Peer Effects and Learning by Doing in the Diffusion of Solar *

Bryan Bollinger
Stanford University

Kenneth Gillingham
Stanford University

PRELIMINARY AND INCOMPLETE - COMMENTS WELCOME

December 2009

Abstract

The solar photovoltaic (PV) industry in the United States has been the recipient of substantial production subsidies at the federal and state level, often motivated both by environmental externalities and dynamic spillovers from learning-by-doing in the diffusion of the new technology. Understanding the nature of this diffusion and documenting the empirical extent of these spillovers are crucial for assessing the benefits of solar production subsidies. Using a rich dataset on solar PV installations in California, we find evidence of peer effects that depend on the installed base at the zip-code level, suggesting the presence of social learning or snob effects for consumers. We then estimate a structural model of supply, using the zip-code level installed base as a demand shifter to instrument for the quantity being installed at any given time. We find evidence of highly non-appropriable contractor LBD at the regional level. These results suggest that there is an economic basis for a production subsidy on solar in California, even if the current subsidies may still be overly generous.

*The authors would like to thank the participants in the Stanford Economics Department Environmental Reading Group, the Stanford MS&E Policy and Economics Research Roundtable, the UC-Berkeley ARE seminar series, and faculty and students from the Stanford Graduate School of Business for their valuable comments and insights. All remaining errors are solely the responsibility of the authors.

1 Introduction

Policies to promote solar photovoltaic (PV) adoption have been gaining momentum throughout the world, as concerns over global climate change and national security externalities continue to grow. In the United States alone, commercial and residential solar installations are eligible to receive a 30% solar energy investment tax credit, which was temporarily converted to a cash grant by the American Recovery and Reinvestment Act of 2009. This federal subsidy comes on top of individual state incentive programs, the most prominent of which is the California Solar Initiative (CSI), a $3.3 billion 10 year program providing substantial subsidies for solar installations through 2016. Moreover, California, along with several other states, has adopted a Renewable Portfolio Standard, which in California's case requires electric utilities to generate 33% of their electricity by 2020, a major boon for centrally generated solar. With such considerable policy interest and high costs of implementation, it is critical to determine whether these policies are justified.

The answer to this question requires weighing the costs and benefits of such subsidies. Proponents of solar energy often point first to the reduced environmental externalities from displacing fossil fuel electricity generation, particularly during the peak hours of the day when wholesale electricity prices are highest. However, there is ample evidence that the environmental externalities are not likely to outweigh the high cost of solar technology (van Benthem et al. 2008; Borenstein 2008).[1] Further justification for incentives must hinge upon the existence of additional market failures relating to the dynamics of the diffusion of solar.

To establish and quantify the nature of these market failures it is necessary to first understand the potential sources of the market failures. Two natural questions arise. First, what are the determinants of the diffusion of solar? One such determinant may be peer effects, whereby the number of current adopters of the technology may influence the choice of whether or not to adopt. Spillover effects in technology adoption have been recognized in classic papers starting with Griliches (1957), and continuing in the innovation diffusion models of the 1960s (Arndt 1967; Bass 1969; Frank et al. 1964) and the more recent network externality literature (David 1985; Liebowitz 1994; Goolsbee and Klenow 2002). In the realm of environmentally friendly technologies, evidence for peer effects has been found in the purchase of hybrid electric vehicles (Axsena et al. 2009; Kakihara 2009), in the adoption of technologies to phase out lead (Newell and Kerr 2003), and in the adoption of ethanol vehicles and fuel (Shriver 2009).

In solar adoption, peer effects may be due to a variety of pathways, including social learning

[1] Note that production subsidies for renewables are a second-best instrument to address environmental externalities, since they effectively lower the price of electricity, reducing the incentives for energy efficiency

about the value of solar and snob effects in which consumers adopt solar to send a message to peers about their "green" choices. The presence of peer effects in solar improves our understanding of the process of adoption of a critical renewable energy technology. It is important to note however, that even if peer effects are present, this does not necessarily mean that there is a market failure: if the peer effects are accounted for by the adopters or by the firms supplying the solar panels, then there may in fact be no externality associated with the peer effects.

Second, what is the nature of the manufacturer and contractor cost structure? One of the most prominent arguments given for promoting solar is the possibility of supply-side learning-by-doing (LBD) in the solar industry, whereby the cost of an installation declines with cumulative experience in installations. The concept of LBD in economics dates to the early 1960s, with theoretical work by Arrow (1962) and empirical work by Alchian (1963). Since then, a sizable literature of both theoretical and empirical work on LBD in economics has developed.[2] Similarly, LBD has long been used to examine new energy technologies, beginning with Zimmerman (1982), and more recently as a common descriptive methodology for modeling technological change in renewables.[3]

Whether learning-by-doing can be considered a market failure depends on whether the learning from the experience by one firm engenders a spillover to other firms in the same industry. For example, if solar manufacturers copy the latest manufacturing techniques that other firms learned from trial-and-error, there would be a learning-by-doing spillover, sometimes called "external learning" effects, or "non-appropriable" LBD. In contrast, if solar manufacturers are able to appropriate all of the gains from their learning (sometimes called "internal learning"), there is no market failure, and hence no motivation for policy based on LBD.

This notion turns out to be critical when examining the costs and benefits of solar policies. Indeed, van Benthem et al. (2008) perform an ex-ante policy analysis and find that the substantial subsidy program in California can be justified on economic grounds based on environmental externalities and LBD - assuming that LBD in California continues to follow the learning previously seen for solar *and* the learning is non-appropriable.

This paper aims to shed light on the empirical evidence for peer effects and LBD in the California solar PV market. Using a doubly stochastic hazard model of adoption, we find statistically significant localized peer effects in the adoption of solar PV. These unique demand shifters are then used as instruments in our structural model of supply for the California solar market. Our

[2]Classic papers using the concept of learning-by-doing include Spence (1981), Fudenberg and Tirole (1983), Young (1991), Lieberman (1984), Irwin and Klenow (1994), and Foster and Rosenzweig (1995).

[3]See Grubb et al. (2002) and Gillingham et al. (2008) for overviews of modeling endogenous technological change in climate policy models.

results suggest important contractor-level learning effects at the regional level. Moreover, we find statistically significant non-appropriable LBD, providing empirical evidence for learning spillover effects.

Our paper is structured as follows. Section 2 provides background on the California solar market. Section 3 describes the panel dataset used for this study. Section 4 presents our hazard model of adoption and the results of the demand estimation. Section 5 develops the structural model of supply, which is estimated and provides evidence of small but statistically significant, non-appropriable LBD at the contractor level. Finally, section 6 concludes.

2 Background

2.1 Solar Policy

There has been a long history of government support for solar energy in the United States. At the federal level, incentives for solar date back to the Energy Tax Act (ETA) of 1978 (Public Law 9-618), which provided a 40% tax credit of up to $2,000 for homeowner installations of solar. This tax credit was phased out in the mid-1980s, but low-level federal support through research and development and pilot projects continued. More recently, the Energy Policy Act of 2005 created a 30% tax credit for residential and commercial solar PV installations, again with a $2,000 limit. The Energy Improvement and Extension Act of 2008 further incentivized solar by removing the $2,000 limit and allowing the credit to be taken against the alternative minimum tax.

On top of this federal policy action, both state and municipal governments in California have been long active in promoting solar - dating back to the creation of the California Energy Commission (CEC) in 1974. For example, in 1984, the Sacramento Municipal Utility District developed a 1 megawatt (MW) solar PV facility, which was later expanded to 2 MW. However, for several decades much of the emphasis was on larger systems and the high cost of residential systems deterred significant adoption. Interest in distributed generation solar PV picked up in the late 1990s. In 1997, Senate Bill 90 set in place the Emerging Renewables Program, and directed investor owned utilities to add a surcharge to electricity bills. The proceeds of this surcharge were then allocated toward a $3 per Watt (W) rebate for solar installations (Taylor 2008).

In 1998, "net metering" was permitted, allowing owners of solar PV systems to receive credit for electricity sold back to the grid during times when the use of electricity was below solar PV output. In 2000, Senate Bill 1345 directed the CEC to develop a additional grant program for distributed generation solar, but the funding for this was short-lived. In addition, Senate Bill 17 of 2001 provided state tax credits for solar PV installations of up to 15%, and these credits remained in place through the end of 2005 (CPUC 2009).

While these incentive programs were substantial, they were renewed on a year-by-year basis, leading to much uncertainty in the solar market. The elements for a longer-term, more predictable policy were put in place in California in August 2004, when Governor Schwarzenegger announced the "Million Solar Roofs Initiative," setting a goal of one million residential solar installations by 2015. In January 2006, the California Public Utilities Commission (CPUC) established the CSI, the $3.3 billion, 10-year program aiming to "install 3,000 MW of new solar over the next decade and to transform the market for solar energy by reducing the cost of solar" (CPUC 2009). The CSI was a unique subsidy policy in that it *counted on* the policy reducing the cost of solar by reducing the subsidy in steps over time as the number of installed MW increases (Figure (1)).

2.2 Solar Costs and Installations

The combination of the CSI and the federal incentives has corresponded with a considerable increase in the number of installations in California (Figure 2). However, the cost per W of solar PV installations in California has not simply declined along with the increase in installations, but actually has been fairly level since 2001, as shown in Figure 3. For reference, cursory calculations suggest that the 2009 residential system average cost of $8 per DC W corresponds to a levelized cost of roughly $0.30-$0.35 per kWh before any incentives.[4] Centrally generated electricity sources, such as coal or natural gas currently have a much lower levelized cost, usually in the range of $0.05-$0.07.

The final installation price paid to a contractor can be broken down into several components. The largest component of the cost is the cost of the PV module itself, which often makes up roughly 50% of the total cost of the installation (Wiser et al. 2009). The PV module market is widely considered a global market, with modules being built in Asia, Europe, and North America for use anywhere in the world, including California (IEA 2009). The best data available on module prices for multi-crystalline silicon modules (the most commonly used module technology in Califor-

[4] These calculations assume a 30 year solar system lifespan, a 30 year mortgage with an interest rate of 3%, an inverter lifespan of 8 years, solar PV system output from Borenstein (2008), no soiling losses, and a PV panel decay for multi-crystalline silicon panels of 0.5% corresponding to the best available evidence (Osterwald et al. 2006). The true levelized cost would vary for individual installations, and alternative assumptions could change the levelized cost calculation significantly.

nia installations) are Navigant Consulting's Global Module Price Index and the SolarBuzz North America Module Retail Price Index. Both of these tend to follow each other relatively closely, and the SolarBuzz North American Retail Price Index follows their European Price Index closely as well - providing further evidence of a global market (SolarBuzz 2009). Figure 3 shows the Navigant Global Module Price Index, both for "quantity" purchases (the price that most contractors would pay), and purchases from very small scale buyers (the price that some very small contractors may pay) (Consulting 2009). These data suggest that for a large contractor in California, modules usually make up roughly 40% of the total cost over the past eight years.

An inverter is needed with every PV installation to convert electricity from direct current (DC) to alternating current (AC). Inverters usually cost in the range of $0.50-$1 per DC W, implying that they may make up roughly 6%-15% of the total cost. Like modules, the inverter market is also considered to be a global market. The remainder of the installation cost is often called the "balance-of-system" (BOS) cost, and is composed of labor costs, marketing costs, overhead, and profits. Among these, labor costs are usually considered to be the largest. In California, there are numerous contractors, and most install solar in only a few counties, as will be discussed below in Section 3.

These facts about the solar market suggest that if there is LBD in California solar PV, as the writers of the CSI hope, then the learning in module costs will likely be at the global level, while learning in BOS costs will likely be at the regional level, corresponding to the geographic scope of most contractors. Indeed, previous studies examining evidence for LBD in the solar industry have focused on the global market for producing modules. The evidence so far for learning in the module cost is very weak at best. Papineau (2004) finds the effect of cumulative experience on total solar PV cost reductions to be significant in some specifications, but in her preferred specification that includes a time trend, cumulative experience is insignificant. Nemet (2006) performs an engineering analysis of the costs of PV modules and finds that learning can only weakly explain cost reductions in the most important components of the cost of producing a PV module.

On the other hand, some authors have suggested that LBD may be important at the regional level in lowering contractor BOS costs (Duke et al. 2005). Along with each additional installation, contractors gain experience with managing and marketing technologically similar installations and their workers gain experience in constructing the installation. If the learning that comes about through these processes spills over to other contractors, the spillovers are most likely to occur at the regional level, to other competing firms in the same geographic region who may be able to replicate the same cost-reductions. Moreover in the solar PV market in California, it is common for contractors to use temporary labor from a limited set of companies for large jobs or during times when there are many jobs - possibly leading to further spillovers at the regional level. Unfortunately, there is a dearth of empirical evidence on LBD in contractor BOS costs, a situation this paper endeavors to alleviate.

3 Data

Our data come from multiple sources. The first source of data is the California Solar Initiative website; the CSI tracks all solar installations since the programs conception in the three, investor-owner utility regions (IOUs) in the state: Pacific Gas and Electric Company (PG&E), Southern California Edison (SCE) and San Diego Gas and Electric (SDG&E). The CSI incentives are only available for installations within these regions, although the regions cover nearly the entire state.[5] The data are very rich and include the IOU, type of installation (residential, commercial, government or nonprofit), size of the installation and incentive, PV installer and manufacturer information, and zip code (and NAICS industry code if applicable) of the customer. In addition, the data include the date when the customer reserved solar incentives for an installation and the date when the project was completed, if it was completed. These two separate dates are extremely useful in an empirical analysis since the reservation date is the one in which a consumer makes the decision to install solar panels and locks in the incentive amount, and the completion date can be used to calculate the cumulative install base at any point in time.

We augment the CSI data, which begin in 2007, with data back to 2001 which were tracked by the CEC's Emerging Renewables Program before the tracking responsibilities were handed off to the CPUC, following the implementation of the CSI. The data also consist of installations, 99.5% of which are in the IOU regions. We restrict the analysis to these regions since they contain almost all of the installations and the data are not available after December 2006 for the non-IOU regions. The CEC data include the same variables as the CSI data with the exception of the installation type. To assign an installation type, we classify installations that are smaller than 10 kW as residential and larger than 10 kW as commercial.[6] The combined data set includes 50,207 installations between January 2001 and October 2009, 30,564 of which are from the CSI period. We augment the data with demographic data from Sourcebook America and American FactFinder.[7]

[5] The CSI incentives are either a one-time payment (for smaller systems) or broken down into smaller payments over the course of five years (for larger systems).

[6] The level of 10 kW has been used in industry reports as a suitable cutoff level and it is the 95th percentile of the residential installations in the CSI data. Since 10 kW is just below the 50th percentile for our CSI commercial installations, some commercial installations will inadvertently be coded as residential in the pre-CSI period.

[7] We also use as data on hybrid vehicle registrations from a proprietary data set as a proxy for green preferences.

We assume that peer effects operate at the zip code level. Without actual address data, we cannot look at a smaller geographic scale. Table 1 contains summary statistics for the zip codes. The average number of installations in a zip code is 31.4 (22.8 in the CSI period). Most of the installations are residential (27.5 per zip) although since residential installations are smaller on average than commercial, government and nonprofit installations, as shown in Table 2, the residential installations make up only 33% of the installed MW. The price variable is adjusted by the CPI to real 2009 dollars per W and all Watts in this paper are direct current Watts.

One implicit assumption in our analysis is that local supply side spillovers are not contained at the zip-code level, which is reasonable since contractors operate in regions that are significantly larger. Table 3 breaks down the installations by contractor and manufacturer. The average contractor operates in 2.8 counties and has performed 26.7 installations. One difficulty we face in determining the extent of localized, nonappropriable contractor LBD is that it is not possible to distinguish between LBD effects and demand-side peer effects if the peer effects operate at a larger geographic unit than the zip code and regionalized nonappropriable LBD operates only at some geographic unit (such as the county level) and is the same for all contractors within that unit.

In reality, LBD spillovers are most likely to occur between competing contractors, and each contractor faces a different set of competitors, due to the different geographic coverage of each contractor. Thus, we create a variable for the installed base of each contractor's competitors, where we define a competitor to be any other contractor who has installations in a county in which the contractor operates (over the years covered in the dataset). Using this variable allows us to avoid the potential issues inherent in using an arbitrary geographic cutoff point for where demand-side peer effects end and supply-side LBD begins.

Table 4 includes summary statistics for the different installed base variables that will be used in the analysis. In addition, we create current contracts variables for a contractor, its competitors, a manufacturer and within a zip code. These quantity variables may capture scale economies, depending on the nature of competition, as will be discussed further in section 5. One possible criticism of estimations of LBD in other papers is that LBD is confounded with economies of scale. Here we can control for both since we know the date at which the installation is requested and the data at which it is finally completed.

4 Solar Adoption

The most common empirical method employed in the technology adoption literature is a reduced-form hazard function approach (Hannan and McDowell 1984; Mulligan 2003; Baker 2001; Engers et al. 2009), especially in the presence of network effects.[8] Diffusion models such as the Bass model (Bass 1969) are derived from hazard models and assume that the probability of adoption depends on the number of previous adoptions. These models include a positive effect of the installed base, captured through a "coefficient of imitation," as well as a market saturation component. The combination results in the well-known "S-shaped" cumulative adoption curve.[9]

As an alternative to the hazard function approach, it is becoming more common to employ a structural model of technology adoption.[10] Structurally modeling the decision to adopt is useful for performing a counterfactual analysis, but is fraught with difficulties such as multiple equilibria (especially in the case of spillovers) and strategic behavior. Some models begin with a latent variable model (Tucker 2004; Gowrisankaran and Stavins 2004; Tucker 2008) and then assume away strategic behavior. In other cases, it is necessary to either find convincing instruments or to simultaneously estimate demand and supply. The popularity of the hazard function approach is largely due to the fact that often the main goal is simply to establish the presence of and perhaps quantify network effects, rather than to perform a counterfactual analysis, which would require a structural approach. Similarly, since our goal is to establish the presence and quantify network effects, and since we do not have the entire menu of prices for all manufacturer-contractor combinations that the consumers can choose from, which would be required in a structural model, we adopt a hazard function approach as well.

One main challenge in the estimation of peer effects is separating the effect of the installed base from correlated unobservables. Most of the empirical applications that estimate duration models assume a specific distributional assumption for the hazard model and may sometimes include fixed effects in the estimation of the hazard rate to control for some of the unobservables. However, even the inclusion of market fixed effects completely ignores the effect of idiosyncratic preferences for individuals within the market of study.

[8] Another common approach uses a normal, censored regression for the date of adoption (Lee and Waldman 1985; Genesove 1999).

[9] Other reduced form models that are used in determining the presence of spillovers include linear models if the outcome variable is continuous, such as the effect of movie reviews on box office revenues (Moretti 2008), or by assuming a linear probability model (Goolsbee and Klenow 2002; Duflo and Saez 2003; Gowrisankaran and Stavins 2004; Oster and Thorton 2009). Other papers exploit a natural experiment such as Kremer and Levy (2008) which studies the peer effects on alcohol consumption.

[10] Recent structural work that incorporates network effects include Hamilton and McManus (2005), Lenzo (2006), Schmidt-Dengler (2006) and Shriver (2009).

The implicit assumption in most hazard models of adoption is that individuals are homogenous, i.e., that there is no heterogeneity component in the hazard function parameters. This assumption allows for easy estimation of the hazard function parameters using a simple maximum likelihood approach. Robustness checks are usually limited to trying other parametric specifications of the hazard function. However, failure to account for individual heterogeneity can lead to severely biased estimates of the duration model parameters, as explained in Heckman and Singer (1984). In addition, Heckman and Singer show that if there is unobserved heterogeneity in the hazard function, the hazard function parameter estimates are sensitive to the distributional assumptions on that heterogeneity. Furthermore, they show that the distributions of the hazard function and heterogeneity are not separately identified.

In our application, the presence of individual heterogeneity within zip codes cannot be disputed. We feel that it is therefore more benign to assume a distributional form for the hazard function than for the heterogeneity component. We assume that solar adoption follows a non-homogenous Poisson process (also known as a Cox process or doubly stochastic Poisson process). Poisson processes are pervasive natural phenomena and have the desirable feature of duration independence, that is, the hazard function is independent of the time since the last adoption, conditional on the hazard rate. However, the Poisson arrival rate is a stochastic function of the independent variables, which may be dependent on time.

4.1 Hazard Model of Adoption

Assume that at any point in time, t, a (residential) consumer has an expected net present value for the installation of solar panels which depend on the costs of the panels, the incentives in place, electricity prices, and the future expectations over these variables. In addition, allow this net present value to depend on the current "local" install base, where we define local to mean other consumers in the same zip code. We assume a Poisson arrival process with the following hazard rate:

$$\lambda(t) = \lambda_0(t) \exp(X_{zt}\beta + \eta_z + \xi_t^d + \epsilon_{zt}), \tag{1}$$

where X_{zt} are market and time-specific explanatory variables, η_z and ξ_t^d are zip code market and time-specific indicator variables, respectively, and ϵ_{zt} is a stochastic term capturing the unobserved heterogeneity. The waiting time between events in a Poisson distribution is exponentially distributed:

$$f(\Delta t) = \frac{1}{\lambda(t)} \exp\{-\Delta t/\lambda(t)\}, \tag{2}$$

and so the log likelihood function takes the following form:

$$L = -\sum \Delta t/\lambda(t) + \log(\lambda(t)), \tag{3}$$

where the hazard rate in the stochastic value given in (1). The first order condition that maximizes this likelihood function can be arranged to give us the following expression for the natural logarithm of time between adoptions:

$$\log(\Delta t_{zt}) = X_{zt}\beta + \eta_z + \xi_t^d + \epsilon_{zt}. \tag{4}$$

By estimating this model of time between successive adoptions, we can allow the hazard rate to depend on time-varying market level regressors X_{zt}, such as the number of previous adoptions in the market and the current number of ongoing installations as well as an observed heterogeneity component. The monthly, time fixed effects capture changes in price level on a month-to-month basis as well as other time-varying factors that are the same across markets. To control for any changes in zip code-specific price levels (pre-incentive), we include the most recent price in the zip code.[11] Figure 4 includes histograms for time between zip code installations and the logarithm of this time, which is the dependent variable in the demand-side regressions. The average time between adoptions is 71 days, and the distribution is skewed with considerable variance. However, the distribution of log time appears to be symmetric; it is centered at 3.22 with a variance of 1.47.

By estimating (4), we assume a distributional form of the hazard function. This, however, does not imply that the rate of adoption is independent of time-dependent regressors, merely that the rate of adoption between adoptions does not change as a function of the time since the last adoption. Furthermore, a distributional assumption is required: either on the hazard function or on the distribution of heterogeneity, which is often assumed away with no discussion. When

[11] As a robustness check, we also estimate the model using a zip-code price index of the average price of the last ten installations and find similar results.

heterogeneity is obviously present, as it is in our application, we feel that an assumption on the hazard function distribution is less restrictive.

In addition, we test for autocorrelation in ϵ_{zt} by estimating equation (4) and regressing the residuals on lags of the residuals. We find that there is autocorrelation in the ϵ_{zt}, but only for one lag. The presence of autocorrelation will not bias the estimates but it will lead us to underestimate the standard errors if it is not accounted for. We therefore report asymptotic standard errors for our estimates assuming an AR(1) process for the errors, $\epsilon_{zt} = \rho\epsilon_{z(t-1)} + e^d_{zt}$.

4.2 Hazard Model Estimation Results

One difficulty in testing for peer effects is controlling for correlated unobservables that shift demand. In the estimation of equation (4), we control for zip code specific demand by using a fixed effects estimator, and we include month-level dummy variables to control for time-varying demand shocks. These dummy variables absorb shocks to overall demand from changes in the economy, advertising campaigns, etc. We report standard errors which allow for AR(1) autocorrelation. This captures time-varying effects that affect only specific zip codes which would accelerate or decelerate adoption over some period of time.

Table 5 contains the hazard model regression results, where the dependent variable is the log of the time until next adoption within a zip code. The first column reports a specification in which zip installed base and current contracts are measured by the absolute number of installations, in hundreds. We include a linear and quadratic term for the zip code installed base in order to capture either increasing or decreasing effects of more installations. The coefficient on the zip code installed base is negative and significant as expected, but there is also a significant, positive coefficient on the quadratic term, meaning that the effect of additional installations declines with the installed base. The interpretation of the size of the effect is that for the first 100 new installations in the zip code (the average amount is 101), the time between adoptions decreases by a little less than one percentage point (decreasing from 0.951% to 0.693%) in the PG&E utility region, and it is twice this size in the SCE utility region.

At first glance, this may seem like a trivial effect on the rate of adoption, but over the period covered by our data, a decrease of one percent in the time between adoptions is considerable. Although the average time between installations across all zip codes is 229 days, this is because there are some zip codes with very little adoption, distorting the mean. The median is much lower, at 90 days, and the 25th percentile is only 28 days between installations. A one percentage decrease in the time between adoption in these zip codes leads to a large increase in the number of total adoptions. Aggregating the effect over these heterogeneous zip codes leads to 524 more adoptions for the years of our data (2001-2009) due to the peer effects.

If solar adoption were to become more prevalent in the zip codes with little adoption up to this point in time, the influence of peer effects would increase. This is particularly relevant because the rate of solar installations have just begun to accelerate in California. While the average number of installations is 101, there is an average of 7,500 owner occupied homes in each zip code, so the market is far from being saturated.

Note that we have not said anything regarding what causes the peer effects, and indeed a study such as ours cannot determine the mechanism behind the peer effects. Ongoing research at UCLA shows some support for the idea that people adopt solar for social image reasons (Lessem and Vaughn 2009), which is sometimes referred to as a ”snob” effect. The authors make use of a measure of ideology and show that while consumers' election of a green energy alternative through their utility provider is driven by their own ideology, the decision to adopt solar is driven more by their neighborhood's average ideology. Of course this could be due to either a snob effect of due to the availability of more information in those areas.

Our research provides some limited support for an informational component to the peer effect. With the conception of the CSI, the utility districts were responsible for their own marketing campaigns to promote solar adoption. The PG&E utility district was by far the most aggressive with such campaigns. We also see that the peer effects are the smallest in this region. While there are many possible reasons for the difference in the size of the peer effect across regions, one possible explanation is that in the PG&E utility region, since more information regarding solar was provided through marketing vehicles than in the SCE and CCSE utility regions, this resulted in a weaker peer effect because part of the peer effect is due to a transfer of information. An informational component of the peer effect is of course only one possible reason for the smaller peer effects in the PG&E region; in all probability, the peer effects result from both informational and snob effects.

We find no effect of the most recent installation price in the zip code on the rate of adoption. Any effect of overall price levels are of course captured by the time dummies, so we would only expect the most recent price to have any effect if there was time varying price disparity across zip codes. We do, as expected, find that as the CSI incentives decline (with step number), that the time between adoptions grows. This increase is not statistically significant which is not surprising since these parameters are only identified through cross sectional variation in the current incentive step: time series variation is captured in the time fixed effects. Finally, we report the extent of the heterogeneity and the amount of autocorrelation. The variance of the zip-code fixed effects is given by σ_u and the variance of ε_{zt} by σ_ε. There is an autocorrelation, ρ, of 0.525.

As a robustness check, we run the model using the fraction of owner-occupied homes who have installed as the units for zip code installed base. The results are reported in the second column of Table 5. The results are all qualitatively the same, in both sign an significance. The first model has more observations since we don't have to worry about any missing demographic data. For this reason, and since we cannot be sure that the number of owner-occupied homes is the best measure of market size, and also because the first specification using absolute installation counts is a better fit to the data, we use absolute numbers in our additional model specifications.

In order to assess the effect of demographic demand variables on adoption, we run the full regression using a zip code random effect and zip code level demographic variables. The results are all as expected. Of course, the larger the owner-occupied installed base, the more solar adoption. The percentage of the population who are male, white, have a college degree, and own a hybrid vehicle also all lead to increased adoption whereas the percentage of people who drive to work, work at home, or have over a half hour commute all lead to less adoption.[12]

To better understand the determinants of the peer effects, we run the regression including the zip-code fixed effects but also with the demographic variables interacted with the zip code level installed base. We find that the number of homes increases the size of the peer effect as does the percentage of people aged 20-45, who use public transportation, and who work at home, whereas the percentage of people who are male, who drive hybrids, or who own homes worth over 400 thousand dollars all lead to smaller peer effects.[13]

So far, the analysis has focused on residential PV installations. We performed the same analysis using government and commercial solar installations, first defining the market as a zip code as done in the residential analysis and second, separating firms by industry type. We find no evidence of peer effects for commercial installations.[14] The lack of peer effects for commercial installations provides no evidence that companies are using solar installations as a green differentiation mechanism, a possibility touted in the California Solar Initiative Business Customer Fact Sheet. It does perhaps lend even more support to our argument that it is peer effects that we are capturing in the residential installation analysis.

[12] At 5 % significance.

[13] At 10% significance.

[14] There are too few government and nonprofit installations to run the analysis.

Although we demonstrate that peer effects exist, this does not necessarily imply that this is a network externality that leads to a market failure. The presence of peer effects could be accounted for through the behavior of the agents. It is possible that people compensate their neighbors in exchange for information regarding PV installations. More likely is the possibility that manufacturers and contractors price their panels to penetrate the market and make use of these peer effects increasing demand, although since the peer effects are not firm specific, the prices would still most likely be too high initially to overcome the market failure. Also, if the peer effects are informational in nature, the first-best solution to addressing the market failure would be to provide consumers with the information and not rely on the peer effects to do so, as the marketing campaigns in the different utility regions have attempted to do.

The next section uses the peer effects found here to help identify the more straightforward LBD market failure.

5 Contractor and Manufacturer Supply

5.1 Model

To determine whether or not the solar industry in California exhibits LBD, we develop a model based on profit-maximizing contractors and module manufacturers. Solar installations are not a homogenous good, since each installation is unique. Thus, the profit for each contractor j at time t is given by the sum over all of that contractor's installations:

$$\pi_{jt}^{c} = \sum_{i=1}^{Q_{jt}^{c}} (p_{ijt} - p_{ijt}^{m} - w_{ijt}^{c}) - F_{jt}^{c}. \tag{5}$$

Here Q_{jt}^{c} is the total number of installations by contractor j, p_{ijt} is the pre-incentive price charged to the consumer for installation i, p_{ijt}^{m} is the cost of the PV module, w_{ijt}^{c} are the BOS costs that include all other costs besides the PV module, and F_{jt}^{c} is the contractor fixed cost.

Similarly, the profits in California of manufacturer k for the module used in installation i at time t are given by

$$\pi_{kt}^{m} = \sum_{i=1}^{Q_{kt}^{m}} (p_{ikt}^{m} - w_{ikt}^{m}) - F_{kt}^{m}, \tag{6}$$

where Q^m_{kt} is the total number of modules produced by manufacturer k at time t, w^m_{ikt} is the manufacturer's marginal cost of production, and F^m_{jt} is the manufacturer fixed cost.

If LBD is an important factor in the production and installation of solar PV, we would expect the cost of production and installation to decline as manufacturers and contractors have more experience. It thus follows that if LBD is influential, costs to both manufacturers and contractors should be a function of the cumulative production of manufacturers and installations of contractors, respectively. Moreover, if learning is not fully appropriable, there will be spillovers leading to costs declining along with the cumulative installations of all relevant competitors. For manufacturers, the relevant market is all of California, so all other manufacturers are relevant competitors. Since contractors tend to be based on a single region, relevant competitors are all other contractors who sell in any one of the counties that the contractor sells in.

Thus, we specify our marginal cost functions as follows:

$$\begin{aligned} w^c_{ijlt} &= \alpha^c + \beta_c b^c_{jt} + \beta_{cc} b^{cc}_{jt} + \beta_{Q^c} Q^c_{jt} + \beta_{Q^{cc}} Q^{cc}_{jt} + \beta_{sc} S_{it} + \theta^c_j + \xi^c_t + \zeta_l + \varepsilon^c_{ijt}, \\ w^m_{ikt} &= \alpha^m + \gamma_m b^m_{kt} + \gamma_{mc} b^{mc}_{kt} + \gamma_Q Q^m_{kt} + \beta_{sm} S_{it} + \theta^m_k + \xi^m_t + \varepsilon^m_{ikt}. \end{aligned} \tag{7}$$

where b^c_{jt} is the contractor installed base (i.e., the contractor's cumulative installations), b^{cc}_{jt} represents the installed base for the contractor and all of the contractor's direct competitors, Q^{cc}_{jt} is the contractor's competitors' current on-going contracts (included to account for the possibility of short-term capacity constraints in availability of temporary labor), S_{it} represents the size of the installation, b^m_{kt} is the manufacturer's installed base, b^{mc}_{kt} is the total installed base, and the ε's are i.i.d. mean zero error terms (we allow for the possibility of heteroskedasticity). The ξ's are time dummy variables to account for unobserved changes in firm costs over time, such as the total global quantity solar panels manufactured. The θ's are contractor or manufacturer dummy variables that are included to account for initial firm specific installation bases (before our dataset begins in 2001) and heterogeneity in firm marginal costs. Finally, ζ_l are dummy variables for the installation type, since contractors may have different costs depending on whether the installation is a residential, commercial, or government installation.

The nature of competition in the California solar market determines the first-order conditions for the firms' profit-maximization problems. Discussions with industry analysts and players in the market suggest that the contractor market is highly competitive; the barriers to entry are relatively low for already licensed general contractors. For example, between January 2007 and October 2009, there were 1,333 different contractors installing solar in California. The module manufacturer market has had documented capacity constraints hindering increased production in the recent past, so it may not have been as competitive, although industry analysts believe that with the recent price decline the market is becoming more competitive.

Since the evidence is only anecdotal, we allow for the possibility of market power in solving the firms profit-maximization problem. We assume each firm charges a constant markup for all installations of each installation type (i.e., $p_{ijt} - p^m_{ijt} - w^c_{ijt}$ and $p^m_{ikt} - w^m_{ikt}$ are constant). This is more flexible than assuming perfect competition in which case the markup is assumed to be zero. The assumption is needed in order to aggregate each firm's profits over its installations and compute its first order condition under Cournot competition. The profit-maximization conditions for contractors and manufacturers are, respectively:

$$\begin{aligned} p_{ijlt} &= p^m_{ijt} + w^c_{ijlt} - \frac{\partial p_{ijt}}{\partial Q^c_{jt}} Q^c_{jt}, \\ p^m_{ikt} &= w^m_{ikt} - \frac{\partial p^m_{ikt}}{\partial Q^m_{kt}} Q^m_{jt}. \end{aligned} \tag{8}$$

Our specification is flexible in that the $\frac{\partial p_{ijt}}{\partial Q^c_{jt}}$ and $\frac{\partial p^m_{ikt}}{\partial Q^m_{kt}}$ terms, which are zero with competitive markets, are absorbed in the current contract variables which are included under either type of competition to allow for changing marginal costs.[15] If there is marker power, we cannot separately identify market power from non-constant marginal costs. Note that we do not assume constant market power, market power is a function of the changing quantity variable. One implicit assumption is that if market power changes with time as well as with quantity, then it changes for all contractors by the same amount (in which case it is then absorbed in the time fixed effects).

We can combine (9) and (8) to yield our final specification, where the price for any installation i of type l by contractor j and module manufacturer k is given by:

$$p_{ijklt} = \alpha + \gamma_m b^m_{kt} + \gamma_{mc} b^{mc}_{kt} + \gamma'_Q Q^m_{kt} + \beta_c b^c_{jt} + \beta_{cc} b^{cc}_{jt} + \beta'_Q Q^c_{jt} + \beta_s S_{it} + \theta^c_j + \theta^m_k + \xi_t + \zeta_l + \varepsilon_{ijkt} \tag{9}$$

where $\alpha = \alpha^c + \alpha^m$, $\beta_s = \beta_{sc} + \beta_{sm}$, $\gamma'_Q = \gamma_Q + \frac{\partial p^m_{ikt}}{\partial Q^m_{kt}}$, $\beta'_Q = \beta_Q + \frac{\partial p_{ijt}}{\partial Q^c_{jt}}$, $\xi_t = \xi^m_t + \xi^c_t$, and $\varepsilon_{ijkt} = \varepsilon^c_{ijt} + \varepsilon^m_{ikt}$.

[15] These could also be interpreted in such a way that our specification follows the conduct parameter approach which nests the models of perfect competition, monopoly pricing and Cournot competition (Corts 1999; Genesove and Mullin 1998).

If there is LBD in the California market, we expect to see significant coefficients on the installed base variables. Moreover, the relative magnitude of the own firm versus firm plus competitors installed base variables should provide a sense for the degree to which the LBD is internal fully appropriable learning or external non-appropriable learning. This formulation differs from some of the previous literature on LBD by estimating an model where the installed base enters additively, rather than multiplicatively.[16] We believe our formulation, based on our structural model, is more likely to capture how experience actually affects costs - the underlying premise of LBD. In addition, we are then able to include the ongoing contracts for contractors and manufacturers in a meaningful way to separate the effects of LBD and non-constant marginal costs.

5.2 Identification and Estimation

Since the purpose of this study is to identify and assess the magnitude of the LBD coefficients, we are most concerned about anything that could impinge upon the identification of these coefficients. However, it is worth noting what our model cannot identify. Foremost in this category are economies of scale or competitive effects. We can identify the sum of these two, but not each individually. Of course, if perfect competition is assumed, then the ongoing contract coefficient estimates are simply the effect of economies or diseconomies of scale.

In addition, an ordinary least squares (OLS) estimation of (9) suffers from endogeneity from the simultaneous determination of price and quantity in the solar market. Similarly, it appears reasonable to assume that the system size S is also simultaneously determined by supply and demand. If Q^m, Q^c, and Q^{cc}, and S are correlated with our variables of interest, we will not be consistently estimating the LBD effect. One approach to address this issue is to explicitly specify a demand equation and estimate the system of equations simultaneously. A second approach is to estimate (9) using two stage least squares, with suitable instruments for our endogenous variables.

We choose the latter approach. Jointly estimating a full system of supply and demand garners increased efficiency, but runs the risk of a misspecification on the demand side corrupting the estimated coefficients of interest on the supply side. Moreover, we have a logical set of instruments that are motivated by our hazard model estimation: the zip-code level localized installed base of solar. The results of the hazard model demonstrate that the localized peer effects influence consumer demand, so these fulfill the inclusion restriction. At the same time, it is difficult to see how peer effects, as quantified by zip-code level installed base, could influence the marginal cost

[16]For example, most estimations of the extent of LBD are based on a specification along the lines of $\log(p) = \alpha + \beta \log(b^c) + \epsilon$, where the learning rate is given by $LR = 1 - 2^{\beta}$.

of contractors and affect supply. While costs may vary across contractors, for a given contractor they are not likely to vary across zip codes.[17] Thus, the exclusion restriction here seems reasonable.

Demographic variables are additional possible instruments that act to shift demand and not affect supply. For demographic variables at the zip code level, the motivation is simple: demographics should not appreciably affect manufacturers and contractors, except through how they influence demand. We include the following demographic variables described in section 3 as instruments: % driving hybrids, population in zip, % with college degrees, % carpooling, % male population, median home value, % of population with a home loan, % of population who have done a major home repair, % of population in 2007 that was white, median household income, % population less than 19 years old, % of population 20-45 years old, % population over 65 years old, % of population who drive, % of population who carpool, % population who work at home, % of population who have a 30 minute commute or greater, and interactions between each of these and the residential zip code installed base.

Finally, the contractor-level fixed effects duplicate the installed base for 501 contractors who only installed one solar system, so we drop these observations in our supply estimation to alleviate this identification issue.

5.3 Supply Estimation Results

The results from estimating our supply model are given in Table 8. We report robust standard errors. The first column presents a simple OLS estimation of equation (9) for diagnostic purposes. The only significant variables are the contractor plus competitors installed base, the system size, the contractor on-going contracts, and the contractor's competitors' on-going contracts. The OLS results suggest a small amount of non-appropriable learning for contractors (e.g., the coefficient indicates that if either the contractor or any of its competitors performs 100 more installations then the cost of an installation will decline by $0.009 per W from an average price of $8 per W of which just under half is the BOS cost subject to contractor learning). The system size coefficient suggests very small economies of scale for a given installation, while the contractor on-going contracts provides evidence for limited diseconomies of scale, as will be discussed further below.

The OLS results are suspect, since the simultaneous determination of price and quantity implies that S, Q^c, Q^{cc}, and Q^m can be expected to be endogenous. The data corroborate this, as

[17] Examining the data, we find low within-contractor variance of price across zip codes. This is informative, but far from definitive, since price may not the same as marginal cost.

the robust Durbin-Wu-Hausman test statistic is 14.1, which corresponds to a p-value of 0.000, giving strong evidence rejecting the null of exogeneity. The second column presents the results of estimating equation (9) using two stage least squares with our demand shifters: residential zipcode installed base (peer effects), and the demographic variables. While it is impossible to know whether the entire set of these instruments truly identify the supply equation, we can perform a Hansen-Sargan over-identification test to give some sense of whether our instruments are reasonable. The test statistic is 31.5, which implies a p-value of 0.39 using a chi-squared distribution with 30 degrees of freedom. Thus, we fail to reject the null that the overidentifying restrictions are valid instruments, providing evidence that we have valid instruments.

The estimation results suggest that there is a statistically significant learning effect. Just as in the OLS result, the contractor installed base coefficient is small and insignificant. This suggests that there is little appropriable LBD for contractors. The coefficient on the sum of the contractor and competitors' installed base is much larger than the OLS result, and remains highly significant. The coefficient suggests that if the contractor performed 100 more installations, that experience would imply a lowering of the price of solar installations by $0.065 per W, out of the average price in the range of $8 per W for residential installations, of which just under half is the BOS cost that contractor learning can affect.[18] If we assume market power does not change with more cumulative experience in installations, then our structural model implies that the marginal cost would decline by the same amount - indicating a relatively small, but significant spillover from experience. As with the OLS results, we find that the manufacturer installed base in California and the total installed base are both insignificant. This suggests there is little evidence for California-level LBD for manufacturers. There may or may not be LBD for manufacturers based on their *global* installed base, but unfortunately, those data are not available.

The coefficients for on-going contracts for contractors and their competitors are positive and significant, but because we can not separately identify them, it is impossible to know whether this is a result of diseconomies of scale or market power. Anecdotal evidence based on conversations with solar companies suggests that capacity constraints may play a role in this market. Solar companies in California often rely on temporary laborers for key phases of many installations. These temporary laborers are usually drawn from a few companies that specialize in renewable energy installations. If solar contractors expand too quickly, they may temporarily exhaust the supply of trained laborers, leading to higher costs. The positive and statistically significant coefficient on the on-going contracts of the competitors corresponds well with this possibility. If the local competitors of any given contractor have more on-going contracts, that may lead to short-term

[18]While our specification does not allow for a direct computation of a learning rate, if we estimate the percent change in cost at the current values of costs and installations, we find a surprisingly small learning rate of just a few percent.

shortages in trained temporary workers. Since our dataset runs from 2001 to mid-2009, a time of rapid growth in the industry, it seems reasonable that we could pick up this effect.

5.4 Robustness Checks

[In progress]

6 Conclusions

With policy interest and activity in promoting solar greater than it has ever been, there is a pressing need for retrospective analysis to understand whether there is evidence for other market failures leading to an under-adoption of solar besides the environmental market failures. This is particularly crucial in a state such as California, where there is also a planned cap-and-trade system to internalize the environmental externality from fossil fuel-based electricity generation.

Our results suggest that there are statistically significant peer effects at the zip code level in the adoption of solar in California. These peer effects may be due to a variety of factors, and understanding these peer effects more deeply may be informative for developing marketing strategies for solar firms. However, the peer effects may or may not represent a market failure that warrants policy intervention, depending on whether the spillovers inherent in the peer effects are internalized by firms and/or adopters of the technology.

On the other hand, non-appropriable LBD represents a clear market failure if it can be demonstrated. The difficulty is demonstrating whether or not LBD exists and is appropriable when there is a paucity of disaggregated data on the cost structure of solar PV. We overcome this by estimating a structural model of solar supply using a rich data set of solar installations in California from 2001-2009.

Our result is that there is statistically significant, non-appropriable LBD (and insignificant, appropriable LBD). This finding has important ramifications for solar policy by providing some economic backing for the current, decreasing installation subsidies as well as potential production subsidies which could correct for the LBD market failure. However, the non-appropriable LBD is reasonably small, and may not come even close to justifying the substantial state and federal subsidies now in place in California.

In addition, to provide motivation for policy action, the LBD market failure should be *particularly* important for solar PV, since some degree of non-appropriable LBD may be a feature in a wide range of technologies and there is a cost associated with any subsidy policy. Policy action is warranted if the benefit from correcting the LBD market failure and the environmental benefits together are greater than the cost of administering the policy and the distortionary cost of raising the revenue. We do not perform this benefit-cost analysis, but a detailed analysis may not be completely necessary.

Based on the results of van Benthem et al. (2008), it is clear that if LBD is relatively small, then the substantial subsidy policy in California is likely not justified on economic grounds. Moreover, the new federal subsidy policy makes it even less likely that the current support for solar is economic efficiency-improving. However, we must acknowledge that this result could turn based on the assumptions in the benefit-cost analysis. Exploring these issues further will be a valuable future research endeavor.

References

Alchian, A. (1963). Reliability of progress curves in airframe production, *Econometrica* **31**: 679–693.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product, *Journal of Marketing Research* **4**(3): 291–295.

Arrow, K. (1962). The economic implications of learning by doing, *Review of Economic Studies* **29**: 155–173.

Axsena, J., Mountain, D. C. and Jaccard, M. (2009). Combining stated and revealed choice research to simulate the neighbor effect: The case of hybrid-electric vehicles, *Resource and Energy Economics* **31**(3): 221–238.

Baker, L. C. (2001). Managed care and technology adoption in health care: evidence from magnetic resonance imaging, *Journal of Health Economics* **20**(3): 395–421.

Bass, F. M. (1969). A new product growth model for consumer durables, *Management Science* **15**: 215–227.

Borenstein, S. (2008). The market value and cost of solar photovoltaic electricity production, *University of California Energy Institute, CSEM Working Paper 176* .

Consulting, N. (2009). Pv global price index.

Corts, K. S. (1999). Conduct parameters and the measurement of market power, *Journal of Econometrics* **88**(2): 227–250.

CPUC (2009). California solar initiative annual program assessment, *California Public Utilities Commission Go Solar California Report* .

David, P. (1985). Clio and the economics of qwerty, *American Economic Review* **75**: 332–337.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment, *Quarterly Journal of Economics* **118**(3): 815–842.

Duke, R., Williams, R. and Payne, A. (2005). Accelerating residential pv expansion: Demand analysis for competitive electricity markets, *Energy Policy* **33**(15): 1912–1929.

Engers, M., Hartmann, M. and Stern, S. (2009). Are lemons really hot potatoes?, *International Journal of Industrial Organization* **27**(2): 250–263.

Foster, A. and Rosenzweig, M. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture, *Journal of Political Economy* **103**(6): 1176–1209.

Frank, R., Massy, W. F. and Morrison, D. G. (1964). The determinants of innovative behavior with respect to a branded, frequently purchased food product, *in* L. Smith (ed.), *Proceedings of the American Marketing Association*, American Marketing Association, Chicago.

Fudenberg, D. and Tirole, J. (1983). Learning-by-doing and market performance, *Bell Journal of Economics* **14**: 522–530.

Genesove, D. (1999). The adoption of offset presses in the daily newspaper industry in the united states, *NBER Working Paper No. 7076* .

Genesove, D. and Mullin, W. P. (1998). Testing static oligopoly models: Conduct and cost in the sugar industry, 1890-1914, *The RAND Journal of Economics* **29**(2): 355–377.

Gillingham, K., Newell, R. and Pizer, W. (2008). Modeling endogenous technological change for climate policy analysis, *Energy Economics* **30**(6): 2734–2753.

Goolsbee, A. and Klenow, P. (2002). Evidence on learning and network externalities in the diffusion of home computers, *Journal of Law and Economics* **45**: 317–344.

Gowrisankaran, G. and Stavins, J. (2004). Network externalities and technology adoption: Lessons from electronic payments, *The RAND Journal of Economics* **35**(2): 260–276.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change, *Econometrica* **25**(4): 501–522.

Grubb, M., Khler, J. and Anderson, D. (2002). Induced technical change in energy and environmental modeling, *Annual Review of Energy and the Environment* **27**: 271308.

Hamilton, B. H. and McManus, B. (2005). Technology diffusion and market structure: Evidence from infertility treatment markets, *working paper* .

Hannan, T. H. and McDowell, J. M. (1984). Market concentration and the diffusion of new technology in the banking industry, *The Review of Economics and Statistics* **66**: 686–691.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data, *Econometrica* **52**(2): 271–320.

IEA (2009). Trends in photovoltaic applications: Survey report of selected iea countries between 1992 and 2008, *Report IEA-PVPS T1-18* .

Irwin, D. and Klenow, P. (1994). Learning-by-doing spillovers in the semiconductor industry, *Journal of Political Economy* **102**(6): 1200–1227.

Kakihara, N. (2009). Effect of social influence in new product adoption: The case of toyota prius in california, *working paper* .

Kremer, M. and Levy, D. (2008). Peer effects and alcohol use among college students, *Journal of Economic Perspectives* **22**: 189–206.

Lee, R. H. and Waldman, D. M. (1985). The diffusion of innovations in hospitals: Some econometric considerations, *Journal of Health Economics* **4**(4): 373–380.

Lenzo, J. (2006). Market structure and profit complementarity:the case of spect and pet, *working paper* .

Lessem, N. and Vaughn, R. (2009). Image motivation in green consumption, *Manuscript, UCLA Economics Department* .

Lieberman, M. (1984). The learning curve and pricing in the chemical processing industries, *Rand Journal of Economics* **15**: 213–228.

Liebowitz, S. Margolis, S. (1994). Net externality: An uncommon tragedy, *Journal of Economic Pespectives* **8**(2): 133–150.

Moretti, E. (2008). Social learning and peer effects in consumption: Evidence from movie sales, *working paper* .

Mulligan, J. G. (2003). Market segmentation and the diffusion of quality-enhancing innovations: The case of downhill skiing, *Review of Economics and Statistics* **85**(3): 493–501.

Nemet, G. (2006). Beyond the learning-curve: Factors influeincing cost reductions in photovoltaics, *Energy Policy* **24**(17): 3218–3232.

Newell, R. and Kerr, S. (2003). Policy-induced technology adoption: Evidence from the u.s. lead phase-down, *Journal of Industrial Economics* **51**(3): 317–343.

Oster, E. and Thorton, R. (2009). Determinants of technology adoption: Private values and peer effects in menstrual cup take-up, *NBER Working Paper 14828* .

Osterwald, C., J, A., del Cueto, J., Kroposki, B., Trudell, D. and Moriarty, T. (2006). Comparison of degredation rates of individual modules held at maximum power, *Conference Record of the 2006 IEEE 4th World Conference on Photovoltaic Energy Conversion*.

Papineau, M. (2004). An economic perspective on experience curves and dynamic economies in renewable energy technologies, *Energy Policy* **34**(4): 422–432.

Schmidt-Dengler, P. (2006). The timing of new technology adoption: The case of mri, *Working Paper* .

Shriver, S. (2009). Market entry with complementary products: An empirical model of ethanol-compatible vehicle and fuel adoption, *working paper* .

SolarBuzz (2009). Solar module retail price index. http://www.solarbuzz.com/moduleprices.htm, *Accessed November 2009* .

Spence, M. (1981). The learning curve and competition, *Bell Journal of Economics* **12**: 49–70.

Taylor, M. (2008). Beyond technology-push and demand-pull: Lessons from california's solar policy, *Energy Economics* **30**(6): 2829–2854.

Tucker, C. (2004). Two-sided network effects in the adoption of electronic payments, *Stanford Institute for Economic Policy Research Discussion Paper 03-11, Stanford, CA* .

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities, *Management Science* **55**(12): 2024–2039.

van Benthem, A., Gillingham, K. and Sweeney, J. (2008). Learning-by-doing and the california solar initiative, *Energy Journal* **29**(3): 131–151.

Wiser, R., Barbose, G., Peterman, C. and Darghouth, N. (2009). Tracking the sun ii: The installed cost of photovoltaics in the us from 1998-2008, *LBNL Working Paper 2674E. Lawrence Berkeley National Laboratory* .

Young, A. (1991). Learning by doing and the dynamic effects of international trade, *Quarterly Journal of Economics* **106**: 369–405.

Zimmerman, M. (1982). Learning effects and the commercialization of new energy technologies: The case of nuclear power, *Bell Journal of Economics* **13**: 297–310.

Table 1: Zip code summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Zip code number of installations	31.425	49.385	1	516	1588
Zip code MW of installations	0.435	0.79	0.001	9.119	1588
Zip code number of residential installations	27.504	44.192	0	476	1588
Zip code MW of residential installations	0.142	0.236	0	2.218	1588
population (100,000s)	0.244	0.214	0	1.095	1268
household size	2.825	0.611	0	5.21	1268
median income	6.374	2.914	0	37.5	1268
% pop male	50.253	3.689	12.4	98.5	1268
% pop who are white	65.070	20.135	4.4	95.2	1268
% pop with college degrees	38.328	17.537	5.215	95.731	845
% pop between 20 and 45	33.481	8.178	3.9	79.600	1268
% pop over 65	12.371	6.476	0	80.900	1268
% pop who drive to work	86.303	10.139	20.34	100	1273
% pop who carpool	14.758	6.925	0.469	83.529	1252
% pop using public transit	3.896	5.656	0.058	42.593	1011
% pop who work at home or walk to work	8.782	6.866	1.617	61.496	1177
% pop with over a 30 min commute	38.174	12.717	5.371	80.881	1099
% pop who drive a hybrid	2.095	2.119	0	20	1449
number of owner occupied homes (1000s)	4.979	4.204	0	18.965	1268
median value owner occupied home	0.537	0.262	0	1	1268
home loan	121.546	68.211	0	576	1268
home repair	123.121	70.837	0	585	1268
fraction of homes worth 0-50K	2.617	3.858	0	53.2	1268
fraction of homes worth 50-90K	2.072	2.861	0	37.3	1268
fraction of homes worth 90-175K	5.888	7.881	0	61.7	1268
fraction of homes worth 175-400K	30.128	23.239	0	89.7	1268
fraction of homes worth 400K+	58.899	29.979	0	100	1268

Table 2: Installations by type

Residential

Variable	Mean	Std. Dev.	Min.	Max.
size (kW)	5.151	3.671	0.114	210.15
price ($/W)	8.42	4.361	0	697.333
N		46,924		

Non-residential

Variable	Mean	Std. Dev.	Min.	Max.
size (kW)	71.761	194.716	0.99	1707.085
price ($/W)	8.303	2.821	0	107.195
N		5,846		

Table 3: Contractor and Manufacturer Installations

Contractor

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Contractor number of installations	26.658	132.49	1	3426
Contractor MW of installations	0.369	2.511	0	70.768
Contractor number of counties	2.852	4.36	1	50
N		1,919		

Manufacturer

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Manufacturer number of installations	656.618	1579.182	1	7290
Manufacturer MW of installations	9.092	21.778	0.002	118.202
N		76		

Table 4: Installed base and ongoing contracts

Variable	Mean	Std. Dev.	Min.	Max.
installed base (1000s)	26.616	13.312	0.001	41.862
residential installed base (1000s)	24.312	12.169	0.001	38.27
zip residential installed base (100s)	0.512	0.558	0	4.28
manufacturer installed base (1000s)	2.054	1.778	0	6.164
contractor installed base (1000s)	0.308	0.487	0	1.814
competitor installed base (1000s)	15.74	10.251	0	38.544
zip contracts (100s)	0.117	0.129	0	1.12
contractor contracts (1000s)	0.134	0.283	0	1.632
manufacturer contracts (1000s)	0.699	0.55	0	2.5
competitor contracts (1000s)	4.128	2.377	0	11.436
N	50,207			

Table 5: Log time regressions for residential installations

	(1)	(2)
zip installed base	-0.951	-0.136
	(0.081)	(0.034)
SCE x zip-code installed base	-0.962	-0.513
	(0.239)	(0.146)
CCSE x zip-code installed base	0.173	-0.820
	(0.615)	(0.434)
zip installed base squared	0.258	0.019
	(0.023)	(0.004)
SCE x zip-code installed base squared	0.663	0.172
	(0.160)	(0.052)
CCSE x zip-code installed base squared	-0.312	0.172
	(0.301)	(0.173)
previous installation price	0.002	0.001
	(0.002)	(0.002)
step==1		3.908
		(0.572)
step==2	-0.360	3.520
	(0.556)	(0.148)
step==3	-0.366	3.502
	(0.558)	(0.155)
step==4	-0.418	3.438
	(0.560)	(0.161)
step==5	-0.323	3.529
	(0.562)	(0.169)
step==6	-0.255	3.599
	(0.565)	(0.179)
ρ	0.525	0.531
σ_u	0.965	1.130
σ_ϵ	1.057	1.057
R-squared	0.190	0.078
N	35424	35308

Table 6: Effect of demographics on adoption

	(Estimate)	(Sd. Err.)
population (100,000s)	0.925	(0.348)
household size	-0.019	(0.099)
median income	0.028	(0.044)
% pop male	-0.032	(0.013)
% pop who are white	-0.010	(0.002)
% pop with college degrees	-0.010	(0.004)
% pop between 20 and 45	0.001	(0.007)
% pop over 65	-0.001	(0.007)
% pop who drive to work	0.017	(0.011)
% pop who carpool	0.005	(0.008)
% pop using public transit	0.007	(0.011)
% pop who work at home or walk to work	0.029	(0.013)
% pop with over a 30 min commute	0.010	(0.002)
% pop who drive a hybrid	-0.031	(0.017)
number of owner occupied homes (1000s)	-0.094	(0.015)
median value owner occupied home	0.266	(0.234)
home loan	-0.006	(0.005)
home repair	0.004	(0.005)
fraction of homes worth 90-175K	-0.006	(0.012)
fraction of homes worth 175-400K	-0.000	(0.007)
fraction of homes worth 400K+	-0.000	(0.008)
ρ	0.523	
σ_u	0.150	
σ_ϵ	1.057	
R-squared	0.288	
N	32461	

Table 7: Effect of demographics on peer effect (installed base x demographic interactions)

	(Estimate)	**(Sd. Err.)**
population x zip-code installed base	5.067	(8.656)
hh size x zip-code installed base	-2.721	(2.575)
med income x zip-code installed base	-0.899	(1.029)
% pop male x zip-code installed base	0.428	(0.271)
% white x zip-code installed base	0.032	(0.048)
% college x zip-code installed base	0.028	(0.066)
% pop between 20 and 45 x zip-code installed base	-0.190	(0.151)
% over 65 x zip-code installed base	-0.029	(0.155)
% drive x zip-code installed base	-0.335	(0.246)
% carpooling x zip-code installed base	0.192	(0.199)
% public transit x zip-code installed base	-0.471	(0.248)
% work at home or walk x zip-code installed base	-0.566	(0.284)
% pop with over a 30 min commute x zip-code installed base	0.000	(0.058)
% driving hybrids x zip-code installed base	0.401	(0.320)
number of owner occupied homes (1000s) x zip-code installed base	-0.403	(0.351)
median home value x zip-code installed base	-4.623	(5.173)
home loan x zip-code installed base	0.151	(0.112)
home repair x zip-code installed base	-0.110	(0.103)
fraction of homes worth 90-175K x zip-code installed base	0.144	(0.356)
fraction of homes worth 175-400K x zip-code installed base	0.174	(0.205)
fraction of homes worth 400K+ x zip-code installed base	0.168	(0.214)
ρ	0.523	
σ_u	0.894	
σ_ϵ	1.048	
R-squared	0.139	
N	31716	

Table 8: Estimation of structural pricing equation

	(1)	(2)
	OLS	2SLS
manufacturer installations	-0.0215	0.222
	(0.05)	(0.539)
all installations	-0.0117	0.0107
	(0.04)	(0.0743)
contractor installations	0.0363	0.181
	(0.36)	(0.695)
contractor and competitor installations	-0.0929	-0.648
	(0.02)	(0.163)
system size	-0.00199	-0.0172
	(0.00)	(0.00522)
contractor on-going contracts	2.427	4.843
	(0.46)	(2.240)
contractor competitors on-going contracts	0.281	3.167
	(0.06)	(0.843)
manufacturer on-going contracts	0.123	-1.458
	(0.18)	(3.239)
Constant	17.43	18.57
	(1.02)	(1.421)
Time Dummies	Y	Y
Contractor Dummies	Y	Y
Manufacturer Dummies	Y	Y
Observations	49,222	42,373

(Robust standard errors in parentheses)

Figure 1: CSI Incentive Schedule (Source: CPUC (2009))

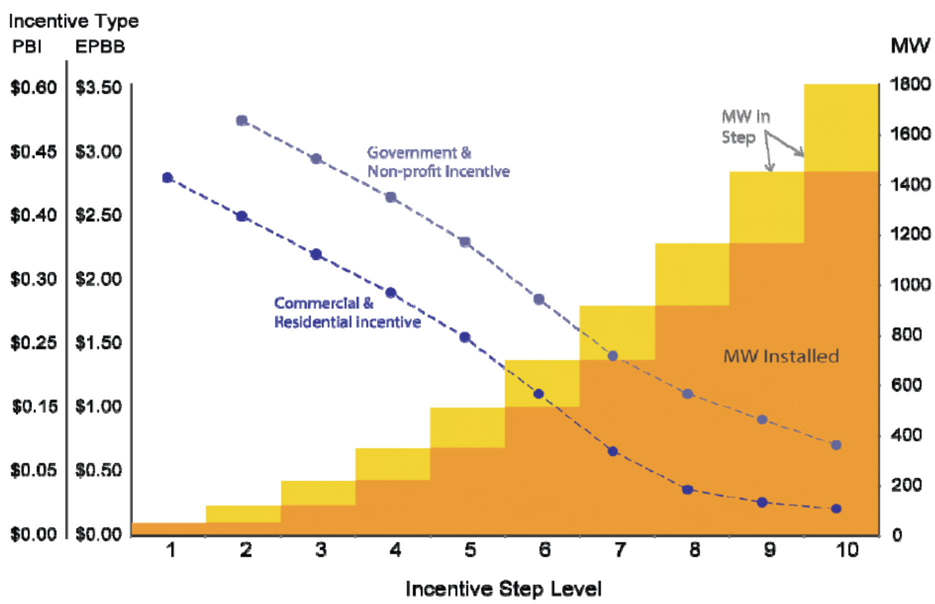

Figure 2: Solar PV Installations in California IOU Regions 2001-2009 (Sources: CPUC, CEC)

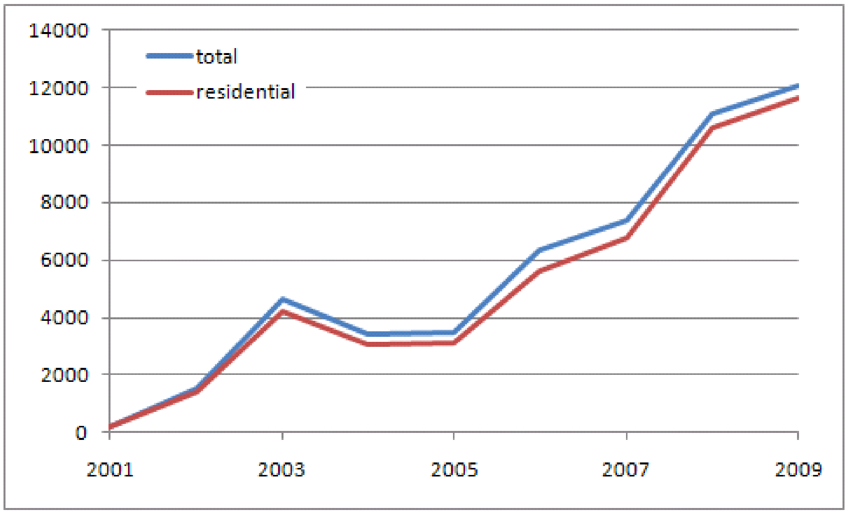

Figure 3: Average Solar PV prices in California 2001-2009 (Sources: CPUC, CEC, Navigant Consulting)

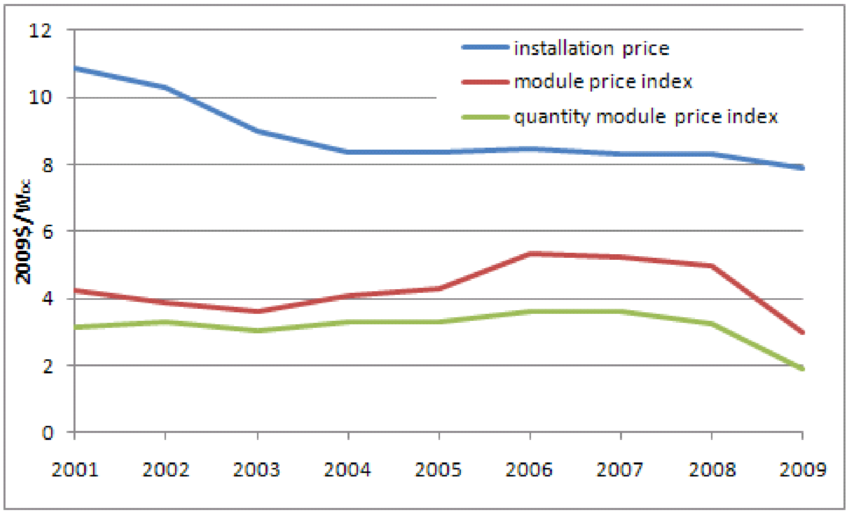

Figure 4: Time Between Zip Code Installations