay an extra $2,254,112 if the incentives were given each year. This cost difference is shown in figure 7 below. It is clear that the government would be best off utilizing these social spillover effects by giving a large incentive early to considerably cut the cost of achieving a certain number of installations.

Figure 7. Cumulative Costs for Different Incentive Structures.



This graphs shows that in 2007 the total cost of the annual $90.50 incentive costs almost four times as much as the single $1000 incentive. Clearly, the most efficient policy is to provide a one-time incentive if the aim is to maximize carbon abatement in a ten-year period.

This argument is even stronger if carbon dioxide is not considered as a 'stock pollutant' – meaning that we value carbon abatement sooner rather than later (and it does matter when this carbon is abated). Figure 8 shows the progression total carbon abatement up to 2007 if we

consider the installs in 1998 to abate 10 times as much as the installs in 2008 because they have 10 years in which to operate.

Figure 8. Carbon Abatement Considered as Time Sensitive.

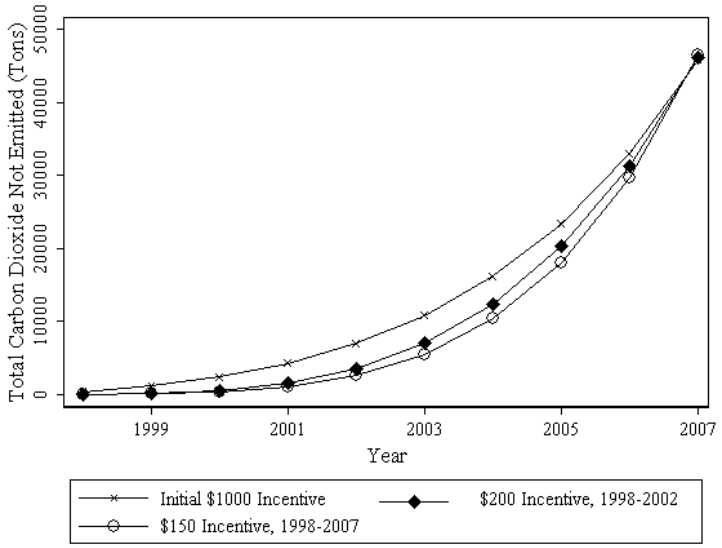
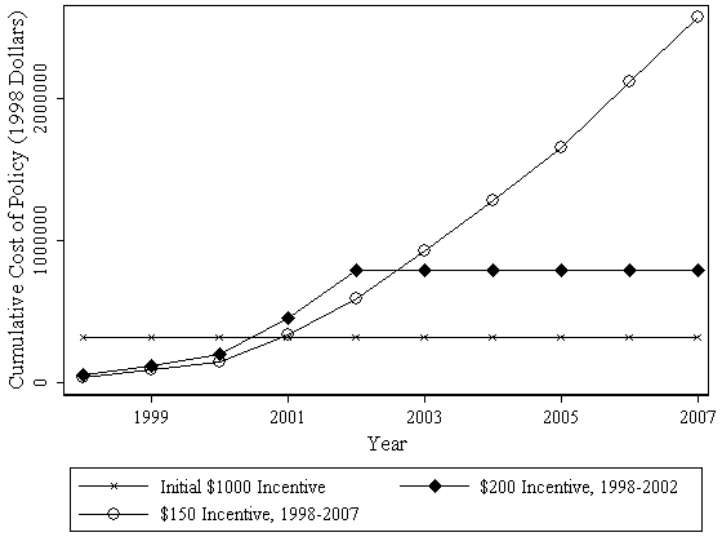


Figure 8 shows that if the government wants to achieve carbon dioxide abatement of around 4500 tons in 2007 they can either choose to implement a $1000 incentive at the start of the 10 year period, or a $200 incentive each year, or a $150 incentive for the first five years. Figure 9 shows the costs of these different incentive strategies.

Figure 9. Cost of Policies When Carbon Abatement Considered as Time Sensitive.



Here we can see again how the costs of implementing an annual policy are much more expensive to the government than using a large upfront incentive that will have spillover effects in a given location.

7. Conclusion

In this paper the effects of current electricity price, incentive levels, previous installations and personal characteristics on the rate of solar PV panel installation in California were recovered in order to understand policy options for solar panel installation in California. Using data separated by zip code locations over a period of 10 years, a combined probit and negative binomial regression shows how chosen input variables influence the rate of solar photovoltaic installation. If the consumer faces a high electricity price they are more likely to install PV panels. This is because they are seeking an alternative method of acquiring power in their homes, substituting away from the source that is most expensive to the cheaper solar power alternative. As past literature had indicated, the cumulative count of installs in a given location has a positive impact on the count of new installs in the current period. This work examined this phenomenon

empirically, and proved it to be true. Also, higher incentives will lead to a higher rate of installation as this will effectively decrease the disposable income needed to cover the sunk cost incurred when the panels are installed. Locations with higher education levels and income were found to be more responsive to installing solar panels.

From the model output a marginal cost of abatement curve for 1998 was recovered. This curve shows how much each additional ton of carbon dioxide will cost if it is to be abated using solar panels. The abatement in the 10 years following 1998 are included in the marginal cost of abatement curve because the model shows the significant effect that previous installations have on future installations in a given location. This marginal cost curve shows that carbon abatement can be achieved using solar panels as the curve falls between the upper and lower bound of social cost of emissions as determined by Stern (2007) and Nordhaus (2008), two renowned environmental economists in the area of climate change.

The most pertinent policy implication seen from these results is the effect of cumulative installs on the timing of incentives. If a large incentive is given in the first period, the additional installs that result from this incentive will cause installs to increase in subsequent periods. Giving a single large incentive is preferable to giving smaller incentives over a large period of time because the cost of repeated incentives is much greater than that of a single large incentive.

As California's installation pattern is so different to the rest of the United States, these policy outcomes cannot be directly extended to other locations at this time. However, this research provides unique insight into the generally dynamics of solar photovoltaic installation, and the positive effects that previous installs have on the future rate of installation.

Appendix A. Negative Binomial Regression

The negative binomial distribution is derived from a Poisson-gamma mixture model. The Poisson regression is the customary method used to model count response data. However, the Poisson distribution assumes that the mean is equal to the variance of the data – a property that is rarely found in sample data. For a Poisson model, the production function is given as:

$$\Pr(Y = y \mid x) = \frac{e^{-\mu}\mu^{y}}{y!}; \quad y = 0,1,2,\ldots; \qquad \mu > 0 \tag{9}$$

where the parameter μ is the fitted mean of the model, and y is the count response. The relationship between μ and the model covariates x_i is parameterized such that $\mu = \exp(x_i\beta)$. To estimate the β coefficients the Poisson log-likelihood function is used:

$$\ell(\mu; y\ ) = \sum\{y_i \ln(\mu) - \mu - \ln(y_i!)\} \tag{10}$$

substituting in, we can see that

$$\ell(\beta; y_i) = \sum\{y_i(x_i\beta) - \exp(x_i\beta) - \ln\Gamma(y_i + 1)\} \tag{11}$$

where $y!$ can be calculated in terms of the log-gamma function $\ln\Gamma(y_i + 1)$. The value of β that maximizes the log-likelihood function is the estimated coefficient vector β. To see this we take the first derivative with respect to β:

$$\frac{\partial(\ell(\beta; y_i))}{\partial\beta} = \Sigma(y_i - \exp(x_i\beta))x_i \tag{12}$$

and set it equal to zero, so that

$$\Sigma(y_i - \exp(x_i\beta))x_i = 0 \tag{13}$$

gives us the solutions to our parameter estimates for the Poisson model.

If the data has a variance greater than the mean it is known as "overdispersed". The negative binomial regression is a standard method used to model this type of overdispersed count

data. The negative binomial can be thought of as an extension to the Poisson model that accounts for the overdispersion of the count variable.

$$\Pr(Y_i = y_i \mid x_i) = \frac{\Gamma(y_i + 1/\alpha)}{\Gamma(y_i + 1)\Gamma(1/\alpha)} \left(\frac{1}{1 + \alpha\mu_i} \right)^{1/\alpha} \left(1 - \frac{1}{1 + \alpha\mu_i} \right)^{y_i} \tag{14}$$

where α identified as the "overdispersion parameter". Like the Poisson distribution the count data is $y_i = 0, 1, 2, \ldots$ and $E(Y_i \mid x_i) = \mu_i$, but the difference occurs in the variance which is seen to be $Var(Y_i) = \mu_i(1 + \alpha\mu_i)$. This is because $\mu = \exp(x_i\beta + \alpha)$ where in the Poisson model α was equal to zero. Thus the log-likelihood function can be written as

$$\begin{aligned} \ell(\mu; y, \alpha) = \Sigma \exp\{ & y \ln((\alpha\mu)/(1 + \alpha\mu)) - (1/\alpha)\ln(1 + \alpha\mu) \\ & + \ln\Gamma(y + 1/\alpha) - \ln\Gamma(y + 1) - \ln\Gamma(1/\alpha)\} \end{aligned} \tag{15}$$

Like with the Poisson model, by taking the derivate of this log-likelihood function and setting it equal to zero we can recover the parameter estimates β for the negative binomial model.

Appendix B. Incentive Tables.

Initial $1000 Incentive

Year	Additional Tons Abated	Cost (per Year)	Cumulative Cost	Cumulative Tons Abated
1998	388.4801095	317923.8955	317923.8955	388.4801095
1999	0	0	317923.8955	388.4801095
2000	74.90279478	0	317923.8955	463.3829043
2001	170.0837756	0	317923.8955	633.4666799
2002	220.0565398	0	317923.8955	853.5232197
2003	283.7809358	0	317923.8955	1137.304155
2004	335.5113792	0	317923.8955	1472.815535
2005	398.272455	0	317923.8955	1871.08799
2006	569.9058215	0	317923.8955	2440.993811
2007	719.0628758	0	317923.8955	3160.056687

$90.50 Incentive, 1998-2007

Year	Additional Tons Abated	Cost (per Year)	Cumulative Cost	Cumulative Tons Abated
1998	31.6830103	22494.54283	22494.54283	31.6830103
1999	45.77832138	29493.07326	51987.61609	77.46133168
2000	63.25286802	35121.89644	87109.51252	140.7141997
2001	164.7733102	111988.0453	199097.5578	305.4875099
2002	229.7706641	149987.0672	349084.625	535.2581739
2003	328.0134009	205265.4056	554350.0306	863.2715749
2004	398.4510613	210876.7367	765226.7673	1261.722636
2005	473.3827318	221500.1624	986726.9297	1735.105368
2006	650.2145196	273697.3204	1260424.25	2385.319888
2007	784.253712	267417.8514	1527842.101	3169.5736

$155 Incentive, 1998-2002

Year	Additional Tons Abated	Cost (per Year)	Cumulative Cost	Cumulative Tons Abated
1998	54.66700001	39219.15913	39219.15913	54.66700001
1999	79.02960617	51449.43418	90668.59331	133.6966062
2000	109.3165269	61365.91953	152034.5128	243.0131331
2001	283.9661136	194734.6491	346769.1619	526.9792467
2002	395.6012419	260696.1903	607465.3522	922.5804886
2003	288.907993	0	607465.3522	1211.488482
2004	342.0777083	0	607465.3522	1553.56619
2005	398.0886605	0	607465.3522	1951.65485
2006	541.2451873	0	607465.3522	2492.900038
2007	656.1849835	0	607465.3522	3149.085021

References

Borenstein, Severin. 2002. "The Trouble With Electricity Markets: Understanding California's Restructuring Disaster". *Journal of Economic Perspectives* (16): 191-211.

Borenstein, Severin. 2007. "Electricity Rate Structures and the Economics of Solar PV: Could Mandatory Time-of-Use Rates Undermine California's Solar Photovoltaic Subsidies?" *Center for the Study of Energy Markets, Working Paper 172*.

Borenstein, Severin. 2008. "The Market Value and Cost of Solar Photovoltaic Electricity Production." *Center for the Study of Energy Markets, Working Paper 176*.

Bradford, Travis. 2006. *Solar Revolution: The Economic Transformation of the Global Energy Industry*. Cambridge, Massachusetts: The MIT Press.

Bubenzer, A., and J. Luther. 2003. *Photovoltaics Guidebook for Decision Makers*. Berlin: Springer-Verlag.

CPUC. 2006. "The California Solar Initiative," Available at http://www.cpuc.ca.gov/static/energy/solar/aboutsolar.htm. Accessed 1 December 2009.

Durham, Catherine A., Bonnie G. Colby and Molly Longstreth. 1988. "The Impact of State Tax Credits and Energy Prices on Adoption of Solar Energy Systems." *Land Economics* (64): 347 – 355.

Dymond, C. 2002. *PV Focus Group Report*. Portland, Oregon: Energy Trust of Oregon.

Energy Information Agency. 2000. "Carbon Dioxide Emissions from the Generation of Electric Power in the United States." *Department of Energy, Washington DC.*

Energy Information Agency. 2009. *Official Energy Statistics from the U.S. Government, 2009*, available at http://www.eia.doe.gov/fuelrenewable.html

Faiers, A., and C. Neame. 2006. "Consumer Attitudes Towards Domestic Solar Power Systems." *Energy Policy*, (34): 1797-1806.

Fry, Gene R. Heinze. 1986. "The Economics of Home Solar Water Heating and the Role of Solar Tax Credits." *Land Economics* (62): 134 – 142.

Fuchs, Doris A., and Maarten J. Arentsen. 2002. "Green Electricity in the Market Place: The Policy Challenge." *Energy Policy* (30): 525-538.

Fujii, Edwin T., and James Mak. 1984. "A Model of Household Electricity Conservation Behavior." *Land Economics* (60): 340 – 351.

Heal, Geoffrey. 2009. "The Economics of Renewable Energy." *NBER Working Paper No. 15081*.

Hilbe, Joseph M. 2007. *Negative Binomial Regression*. New York: Cambridge University Press.

Houriuchi, Catherine. 2006. "California and the Implementation of Renewable Energy Technologies." *Sustainable Energy and the States: Essays on Politics, Market and Leadership,* edited by Dianne Rahm. North Carolina: McFarland & Company, Inc. 64-80.

Intergovernmental Panel on Climate Change. 2007. "Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change". New York: Cambridge University Press.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem". *The Review of Economic Studies Ltd* (60): 531-542.

Maslin, Mark. 2009. *Global Warming: A Very Short Introduction*. New York: Oxford University Press.

Metcalf, Gilbert E. 2009. "Tax Policies for Low-Carbon Technologies." *NBER Working Paper 15054*.

Nordhaus, William D. 2008. *Weighting the Options on Global Warming Policies*. New Haven: Yale University Press.

Nordhaus, William D. 2009. "Economic Issues in a Designing a Global Agreement on Global Warming." Keynote address Prepared for *Climate Change: Global Risks, Challenges, and Decisions*. Copenhagen, Denmark.

Papineau, Maya. 2006. "An Economic Perspective on Experience Curves and Dynamic Economics in Renewable Energy Technologies." *Energy Policy* (34): 422-432.

Roaf, Susan, and Rajat Gupta. 2007. "Solar Power: Using Energy from the Sun in Buildings." *Sustainable Energy: Opportunities and Limitations*, edited by David Elliot. New York: Palgrave Macmillan. 84-107.

Roberts, Billy. 2008. "Photovoltaic Solar Resource of the United States". *National Renewable Energy Laboratory*, Colorado.

Siebert, Horst. 2008. *Economics of the Environment*. New York: Springer. Seventh Edition.

Silverman, Dennis. 2008. "An Analysis of California's Million Solar Roof Initiative." Presentation for the California Electric Commission. Department of Physics and Astronomy, U.C. Irvine.

Stern, Nicholas. 2009. *A Blueprint for a Safer Planet*. London: The Bodley Head.

U.S. Census Bureau. 2009. *2006-2008 ACS 3-Year Accuracy of the Data (US).* Accessed 3 November, 2009: http://www.census.gov/acs/www/UseData/Accuracy/Accuracy1.htm

Van Benthem, Arthur, Kenneth Gillingham, and James Sweeney. 2007. "Learning-by-Doing and the Optimal Solar Policy in California." *The Energy Journal* (29): *Issue 3*.

Winkelmann, Rainer. 2008. *Econometric Analysis of Count Data*. Berlin: Springer.

Wiser, Ryan, Andrew Mills, Galen Barbose and William Golove. 2007. "The Impact of Retail Rate Structures on the Economics of Commercial Photovoltaic Systems in California." *National Renewable Energy Laboratory*, Colorado.

Wiser, Ryan, Galen Barbose, Carla Peterman and Naïm Darghouth. 2009. "Tracking the Sun II: The Installed Cost of Photovoltaics in the U.S. from 1998-2008." *Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory*.