An Unintended Effect of Subsidizing Green Technology: Lessons Learned from California

Jin Chen*

November 12, 2018

Latest version is available here.

Abstract

This paper presents a previously unrecognized but important consequence of subsidizing green technology in the context of California residential solar market. I show that this market is characterized by many small and local solar suppliers as well as a few large solar suppliers. Large firms advertise heavily and pursue primarily the environmentally focused consumers, whereas small and local suppliers have a limited ability to advertise and pursue primarily the economically focused consumers. Using rich micro-level data, I estimate how subsidies are apportioned between sellers and consumers among different seller-customer groups, and to what extent subsidies have stimulated adoption. The estimate is robust to a spatial discontinuity design. Contrary to many studies in the previous literature, I find very different subsidy pass-through among different groups. Large suppliers capture almost all the subsidy benefits whereas small and local suppliers concede the majority of the benefits to customers. Since customers have different motives to go solar and the solar market is opaque, endogenous matching of consumers and sellers causes this large variation in subsidy pass-through. Over half of the subsidies fail to reduce prices to consumers and encourage adoption. Instead, a large fraction of subsidies mainly yields rents to large solar suppliers to sustain their advertising-supported business model.

*Department of Economics, Stanford University, Stanford, CA 94305 (e-mail: jinc@stanford.edu). I am indebted to my advisor Frank Wolak for his guidance and encouragement throughout this project. I would like to thank the other members of my dissertation committee: Lawrence Goulder and Peter Reiss for generously offering their time, support and guidance. I am also grateful for helpful comments from Tim Bresnahan, Liran Einav, Mathew Gentzkow, participants at the Berkeley Energy Camp 2018, and seminar participants at the Stanford IO Workshop and Stanford SEEPAC Lunch. I am grateful for financial support by the Bradley Graduate and Postgraduate Fellowship through a grant to the Stanford Institute for Economic Policy Research. All errors are my own.

1 Introduction

The typical motivation for a subsidy is to reduce the price paid by consumers in order to increase demand for a product. In a competitive market, the extent to which a subsidy achieves this goal depends on the slope of the aggregate supply curve. A flat supply curve implies that a greater share of input cost increases or decreases is passed through to the market price. Under these conditions, the price paid by consumers will fall by the amount of the subsidy paid, which stimulates demand for the product.

In the context of environmentally friendly (or "green") products, the impacts of prices to consumers can be more complicated. That is because consumers have different motivations for purchasing green products. A green product provides the household with an environmental "warm-glow" as well as a replacement for the service provided by a less environmentally friendly product. For example, a rooftop solar system allows a household to reduce its carbon footprint by replacing grid-supplied electricity with energy from its solar panels. Depending on the price a household pays for a rooftop solar system, the energy it supplies may be less expensive than grid-supplied electricity.[1]

These two motivations for purchasing a green product, such as a rooftop solar system, imply that suppliers of these products may pursue different marketing and sales strategies. One set of suppliers might focus on the environmentally motivated buyers that purchase primarily for the environmental "warm-glow" and another set of suppliers might focus on the economically motivated buyers that purchase primarily to save money on their monthly electricity bills.

This paper analyzes and documents the market segmentation that has emerged in the California rooftop solar industry and examines the implications of market segmentation for the effects of the rooftop solar subsidy among different consumer-supplier groups. I provide evidence to support the logic that suppliers of rooftop solar systems in California have segmented themselves into those that primarily pursue the more environmentally motivated consumers and those that primarily pursue the mainly economically motivated consumers. The suppliers that focus mainly on the more environmentally oriented consumers tend to spend large sums of money on advertising. These advertisements focus on highlighting the climate benefits of installing a rooftop solar system. The suppliers that focus on the less environmentally oriented consumers tend to spend very little on advertising, primarily on the design of their web-sites. The advertisements on their web-sites are mainly about the economic benefits of a rooftop solar system. These two strategies have different implications regarding the extent to which subsides reduce the prices to consumers and increase the subsequent sales volume.

[1] For instance, the levelized cost of electricity of rooftop residential solar dropped from \$0.18–\$0.30/kWh in 2015 to \$0.14--\$0.22/kWh in 2016 by calculations of Lazard. As a comparison, California households paid on average \$0.18/kWh for electricity in November 2016.

One would expect the purchase decision of environmentally motivated consumers to be less sensitive to the price of the rooftop solar system than is the case for economically motivated buyers. This is borne out by my analysis of the pass-through rate for subsidies for solar systems paid under the California Solar Initiative (CSI). As a result of the market segmentation, the more environmentally motivated consumers end up receiving a smaller price reduction from the subsidy than consumes that focus more on cost. Buyers from the environmentally focused firms also tend to live in higher-income areas, although the income gap between these two buyer groups appears to have closed over time.

I provide further evidence that environmentally focused buyers purchase primarily for the warm glow. I find that the same kilowatt capacity rooftop solar system installed by an environmentally focused supplier produces significantly less energy than a system installed by an economically focused supplier. Consequently, environmentally motivated buyers pay higher net-of-subsidy prices for lower quality systems.

This result has implications for the economic impact of the rooftop solar subsidies provided by the CSI. Specifically, I find that the majority of subsidies were paid to systems sold by environmentally focused firms. This means that a large proportion of the subsidies provided by the CSI yielded rents to these firms and did not stimulate as much of an increase in the demand for rooftop solar systems as would have been the case if the systems had been sold by economically focused firms. For these economically focused firms, I find that virtually all of the subsidy was passed through to consumers in the form of lower prices for solar systems.

California's distributed solar market is an ideal setting for the empirical analysis. California is the leading rooftop solar market in the US. In addition, California had a large solar incentive program (the CSI),[2] which was one of the largest solar energy subsidy programs in the US. The unique design of the subsidy program causes multiple sharp jumps in rebate levels across time and regions, which allows for significant price and rebate variation to identify the subsidy impact on prices. In particular, this subsidy program is divided into three independently administrated regions based on the utility service territory and within each region the subsidy level fell to a lower level whenever the cumulative installed solar capacity reached a pre-determined quantity target. There are nine quantity targets for each region and these targets also differ across regions. Third, the rich micro data that I have compiled on solar suppliers allows me to document their marketing strategies and to examine the market structure. Finally, thanks to the rich data and the way how a solar system works, I am able to construct a performance metric to measure the quality of the product.

[2]Besides the California Solar Initiative, there is a 30% federal tax credit program at the federal level. However, this is not a big concern as this federal program remained (and general buyers also expected) a constant subsidy rate during the sample period being studied.

Using transaction-level data from CSI, I first estimate how subsidies are apportioned between consumers and sellers in the California Solar Initiative program by measuring an overall subsidy pass-through rate. I use a difference-in-difference (DD) strategy which follows the method of Henwood (2014) to estimate the subsidy pass-through rate. The idea is that since the subsidy level is a function of both pre-determined targets and realized aggregate installed capacity, a single firm or consumer has limited capability to manipulate the timing of a subsidy change. As a result, an individual firm or individual consumer treats the subsidy level as given, and how much between-region price level variation can be explained by between-region rebate level variation serves as a good estimate of subsidy pass-through. Using this approach, I find that the overall subsidy pass-through rate is around 0.63 to 0.86 depending on the model specification, meaning that the majority of the subsidy goes to solar retailers and the subsidy has a limited effect in terms of encouraging solar adoption.

I then take a closer look at the micro-data and examine subsidy effects among different consumer-supplier groups. By enriching the CSI data with California contractor database and by analyzing the text contents of solar supplier's websites, I document that (1) the market is dominated by the top few sellers. Among the over two thousand solar sellers, the total market share of the largest 1% sellers is over 50%. However, the total market share of the bottom 50% sellers altogether is only around 1%; (2) top sellers and small/local sellers have distinct advertising and marketing strategies. Top sellers spend a significant share[3] of their revenue on various kinds of sales and marketing activities, and are more likely to focus on pointing out the climate benefits of solar energy on their websites. On the contrary, small sellers barely advertise. Most of them do not even have a website; and (3) top sellers serve customers in regions that are wealthier or more environmentally-conscious compared with the ones served by small and local sellers. I argue that this evidence suggests that the market is segmented into two groups. One group primarily consists of large suppliers and consumers who are more environmentally motivated. The other group primarily consists of small and local suppliers and consumers who care about saving on energy bills.

Given these stylized facts that are in support of market segmentation, I categorize the solar companies into three groups: large sellers (top 1% companies), medium-sized sellers (top 1%–10% companies), small sellers (all other companies) in terms of total sales, and analyze the subsidy effects in each group. I find that large companies nearly capture all of the subsidy benefits which is much higher than what medium-sized and small solar companies capture. In fact, small and local suppliers concede almost all subsidy benefits to consumers. Large companies also charge higher unit price than small and local companies, which is in line with a recent research (O'Shaughnessy and Margolis, 2017) suggesting that large solar companies quote \$2000–\$5000 more for the same solar system than small companies. Why do top solar firms capture so much more subsidies than

[3] For the publicly listed companies, the ratio of sales and marketing expenses to total revenue is around 20%–60%.

smaller firms? By utilizing a detailed data of real energy outputs of around 500 solar systems for 5 years across California, I construct a performance metric for solar systems. This performance metric can be thought of as how a solar system really performs given its nameplate size.[4] I find that this performance metric is an excellent predictor of solar system's future kWh output normalized by the upfront cost. However, top 1% sellers consistently sell solar systems that are poorer in terms of the performance metric compared with other sellers. This rules out the possibility that the high subsidy pass-through enjoyed by the large sellers is because they provide better quality products.

Solar system prices are formed through bilateral negotiation between consumers and sellers, so one might worry that such unobserved bilateral negotiation (which affects price), when aggregated at the region level, may also be correlated with the realized subsidy levels. For example, if households in San Diego have become increasingly more passionate about solar energy than households in the Los Angeles region, they would have less bargaining power over the years compared to their counterparts in LA. This unobserved factor leads to an increasingly higher price level (households willing to pay more) as well as a lower subsidy level (higher than expected demand causes the subsidy to step down more quickly) in the San Diego region. To resolve this issue, I propose a spatial discontinuity strategy as a robustness check. For historical reasons some households in the same city limits or even in the same zip-code area are assigned and served by different utility companies,[5] which means they could face very different subsidy levels at the same time. By exploiting this spatial discontinuity, I can analyze how solar system prices differ for nearly identical groups of households living in the same city but facing different subsidies. Results of this robustness check are consistent with my findings that top solar sellers, which altogether constitutes a big market share though there are only few of them, capture nearly all the subsidy benefits. On the other hand, medium-sized and small solar sellers concede almost all benefits to consumers.

The rest of paper is organized as follows. Section 2 presents an overview of the related literature. Section 3 provides background knowledge about the California solar market and the solar subsidy program. Section 4 describes the data. Section 5 presents evidence in support of market segmentation. Section 6 describes the empirical strategy used to estimate the subsidy effects. Results and discussions of this analysis as well as the robustness checks are reported in Section 7 and 8. Finally, Section 9 concludes.

[4] Nameplate size, as known as the nominal capacity or rated capacity, is the nameplate wattage of a solar system.

[5] This fact was firstly used by Ito (2014) to examine consumers' response to electricity price.

2 Related Literature

This paper is related to prior literature in three main ways. First, this paper extends earlier work on consumer heterogeneity in greenness and market segmentation to assess their impact on the effects of subsidies. A series of papers and articles (Kahn and Matsusaka, 1997; Sterner and Coria, 2003; Kahn, 2007) document the existence and emergence of the environmentally-conscious consumers. Though there is abundant evidence of the emergence of green consumers, very little empirical work has discussed the implications of green consumers for market segmentation. These studies usually rely on the analysis of survey data. For example, Paco and Raposo (2010) survey what demographic variables are significant in differentiating between the "greener" consumer group and the other segments in Portugal. They find that green customers are associated with higher education levels and earning higher incomes. Also relying on survey data, Zhang and Wu (2012) provide evidence that households with high income and higher education tend to have a higher willingness-to-pay for green electricity in China. However, according to a long strand of theoretical papers (Dickson and Ginter, 1987), the multi-attributes property of a good[6] is likely to cause heterogeneity in demand functions and further may cause market segmentation, where the total market consists of segments whose within-group differences in individual demand functions are relatively small compared with the between-group difference. There are research opportunities to use rich micro-level data to test for this theoretical prediction and to examine the implications of market segmentation for subsidy effectiveness.

Second, the paper adds another dimension to evaluate environmental subsidies. Existing literature has shown mixed evidence of the effects of solar subsidies in the California context. Henwood (2014) is among the first to estimate subsidy pass-through in California's solar market and concludes that about half of the subsidy benefits go to consumers. On the contrary, using the same data, Dong et al. (2018) find that almost all subsidy has gone to consumers. Pless and van Benthem (2017) find similar results for customer-owned solar systems to the ones in Dong et al. (2018) and go beyond by collecting contract data of the third-party owned (TPO) systems, and find an overshooting subsidy pass-through rate in the TPO market. In this paper, I show that by considering consumer heterogeneity and market segmentation these estimates can actually be harmonized. Besides the solar subsidy, there are papers examining the effects of subsidies of electric or hybrid vehicles. For instance, Sallee (2011) finds a high subsidy pass-through to Prius consumers and suggests that this is due to an inter-temporal link in pricing that stems from search friction. For evaluating all these green subsidy programs, this paper highlights the importance of consumer purchasing motives, which affect the market structure, in determining a subsidy's effectiveness.

Third, the paper contributes to explaining the price dispersion that has been found in the solar

[6] A green good, with environmental concern as one attribute, usually has more than one attributes.

market. A series of empirical studies and reports have documented wide price dispersion in the residential solar market (Gillingham et al., 2014; Nemet et al., 2017a; Nemet et al., 2017b). In addition to this wide dispersion of equilibrium prices, quote prices also vary significantly according to Dorsey (2017) and O'Shaughnessy and Margolis (2017). These studies highlight the importance of information and market structure in the price formation process. I illustrate in this paper that a substantial part of the price dispersion can be explained by the dramatically distinct subsidy pass-through for different consumer-supplier groups.

3 Background

3.1 Green Technology and Subsidies

Green technology, or environmental technology, is the application of science and technology to model, monitor, and conserve natural resources, in order to mitigate the negative impacts of human activities on the environment. In other words, green technologies aim to produce energy in an environmentally friendly way. Examples include power generation using renewable resources such as wind, sun or water, sustainable transport such as electric or hybrid vehicles, and home appliances, improvements or equipment that are energy efficient.

Subsidies for green technology have been on the rise. For example, in the US, subsidies for renewable energy at the federal level have grown tenfold from $1–2 billion in 2000 to over $15 billion in 2010.[7] Globally, total green incentives were about $66 billion in 2010. The International Energy Agency forecasts that the total green incentives will reach $250 billion by 2035.[8] A common type of subsidy for green technology is direct financial transfers or tax credits to consumers. For instance, consumers who buy an electric vehicle in the US are eligible for a tax credit of $2,500 to $7,500. There are also numerous rebate programs for Energy Star home appliances.

The goal of these incentive programs is to reduce the price that consumers pay and to stimulate the adoption of green technology. From the classic view on subsidy, the extent to which this goal is achieved depends on the relative slopes of the demand and supply curves. If the supply curve is relatively inelastic, then the subsidy will result in an increase in price but little increase in quantity. In this case, though consumers get the rebate or tax credit from the subsidy program, they have to pay a higher price to suppliers. In an extreme scenario where the supply curve is vertical, consumers get zero benefits, whereas suppliers capture the full subsidy benefits. In this case, the subsidy program effectively doesn't stimulate adoption at all. On the other hand, a relatively elastic supply curve implies that the market equilibrium price will not be very responsive when a subsidy

[7] Source: U.S. Energy Information Administration, Direct Federal Financial Interventions and Subsidies Reports.
[8] Source: IEA World Energy Outlook 2011.

kicks in. If the supply curve is very flat, then consumers capture the full benefits and the subsidy program achieves the most effective result in terms of stimulating adoption.

The implications of subsidies for green technology, however, are more complicated than what the discussion above suggests. This complexity is due to the unique nature of green goods. A green good is a substitute for a conventional good. Therefore, first, it produces or provides the service that can also be supplied by a conventional good. For example, households can purchase a home solar system for its electricity outputs so that they do not need to pay monthly energy bills to the utility company. Under some circumstances, the levelized cost of electricity of a home solar system could be even lower than the average electricity price. If this is the case, it is economically beneficial to adopt the green product. I call this the *practical-use attribute* of the green good and call consumers motivated by this attribute the *economically focused* consumers. Second, a green good can also provide the consumer with an environmental "warm glow." For example, adopting an environmental good can demonstrate the buyer's social standing, social identity, or views and attitudes toward the environment. I call any consideration (environmental consciousness, social statement, or pure altruism) other than purely practical or economic concerns the *environmental "warm glow"* motive.

The practical-use attribute and the environmental "warm glow" attribute of a green good can cause market segmentation because different consumers may care about different attributes and firms may develop distinctive sales and marketing strategies to pursue each consumer group. In each market segment, a subsidy program is likely to have different effects and may also affect the market segmentation equilibrium.

To see this, suppose the total market demand is represented by

$$Q = F\left(p, x_1, x_2, ..., x_n\right).$$

Here the demand function $F(.)$ is consistent with the multi-attribute model commonly used in the marketing literature. It is a function of both the product price and a vector of consumer perceptions[9] of product characteristics. The product characteristics include both physical product attributes (e.g., quality and efficiency of the components) and nonphysical attributes (e.g., the product's greenness and image).

It is unlikely for all consumers to respond equally to a market offering $(p, x_1, x_2, ..., x_n)$. Therefore, the market demand can be dis-aggregated into segments $i = 1, 2, ..., I$ with distinct demand functions $F_i(.)$ where within each segment consumers have similar responses to a market offering.

[9] It is assumed here that consumers may not observe full information on product characteristics. Consumers make the purchase decisions on their perceptions of the product characteristics.

I define market segmentation as a state of demand heterogeneity such that the total market demand can be dis-aggregated into segments with distinct demand functions, and consequently, distinct market offerings. Its formula is as follows.

$$\begin{aligned} Q &= F\left(p, x_1, x_2, \ldots, x_n\right) \\ &= \sum_i Q_i \\ &= \sum_i F_i\left(p, x_1, x_2, \ldots, x_n\right) \end{aligned}$$

In our context, the market is likely to be segmented into an environmentally focused consumer segment ($F^G\left(p,g,e\right)$) that puts a relatively higher weight on the green attribute g in the purchase decisions and an economically focused consumer segment ($F^E\left(p,g,e\right)$) that puts a relatively higher weight on the quality attribute e.

$$Q = F^G\left(p,g,e\right) + F^E\left(p,g,e\right) \tag{1}$$

The intervention of a subsidy program may result in distinct effects on the equilibrium market offering $\left(p^*, g^*, e^*\right)$ in different segments. To achieve the most effective outcome of a subsidy program[10] in terms of reducing the price that consumers pay and stimulating adoption, we want, in each segment, that p^* to remain low when the subsidy goes into effect so that the subsidy effectively reduces the price paid by consumers. In addition, for those subsidies captured by suppliers, we want them to be captured by the suppliers in the segment where e^* is larger so that the subsidy appropriately creates incentives for firms to invest in R&D and to improve product quality. I show empirically in this paper that this is highly unlikely to be the case because of the environmentally focused segment. Moreover, I demonstrate that the environmentally focused segment is quite large and causes an undesirable consequence, namely, a substantial transfer of the rate-payer funded subsidy to several large solar firms that provide low quality products.

3.2 Distributed Solar Systems

The particular green technology examined in this paper is distributed solar systems. These are small solar systems, usually less than 10 kilowatts in capacity, that can be installed on a house roof. They are also called rooftop solar systems, and are typically connected to the local utility distribution grid. By installing a solar system, households can pay less to the utility company for electricity.

[10] Assume it is a consumer subsidy program in which consumers get a rebate after buying.

Consumers can choose two different ways to adopt a solar system. First, they can choose to buy and own the system. This is the so-called customer-owned system. Second, they can choose to lease from or sign a Power Purchase Agreement (PPA) with a third party, which is the real owner of the solar system and usually is the solar system seller. This is the so-called third-party-owned system. For lease systems, consumers pay a fixed monthly fee to the third party for the electricity outputs produced by the solar system. A PPA system is similar. The only difference between a PPA and a lease is that the monthly fee of a PPA is a function of power produced by the solar panels so that it could vary by month. Because of data limitations,[11] this paper will only examine consumer-owned solar systems.

In this paper, I construct a performance metric for solar systems, thanks to the mechanism of how it works. A typical distributed solar system consists of two main parts - solar modules/panels and an inverter (Figure A.2). Solar panels collect the sunlight and turn it into direct current (DC) electricity. The DC signals are fed into an inverter that converts the DC into grid-compatible alternative current (AC) power, which is what we finally consume. Every solar panel has a published power rating under Standard Test Conditions (STC), also called Peak rating. This measures the DC power that solar panels can produce under a set of ideal conditions. Though the STC-DC rating is an industry standard and is known by the general solar consumers, a solar-savvy consumer may be more interested in the CEC-AC rating.[12] This rating measures the final AC power output that can be directly consumed of a solar system under a more realistic rooftop environment, which also accounts for the inefficiency introduced by the DC to AC inverter.[13] Therefore, the ratio of the CEC-AC rating and STC-DC rating can be used as a performance metric of the solar system as a whole, which I will discuss more in a later section. This ratio can be thought of as the how the solar system really performs given its nameplate size.

$$\text{performance metric} = \frac{\text{CEC-AC rating (CSI rating)}}{\text{STC-DC rating (Nameplate rating)}}$$

3.3 California Solar Initiative Program

California has a big solar market. Almost 40% of America's distributed solar capacity is located in California and the majority of the distributed capacity was installed over the last decade as shown in Appendix Figure A.3, thanks to the comprehensive California Solar Initiative (CSI) Program.

[11] It is difficult to observe the price of a PPA/lease system, because it is a contract between the buyer and seller and its details are usually not shared with researchers.

[12] CEC-AC rating is harder to be observed by consumers, as it is in general not known until the purchase process comes to the last few steps, e.g., under contract.

[13] STC rating is measured in terms of direct current (DC) and under the conditions of 1,000 Watt/square meter solar insolation and 25 degree Celsius cell temperature (which assumes roughly 0 degree Celsius ambient temperature). The "CEC-AC" rating standards are based on 1,000 Watt/square meter solar irradiance, 20 degree Celsius ambient temperature, and 1 meter/second wind speed. It is measured in terms of alternative current (AC).

This subsidy program started in 2006 and ended in 2015. This program is under the Go Solar California campaign, which is a joint promotion program led by the California Energy Commission and the California Public Utilities Commission. Its goal is to encourage California residences to install solar energy systems on homes and/or businesses with a target of 2000 MW of new solar generation capacity. The funding for the CSI program comes from electric ratepayers. Rebates are offered through the program administrators in the form of cash.

To be specific, in the California Solar Incentive Program, solar buyers get a direct cash rebate based on the solar capacity they install. For example, a consumer who buys a solar system of C kW can get a cash rebate of $R = L \cdot C$ dollars, where L (dollars/kW[14]) is the rebate level. The program is divided into three utility regions, and each region is exclusively administered by one of the three investor-owned utilities (IOUs) in California—the Pacific Gas and Electric company (PG&E), Southern California Edison (SCE) and San Diego Gas & Electric (SDG&E).[15] Consumers have to participate through their assigned IOU. Within each utility region, the rebate level L (dollars/kW) is a non-increasing step function of the cumulative capacity of solar systems installed in that region. The trigger values in each utility territory which determine the shape of the function are pre-determined by the program administration before the program started. To be more specific, the trigger values are three sets of capacity sequences for each of the IOUs as illustrated in Table 1. In each utility territory, whenever the cumulative solar capacity reaches a pre-specified trigger value, the rebate automatically steps down to the next level. Therefore, at the same point of time different utility territories could have different rebate levels.

In the data, I observe nine rebate changes for each utility territory in total and the realized rebate schedules are different across regions. This cross-region rebate level variation, therefore, can be utilized to measure the subsidy effects. For example, the equilibrium price in utility PG&E can be treated as the basis for the counterfactual price in utility SCE if the subsidy level was higher (or lower). Essentially, this allows us to use differences-in-differences statistical technique to estimate the subsidy effects.

The typical procedure to participate in the rebate program and to get the rebate is as follows. First, potential consumers may check with local utility company to get information about their household energy usage. Then consumers will reach out and find qualified solar sellers that meet their needs and budget. Solar sellers are supposed to help consumers understand the costs and benefits of installing solar systems according to consumer's specific home attributes. Consumers then choose a solar seller and a contract plan. Once the contract is signed, the seller would help consumers prepare paperwork and apply for the rebates on consumer's behalf. When the complete application is received by program administrator, rebate fund is reserved for the project. The fund

[14] Here kW is AC-rated.

[15] Center for Sustainable Energy (CSE) administers the CSI Program in SDG&E's service territory.

reservation is based on a first-come-first-reserved criteria. The next step would be solar system installation, typically by the same company who sold the system. Finally, consumers will receive cash rebates after installation is verified by the incentive program administration. Consumers rely heavily on solar sellers to obtain relevant information, choose between different solar brands and contracts, and participate in the rebate program.

4 Data

4.1 Main Data Sources

The paper uses data from the California Solar Initiative (CSI). The CSI data is built by collecting and reviewing information for each of the solar project that has claimed rebates in the CSI program. It collects 126 variables for over 174,000 observations during the program period 2006–2015. One observation is simply one solar project. For each solar project, we see a sequence of dates that specify the timeline of project application, submission, completion and verification. The data also tells the detailed engineering information of the solar system, including system size, panel brand, model and quantity, inverter brand, model and quantity, as well as price information, including transaction price and rebates of the project. Information of consumer (zip-code location, consumer-owned or TPO, utility territory, rooftop environment) and seller (name, contractor license number) for each solar project are also provided.

Since the transaction price of solar systems is crucial in our analysis of subsidy pass-through, understanding the differences between customer-owned and third-party-owned systems is necessary. According to the rebate program administration, the reported costs for customer-owned systems are simple, as they reflect the purchase price inclusive of parts, labor, permitting fees, overhead, and profit. Third-party-owned systems, on the other hand, could be reported in a variety of ways, and may also capture costs for additional services. To be specific, there are at least three different ways third-party owners may report their system costs. First of all, third-party owner may report the sale price as the system price, which does not include the overhead and profit of third-party owner. Second, the third-party owner may report the Fair Market Value, which is an estimate of the market value of the solar property. Finally, the third-party owner may report the appraised sum of cost inputs. Moreover, some third-party-owned systems also include the other costs like inverter replacements, roofing replacements, etc. Therefore, I drop the third party-owned systems when conducting price related statistical analysis. This reduces my sample by half.

Since I focus on the residential solar market, I also drop observations of large commercial projects, installations for low-income residents, and systems for multifamily affordable housing.

The final dataset also drops projects that are canceled, withdrawn, or delisted, and only keeps projects with a reported rebate amount smaller than its reported transaction cost. The following summary statistics and analysis are based on the final dataset.

4.2 Other Data Sources

I supplement the main CSI data with auxiliary solar seller (solar contractor) information, consumer demographics, solar system performance, and marketing information of solar companies.

Solar Seller Information

CSI requires all participating solar contractors and sellers to register on the CSI database so that potential consumers can obtain relevant information on sellers and contractors through an online search tool. I scraped each contractor's online page in this database. Each page reports the company's website link (if available), physical location, service area, type of services offered, number of employees, CSLB (Contractors State License Board) license number, year of the license establishment, and type of contractor licenses that are held.[16] The total number of contractors/sellers is 3708, among which 1370 can be finally matched to the main CSI data through a mapping using either the contractor license or seller name. Given the CSLB license number, I can also identify the license expire date for each contractor from the CSLB database. This data is helpful when examining the market structure.

Consumer Demographics

The consumer demographic data comes from the census ACS 2010–2014, which is at the ZIP Code Tabulation Area (ZCTA) level. I merge the demographics with the main CSI data, which is at the zip-code level, through a ZCTA-zipcode mapping. Demographics include median household income, median home value, and education level in each zip-code block. I also collect motor vehicle registrations by fuel type in California at the zip-code level. This data is used to construct a measure of the "greenness" of consumers.

[16] According to the CSI clarification, in order to be eligible for incentives, solar energy systems shall be installed by individuals with a current A (general engineering), B (general building), C-10 (electrical), or C-46 (solar) contractor license, with a limited exception for systems that are self-installed by the purchaser. Roofing contractors with a current C-39 license may place PV panels in accordance with the limitations of their license; however, electrical connections shall not be made by a roofing contractor. The database illustrates that other (not required) common contractor licenses held by a solar contractor include C-20 (warm-air heating, ventilating and air-conditioning), C-36 (plumbing), and C-53 (swimming pool) as sometimes a solar project is connected to these systems.

Solar System Performance

The CSI program collects solar energy output information of 543 solar systems since its installation for 60 consecutive months. This piece of information, combined with the main CSI data-set, helps me identify the key predictors of solar system output/performance.

Marketing Information of Solar Companies

I scraped the web-pages of approximately 600 solar companies with websites. For solar companies that are publicly listed, I collect and compute their annual sales and marketing expenses since IPO. This data provides information on firms' sales and advertising strategies.

4.3 Descriptive Statistics

As an overview, Figure A.4 shows the distribution of the CSI solar projects across the three utility regions in California. The larger markets are the San Francisco area (PG&E) and Los Angeles area (SCE). The San Diego area is a relatively small market compared with the other two. Table A.1 shows that the average cost of a solar system in the sample is about 35,000 dollars and on average around 10% of the cost is subsided by the incentive program. Figure 1 illustrates the variation of rebate levels across time and across region.[17] At the same point of time, the subsidy level could be very different among two utility territories. For instance, around the beginning of 2010, the rebate was 1100/kW in PG&E and CSE; whereas it was 1900/kW in SCE, which is 73% higher than the level in the other two regions. Figure 2 shows how consumers respond to the rebate steps in each of the three utility territories. Transaction bunching (the green line) occurs as a subsidy step is being approached, implying that consumers and/or contractors are aware of the particular incentive design and respond to it. Figure 2 also shows that the cost of the solar system is reduced dramatically over the sample period from around $10000/kW to only $5000/kW.

Dates

A sequence of date variables in the data sample describes the application and installation cycle of the solar project, such as the first new application request date, the first online reservation request date, and the first reservation request review date. Among these dates, I use first reservation request *review* date (I call it the "review date" for short) as the representative date of application for the following reasons. First, this date is the indicator of application completion and rebates are

[17] The CSI Program is directly administered by each investor-owned utility, with the exception of SDGE which is administrated by CSE instead.

reserved based on this date. In the data sample, the review date is the same date of the submission of complete application. Second, applications are reviewed on a first-come-first-reviewed basis. The review date therefore can nicely keep the order of application. Third, the review date has the least missing values compared with other date variables. 97.8% of the observations have a review date.

Transaction Prices

The transaction price of a solar system includes three components. The first component is physical equipment cost like the costs of solar panels and inverters. The second component includes other non-physical costs like marketing and operation costs, installation costs, and warranty. The third component is markup of sellers.

There is evidence that firms cut prices near rebate steps. To show this, I compute the average prices before and after a subsidy change date. Figure 3a illustrates how daily pre-rebate transaction prices evolve before and after the rebate changes in the PG&E territory. The horizontal-axis shows the days before and after changes. Day 0 indicates that rebate change is undergoing on that day. Negative days indicate days before changes, whereas positive days indicate days after changes. Two polynomial fitting curves of degree 2 that are before and after rebate change illustrate that pre-rebate transaction prices right before rebate changes are lower than those right after rebate changes. However, the subsidy level is higher before a rebate decrease and that is why the post-rebate prices to the left of the rebate change date shift even lower, as in Figure 3b. This implies that sellers are actually cutting prices right before rebate changes and the price cuts are large in magnitude. This finding[18] motivates me to do a robustness check by dropping out observations near the rebate change date in Section 7. The size of the dots represents the number of transactions. It shows that consumers are scheduling projects right before rebate changes.

Licensed Contractors (Sellers)

As discussed in Section 3, consumers rely heavily on solar contractors (sellers) to understand the benefits and costs of a solar system. In California, all solar systems must be installed by a licensed solar contractor. Solar contractors are required to register with the California Energy Commission through its online database with a unique contractor license number. Throughout this paper, a solar company or a solar seller is defined as a solar license holder.

[18] Similar price and quantity patterns are found in the other two utility territories.

Market Structure

Define the overall HHI of the California market as a whole of each month as follows.

$$\text{HHI}_t = \sum_{i \in I} \left(s_t^i\right)^2$$

Here $s_t^i = \frac{q_t^i}{\sum_{i \in I} q_t^i}$ is the market share of firm i in a certain month t; q_t^i is the number of systems sold by firm i in month t; I is the set of firms that made positive sales in month t. Figure 4 shows that the HHI is in the range of 0.01 to 0.08, implying that overall it is a quite competitive market. The overall HHI went up after year 2010, which indicates that the market became a little bit more concentrated towards the end of the sample period. Figure 5 shows the share of customer-owned systems against the share of TPO (leased or PPA) systems. Significantly more systems are sold as TPO rather than customer-owned after 2010. This pattern is mainly driven by the top firms in the market (Figure A.5). On the other hand, the customer-owned system market does not witness a dramatic change in the market share of top, medium-sized, and small firms over time (Figure A.6).

5 Evidence of Market Segmentation

In this section, I show that the suppliers in the California residential solar market have segmented themselves into two groups. One group, consisting mostly of large solar firms, spends large sums of money in sales and advertising and is primarily focused on consumers that are more environmentally motivated. The other group, mostly consisting of small and local solar sellers, spends little in sales and advertising and is primarily focused on consumers that care more about saving on energy bills by installing a solar system.

5.1 Market is Dominated by the Few Top Firms

Although there are over two thousand solar sellers in California, the residential solar market is dominated by a few companies. To show this, I compute the total number of sales for each seller in the sample and rank them from the highest-sales company to the lowest-sales company. Table 2 reveals the average solar company sold about 58 solar systems during the sample period. However, a top 1% seller on average sold more than 3000 solar systems and their total market share is over 50%. A typical bottom 50% solar company only made 1 or 2 sales during the whole sample period (which is 8–9 years). Their total market share is just about 1%. Top 1% sellers also charge a higher price than other sellers. Figure A.1 shows how average transaction prices differ among top 1% sellers and other sellers. In each utility territory, price offered by top sellers was almost always larger than the one offered by other sellers except for those initial periods.

The reason for this dramatic difference in firm size is that solar suppliers have very distinct business models. Top firms, like SolarCity and Sunrun, are usually national companies and publicly listed companies. These firms have a financial advantage and can purely focus on the sales of residential solar systems. Small sellers, on the other hand, are usually contractors that were contractors for other house construction or improvement services before the emergence of rooftop solar. Their business model doesn't rely on the sales of rooftop solar, because they also provide other home improvement services besides selling and installing the solar systems.

5.2 Firms of Different Sizes Have Different Sales and Marketing Strategies

In order to show that top firms and small firms advertise differently and focus on primarily different attributes of the solar energy to pursue different consumers, I conduct text analysis on the solar seller web-pages. Among the 2309 solar sellers, about 74% of them do not have a website[19] and I call them Group 2. For those companies that have a website (589 companies), I scrape the texts of their homepages and apply an unsupervised machine learning method to categorize them into two groups[20] — Group 0 and Group 1. For Group 0, the predictive words are "sunpower", "best", "repower", "america", etc. For Group 1, the predictive words are "construction", "electrical", "service", "projects", etc. Group 0 sellers are focused more on the "green" and "cool" attribute of the solar system, whereas Group 1 sellers emphasize the engineering or practical-use attributes of the product. These descriptive words reflect the way how firms advertise.

Larger firms are more likely to be categorized into Group 0 and less likely to have no website, as shown in Figure 7. On the contrary, it is very unlikely for small and local sellers to be categorized into Group 0. In fact, the majority (77%) of them do not even have a website. This is in line with our discussion above. Due to their business model, small and local solar sellers do not have the financial capability to do market research and sophisticated advertising. However, top firms, which receive findings from venture capital or IPO, have the financial capability to develop sophisticated sales and marketing strategies. For instance, SolarCity spends hundreds of millions of dollars in sales and marketing.[21] In certain years, SolarCity spent more in sales and advertising than their total revenue (Figure 8).

I have shown that (1) firms are dramatically different in their sizes and business models and (2) firms of different sizes have different marketing strategies. In particular, top firms tend to do

[19] Or I got access error when I try to scrape it.

[20] The advantage of unsupervised learning is that the clustering is purely objective (purely let the data talks). I can also count the frequency of "environmentally-related" words against the "practical-use-related" words. But this method is more subjective (I have to impose a subjective set of environmental words and a set of non-environmental words). Here the unsupervised learning method I use is k-means. More details are in Appendix.

[21] SolarCity market and sell its products and services through a national sales organization that includes a direct outside sales force, a door-to-door sales force, call centers, a channel partner network and a robust customer referral program, which is far beyond the website design.

sophisticated advertising and give a "green" impression. Small and local firms do little advertising and focus more on the practical-use attribute of the product. Next, I show evidence that these two types of firms are also focused on different types of consumers.

5.3 Endogenous Matching Between Consumers and Suppliers

Solar firms of different sizes primarily pursue different consumer groups in terms of household income[22] and greenness. To show this, I again divide firms into three groups based on total sales: top 1% firms (top firms), top1%–10% firms (medium-sized firms), and bottom 90% firms (small and local firms). For each firm category and for each year-quarter, I compute the median of the service area's median household income and average green car penetration at the Zip-code level, weighted by the number of transactions. These two indicators reflect the wealthiness and greenness of the consumers that each firm category has served in each period. As illustrated in Figure 9, the top firms served the most affluent and green areas whereas medium-sized and small sellers served the areas that are much less affluent and green, although over time this gap has narrowed.[23]

I have shown evidence that is in support of the market segmentation hypothesis due to the two motivations of consumers. What are the implications of market segmentation when evaluating the subsidy effects? In the next section, I discuss the empirical strategy to estimate the overall effects of the subsidy program as well as separate subsidy effects in each segment. After that, I will discuss the empirical results and their implications.

6 Empirical Strategy

This section presents the empirical strategy to estimate the overall subsidy effects as well as the effects in each consumer-supplier segment. The subsidy effects are two-fold. First, as has been discussed above, the more subsidy goes to consumers, the more effective the subsidy program is in terms of stimulating volume. Second, for those subsidies that finally have gone to suppliers, we would like them to be captured by suppliers that have made efforts in improving product quality, so that the subsidy program would have provided the right incentives for firms to invest in R&D. I approximate the first effect by estimating the subsidy pass-through rate and examine the latter effect by constructing a quality metric of the solar system.

[22] In literature, there is abundant evidence that earning a higher income is a strong indicator of the green activists (e.g., see Paco and Raposo, 2010).

[23] This convergence may imply the exhaustion of the environmentally focused consumers.

6.1 Subsidy Pass-through

To quantify the share of benefits customers and sellers get from the California Solar Initiative (CSI), and to examine to what extent this program has helped to boost solar adoption, I illustrate that the subsidy pass-through rate is a sufficient statistic and can achieve the two purposes.

Figure 6 shows three scenarios that illustrate how subsidy pass-through could reflect the effectiveness of a subsidy program. As shown in sub-plot (a), if the supply curve is flat, the subsidy achieves a maximum boost in solar adoption from Q* to Q** without any increase in solar price. Whereas if the supply curve is very steep in sub-plot (b), the subsidy only results in higher solar price but no increase in solar adoption.

In reality, the supply curve is more likely to be upward sloping so that the retail price ends up between $6000/kW and $7000/kW as illustrated in sub-plot (c), meaning the subsidy benefits are shared between retailers and consumers. To what extent the retailers can increase the sales price relative to subsidy indicates the distributional effects of the subsidy. But it is an equilibrium outcome depending on the nature of the market, in particular the relationship between the price elasticity of retailers (the shape of the supply curve) and the price elasticity of consumers (the shape of the demand curve).

Therefore, who benefits how much from the solar incentive program and how much the subsidies have assisted in boosting solar adoption are two empirical questions and have to be answered using real world data. In fact, the answer can be summarized as a single sufficient statistic, called subsidy pass-through, and can be derived from the following equation. As we have seen above, with higher subsidy pass-through, the more share of benefits goes to retailers and the less effective the subsidy program is in terms of increasing solar adoption.

$$\text{subsidy passthrough rate} = \frac{\text{equilibrium price with subsidy - equilibrium price without subsidy}}{\text{subsidy level}}$$

If we are able to observe the market equilibrium prices with and without subsidy at the same time, then the subsidy pass-through rate can be computed from the data. However, it is impossible for one to observe both of the equilibrium prices at the same time, as one can only see the result of one world whereas the price in the counterfactual world is unobserved. Next, I show that this obstacle can be overcome through a careful empirical study on the particular subsidy scheme in California.

Baseline Model

Thanks to the independent design of the subsidy scheme across utilities as shown in Section 3.3, I measure the subsidy pass-through rate by taking advantage of variation in the rebate across utility

territories and time. Consider the following regression equation

$$P_{ikt} = \beta_0 + \beta_1 \cdot R_{(i)kt} + \text{FE}_k + \text{FE}_t + X_{ikt}\Gamma + \varepsilon_{ikt} \tag{2}$$

where P_{ikt} is the pre-rebate unit transaction price (dollar per Watt[24]) of the solar system i sold at date t in utility territory k. $R_{(i)kt}$ is the rebate level (dollar per Watt[25]). FE_k and FE_t are utility fixed effects and date (or year-month) fixed effects, respectively. X_{ikt} are control variables that include solar system characteristics including the brand and model of the physical components and consumer demographics at the zip-code level including median household income, median home values and education levels. ε_{ikt} is the error term. The coefficient on rebate level (β_1) is the parameter of interest and indicates the subsidy pass-through rate. This number should be in the range of 0 to 1 if the market is completely competitive. The larger it is, the more subsidy goes to solar sellers and the less effective the subsidy program is in stimulating adoption. For instance, if β_1 is estimated to be 1. This implies that although consumers get rebates from the subsidy program, they have to concede all the benefits back to the suppliers by paying a higher price which is exactly the original market equilibrium price plus the subsidy. In this case, consumers buy as if there is no subsidy and the subsidy program doesn't increase demand.

To interpret β_1 as the subsidy pass-through, it requires that variations in the rebate level across utility territories are exogenous. Detailed discussion of the identification of β_1 is illustrated in Appendix C.

6.2 Quality Metric

As discussed in Section 3, the ratio of the CEC-AC rating and STC-DC rating can be used as a performance metric of the solar system as a whole. This ratio, defined below, can be thought of as the how the solar system really performs given its nameplate size.

$$\text{performance metric} = \frac{\text{CEC-AC rating (CSI rating)}}{\text{STC-DC rating (Nameplate rating)}}$$

I find that this performance metric is an excellent predictor of solar system's future kWh output normalized by the upfront cost or by the nameplate system capacity. A one percent increase in performance metric is associated with at least 0.65 percent increase in the future energy output (see Appendix Table A.2). This performance metric I construct helps control for product quality when estimating subsidy pass-through rate.

[24] To be more specific, it is dollar per AC-rated Watt. Since AC-rating/DC-rating reflects the economic performance of the solar system, this unit price ($/AC-rated Watt) is actually "quality-adjusted" price.

[25] CSI rebate is paid based on the AC-rated capacity.

7 Results

This section presents the empirical results of the subsidy effects in the California solar market. I first discuss the results of the overall subsidy effects. Then I explore estimates for each market segment.

7.1 Overall Estimate

Table 3 reports an overall estimate of the subsidy pass-through based on Equation (2) using all the non-TPO sample. From column (1) to column (7), I gradually add more controls into the regression model. The dependent variable of the regression is unit transaction price (total transaction price divided by the system capacity in terms of the AC-rating). Column (1) is the baseline regression where I include only utility territory fixed effects and year-month fixed effects and do not control for component brands/models and consumer demographics. The overall estimate is about 0.8 and statistically significant, which suggests that about 80% of the subsidy benefits go to the suppliers and the program has had a limited effect in stimulating adoption. If replacing the year-month fixed effects with the date fixed effects or replacing the utility territory fixed effects with the county fixed effects, the estimate changes slightly from 0.8 to about 0.86 or 0.83, as shown in columns (2) and (3). However, when I add controls of components fixed effects and consumer demographics including median household income, median home value, and education levels at the zip-code level, the estimate drops from 0.8 to about 0.64–0.66 (columns (4) and (5)). This change suggests it is important to control for heterogeneity in consumer demographics and product characteristics. In order to consider all unobserved heterogeneity in consumers at the zip-code level, in columns (6) and (7), I replace county fixed effects with zip-code fixed effects and include year-month or date fixed effects. This does not change the estimate much. Overall around two thirds of the subsidy benefits have gone to suppliers, which implies that overall the subsidy program is not very effective in encouraging solar adoption.

7.2 Estimate in Each Market Segment

From the discussion in Section 3 and Section 5, it is likely that the subsidy has different effects in different segments. For example, environmentally focused consumers may be willing to concede more subsidy benefits to the suppliers (large firms), whereas economically focused consumers are likely to concede much less benefits to the suppliers (small and local sellers). Therefore, I test for this hypothesis by adding an interaction term between firm size category and the rebate level. The results are shown in Table 4. The estimated subsidy pass-through for top firms almost doubles

the pass-through for small and local firms. In fact, the estimate for top firms is around 90%, which means that top firms capture almost all the subsidy benefits if consumers buy from them. medium-sized firms capture around 55% benefits, which are also statistically higher than the 45% pass-through of the small firms.

Researchers might worry that firms of different sizes may also charge different price levels. To control for it, I add firm fixed effects into the model and this changes the overall estimate of subsidy pass-through dramatically. As shown in column (1) in Table 5, now the estimate is only about 0.36, which is much smaller than the previous estimate 0.66 in column (4) of Table 3, which is produced by the similar model but without firm fixed effects. This change may suggest that top firms may have redistributed transactions in areas that are of higher price level or higher rebate level and in order to do so they are willing to concede a bit of more subsidy benefits to the consumers. This change in estimate also suggests that firms indeed charge different prices. In fact, the between-segment price dispersion might be much higher than the within-segment price dispersion. Therefore, I relax the constraint that all firms (even if they are very different in scale or adverting strategies) face the same time fixed effects. Here the time fixed effects can be thought of as the average price set by firms net of the subsidy effects. It is highly possible that firms of different sizes set different prices over time. I re-do the regression but separately for each firm category so that each firm category could have a different set of time fixed effects.

Columns (2), (3) and (4) of Table 5 show results for the top firm category, medium-sized firm category and small firm category, respectively. According to this less restricted model, top firms enjoy almost 100% of the subsidy benefits, whereas medium-sized and small firms essentially concede the majority of the benefits to the consumers. These results are consistent with the findings above. I then plot out the time fixed effects for each firm category, as shown in Figure 10. There are two main takeaways from this figure. First, the average price set by small firms has been falling down very quickly. However, top firms were reluctant to reduce price especially in the early periods. There is a substantial price gap between top firms and the other firms, which may indicate the markups of top sellers. Second, by plotting each firm category's smoothed average price against the factory-gate price of solar modules, the lower panel of Figure 10 illustrates that the price set by small firms falls even faster than the price of solar modules, which may suggest that these firms may have successfully reduced the balance of system costs over the years.

To sum up the findings so far, on average approximately two thirds of the subsidy has gone to suppliers and this implies the subsidy has had a limited effect in stimulating adoption. By considering the fact that the market is segmented into the environmentally focused group (mostly large firms and "greener" consumers) and the economically focused group (mostly small and local firms and less "green" consumers), I find that economically focused suppliers are willing to concede more subsidy benefits to the economically focused consumers, which suggests that a lot of the

economically focused consumer would not buy if there was no subsidy. In this segment, the subsidy program is quite effective. However, environmentally focused consumers almost capture no benefits and they buy as if there was no subsidy program. For this segment, all subsidy benefits have gone to the pocket of the few top solar companies. And the subsidy program hasn't boosted adoption that much.

I now explore whether the high subsidy pass-through enjoyed by the top sellers is because they provide better quality products. To do this, I compute and compare the average product quality of top firms and other firms. As shown in Figure 15, top 1% sellers consistently sell solar systems that are poorer in terms of the performance metric compared with other sellers. In fact, in my regression analysis above, the dependent variable, unit price, is already quality adjusted. Therefore the high subsidy pass-through of top firms already reflects the fact that in this segment, consumers are more willing to buy a low quality product. This finding, again, supports the hypothesis that in this segment, consumers put less weight on the economic performance and put more weight on the environmental "warm-glow" attribute of the product.

Next, I want to show that the way I categorize the market into top-firms-consumers segment which is supposed to be environmentally focused and other-firms-consumers segment which is supposed to be economically focused is roughly right. The idea is that in the environmentally focused segment, the "greener" the areas are, the higher the subsidy pass-through should be. This is because all consumers in this segment are likely to be environmentally motivated but may be motivated at different levels. A hypothesis is that if a consumer lives in a "greener" area, it is likely that this consumer is more environmentally motivated. Whereas in the economically focused segment, what consumers care about is the economic performance rather than the "greenness" of the product. Therefore, where the consumers live should not affect the subsidy pass-through. A consumer could be very economically motivated even if he or she lives in an environmentally conscious area. Therefore, I add an interaction term between the "greenness" of consumers at the zip-code level and the subsidy level in the model. I use the percentage of electric or hybrid car adoption among all car registration as the proxy for the "greenness" metric. In columns (2-g), (3-g) and (4-g) of Table 5, it is shown that the coefficient on the interaction term is positive and statistically significant only for the "green" segment. For the other segments, it is not statistically different from 0. The estimate in column (2-g) implies that in the top-sellers-consumers segment, one standard deviation increase in the "greenness" of the area that consumers live would lead to an additional 6%–7% in subsidy benefits enjoyed by the firm.

All the regression analysis so far is based on "quality-adjusted" prices. The two effects of subsidy—price effect and quality effect—cannot be disentangled. One way to separately see the quality effect and the pure price effect is to re-normalize the price and subsidy by the name-plate rating. Consider the following estimating equation

$$
\begin{aligned}
P_{ikt} = \beta_0 &+ \sum_{q=1}^{4} \beta_1^q \cdot \text{perf\_metric\_q}_i \cdot R_{ikt} \\
&+ \sum_{q=1}^{4} \beta_2^q \cdot \text{perf\_metric\_q}_i \cdot R_{ikt} \cdot \text{green\_car\_pct}_{ik} \\
&+ K_k + T_t + X_i\Gamma + FE^{firm} + \varepsilon_{ikt}
\end{aligned}
$$

where q represents the quality quantile. I set 4 quantiles, where 0–25% represents lowest quality and 75%–100% represents the highest quality. The quantile of a solar system is based on its cohort (a rolling cohort of 100 solar systems ordered by date). Again, green_car_pct is the percentage of electric or hybrid car registrations in each zip-code location. Price and rebate are normalized by nameplate DC rating.

Results are illustrated in Table 10. We can see that even when controlling for product quality, the estimated subsidy pass-through enjoyed by the top firms is still much larger than the one captured by the other smaller suppliers. I next add an interaction term between the greenness of consumers and rebate level into the model. Table 11 shows that if a consumer buys a very low quality solar system from a top seller, then there is statistical evidence that the more environmentally motivated the consumer is, the more subsidy the consumer concedes to the suppliers (one standard deviation increase in the "greenness" of the area that consumers live would lead to an additional 5% in subsidy benefits enjoyed by the firm).

7.3 Robustness Checks

A threat to the baseline analysis is that due to the discrete subsidy scheme, purchases are bunching before the subsidy jumps. Researchers may concern that consumers before and after rebate changes are not necessarily the same in terms of price elasticity or bargaining power, which in turn affects the subsidy pass-through identification near the subsidy changes. One approach to tackle this concern is to re-do the baseline regression by dropping out observations around rebate changes to see whether the estimates are consistent.

A second concern is that there might be unobserved factors that have changed differently across regions over time. In other words, the common trend assumption might be violated if this is the case. For instance, consumers living in the SDG&E service territory might have become much more enthusiastic about solar power during the sample period. This would cause the rebate level to fall down more quickly in the San Diego area. In addition, it would also cause an increasingly relatively higher price level in San Diego. The estimate of the subsidy pass-through is then biased. The approach to deal with this issue is to exploit an interesting fact that due to historic reasons

households in the same city or zip-code area in Orange county may face different subsidy levels, so that a "cleaner" estimate is obtained by limiting the sample to these cities or zip-code areas where households are supposed to experience common trend except the rebate level.

7.4 Robustness Check I

As discussed in Section 4, in Figure 2, Figure 3b and Figure 3a, we observe that right before the rebate change, the number of transactions increases dramatically. For instance, in PG&E territory, about 85 transactions were made on average on the day right before each rebate jump, whereas on the day right after rebate jumps, in general only 40 transactions were made. More interestingly, temporary price reduction happened right before rebate jumps. Such quantity and price patterns suggest that consumers who made purchase before and after rebate changes are possibly different in terms of time preference, search cost, and/or price elasticity. As for sellers, they also could have different pricing strategies before and after rebate jumps.

To see whether consumers before and after rebate jumps are indeed different in terms of demographics, I compute the average demographics at the zip-code level for before and after consumer groups. Table 6 shows that consumers before subsidy changes tend to earn higher income, live in home of higher values, and hold a higher education degree. Statistically, the null hypothesis that household characteristics are the same before and after subsidy jumps is rejected by each of the individual t-test.

Therefore, I exclude observations around rebate jumps and re-do the regression analysis. As is shown in Table 7, the subsidy pass-through estimates do not change much compared with the baseline regression model, which makes the previous findings more credible.

7.5 Robustness Check II

My main strategy to estimate the subsidy pass-through is a difference-in-difference technique. Time fixed effects and region fixed effects altogether control for any factors that develop in parallel across regions. Ideally, we would like to verify if the common trend assumption is reasonable in this difference-in-difference framework. One way to do that is to see whether the pre-trends were developing in parallel. However, in this context, I do not observe enough observations for the pre-trend period. My strategy then is to limit the observations to those that are supposed to be very similar to each other so that any unobserved factors should have been developing in parallel. If similar estimate is produced by doing the regression analysis on this limited sample, then the baseline results should be more credible.

To be specific, I exploit subsidy variation at spatial discontinuities in electricity service areas, where households in the same city or even in the same zip-code area face substantially different

subsidy levels. Figure 11 shows the utility territory border in Orange County between SCE and SDG&E. Since the border lies within city limits, households in the same city are administrated by different companies to apply for the subsidy. The reason that the utility border lies within city limits is that SCE and SDG&E connected their transmission lines in this area and established the territory border as early as in the 1940s. Whereas the city limits in this area were not established until the 1980s (Ito, 2014). Because the subsidy program is divided and administrated by the three utility companies in California (for instance, Los Angeles residents are administrated by SCE. San Diego residents are administrated by SDG&E (CSE)), residents are rebate level taker and do not have the freedom to choose their subsidy administrator. Therefore, at the same point of time, households in the same city along the border line could experience substantially different subsidy levels. As can be seen in Figure 12, for all of the time, subsidy level in SCE was higher than the one in SDG&E if not equal to. For some periods in 2011, such difference was as high as over 1000 dollars per kW. This means that for an average size solar system, say 4–5 kWs, a SCE resident could save as much as 5000 dollars than his counterpart living in the same city but in the SDG&E territory. Consumers living along the border also have similar demographics (Figure 13).

One of the concerns to limit the sample on the border is that we may end up with a very small sample. I show that although it is a small sample, the sample is quite balanced across regions and over time. Table 8 shows the distribution of solar systems sold in the cities along the border of SCE and SDG&E. We can see that though the sample is small (less than 1000 observations), it is a good mix of both utility territories in most of the cities. In Appendix Figure A.7, it also shows that the sample is quite balanced over time. The following analysis in this subsection is on this limited sample along the border between SCE and SDG&E.

Before conducting the regression analysis, I first provide graphical evidence that solar systems sold in SCE territory are relatively more expensive compared to SDG&E systems when the relative SCE subsidy level is high, even controlling for system size. In Figure 14, I show how SCE's average solar system capacity and average system total price evolved after normalized by base year 2008 relative to SDG&E. In particular, the green line shows the difference-in-differences (DD) in the mean of the system capacity for SCE customers relative to SDG&E customers. For each year, I calculate the average solar system capacity in each year for both service areas. Then, I normalize the annual average capacity by subtracting the average capacity of 2008 from the annual averages for each of the following years. Finally, I obtain the DD by subtracting SDG&E's annual averages from SCE's annual averages. The DD estimate thus shows how SCE's solar system capacity evolved since 2008 relative to SDG&E. I call it the relative change in system capacity. In the same manner, I compute the DD in the means of the total system prices, as is indicated in red line in Figure 14. Essentially, Figure 14 shows that compared to base year 2008 and compared to SDG&E, (1) the average system size sold in SCE was relatively smaller in the four consecutive

years following 2008 and then became relatively larger in 2013 and 2014; and (2) the average total price of a system in SCE was in general relatively higher in all years following 2008 but in 2014. This pattern implies that during 2009–2012, solar systems in SCE were relatively more expensive than SDG&E especially in year 2011, which is also the period when the subsidy level difference was large. To conclude, Figure 12 and Figure 14 illustrate that SCE's subsidy level relative to SDG&E is positively correlated with SCE's price level relative to SDG&E.

For rigorous statistical analysis, I next run the following regression on this limited sample

$$P_{ikt} = \beta_0 + \beta_1 \cdot R_{(i)kt} + FEs + X_i\Gamma + \varepsilon_{ikt}$$

where $R_{(i)kt}$ is again rebate level; FEs are fixed effects dummies including service utility fixed effects, year-month fixed effects, and solar firm fixed effects; To control for solar system heterogeneity and consumer heterogeneity, X_i include fixed effects of system component brands and controls for consumer demographics at the zip-code level. I also control for residential electricity prices by adding annual Tier 5 electricity price or average electricity price faced by a single-family home household. Table 9 illustrates the regression results. The model specifications of columns (1) and (2) do not include controls for electricity price. The coefficient on rebate in model (1) shows that on average about 45% of subsidy goes to sellers. When I add the interaction term between rebate and an indicator that equals one for top firms, as shown in column (2), the estimates imply that top firms capture the full subsidy benefits whereas other firms concede all benefits to consumers, which is similar to what I find in the baseline model. These findings are very robust when I control for Tier 5 electricity price or average electricity price, as shown in column (3)–column (6). This robustness check exercise suggests that it is unlikely that the common trend assumption is violated in this context.

8 Discussion

Why do top firms capture so much of the subsidy and yet provide products of lower quality? I have shown in Section 5 that consumers living in high income or high green car penetration areas (which are established indicators of the greenness of consumers) are more likely to buy from top companies (which advertise heavily and focus on advertising the green attribute of the product). On the contrary, small and local solar companies are more focused on local markets and have little capability to advertise. They serve consumers that are more economically focused and less environmentally focused. The market segmentation causes dramatically different subsidy effects among different consumer-supplier groups, as illustrated in Equation (1). I show empirically that the equilibrium market offering (p^*, g^*, e^*) are indeed different in the two segments.

This is an unintended effect, given that the intention of the subsidy program is to reduce the price that consumers pay and stimulate solar adoption.[26] In the environmentally focused segment, consumers buy a solar system as if there were no subsidy benefits. These consumers are also those who are likely to eventually adopt the solar system anyway. In this sense, the subsidy program fails to achieve its purpose for these consumers.

Quantifying the Unintended Effect

To quantify this unintended effect, I conduct a back-of-the-envelope calculation based on the estimated subsidy pass-through rate for each consumer-supply segment.

First, I calculate the subsidy incidence. I assume that the subsidy pass-through rates for the third-party-owned systems are the same as the rates for customer-owned systems.[27] Given the number of projects, the average system capacity, the average rebate paid per capacity, and the estimated subsidy pass-through rate for all the three segments $j \in \{L, M, S\}$ (see Table 12), the subsidy incidence captured by firms can be expressed as

$$
\begin{aligned}
&\frac{\text{Subsidies captured by firms}}{\text{Total subsidies}} \\
=&\frac{\sum_{j\in\{L,M,S\}} \text{Num of projects}^j \times \text{Avg capacity}^j \times \text{Avg rebate}^j \times \widehat{\text{Subsidy pass-through}}^j}{\sum_{j\in\{L,M,S\}} \text{Num of projects}^j \times \text{Avg capacity}^j \times \text{Avg rebate}^j} \\
=&\frac{72,325 \times 5.61 \times 0.60 \times 1 + 44,491 \times 6.08 \times 0.76 \times 0.046 + 17,091 \times 6.25 \times 0.76 \times 0.188}{72,325 \times 5.61 \times 0.60 + 44,491 \times 6.08 \times 0.76 + 17,091 \times 6.25 \times 0.76} \\
=&\frac{\$243.446\text{M} + \$9.457\text{M} + \$15.262\text{M}}{\$243.446\text{M} + \$205.584\text{M} + \$81.182\text{M}} = \frac{\$268.165\text{M}}{\$530.212\text{M}} \approx 50.6\%.
\end{aligned}
$$

Over 50% of subsidies have gone to firms. This is due to the fact that the green consumers in the California market represent about $\frac{72,325}{72,325+44,491+17,091} \approx 54\%$ of the market. The following calculation illustrates that over 90% of the subsidy benefits captured by firms are captured by the top firms.

[26] The learning-by-doing effect is another concern. For example, the more solar systems are installed, the more quickly the cost of solar would reduce. For this purpose, it is still ideal if the subsidy program can stimulate volume as large as possible.

[27] For TPO solar systems, subsidy is paid directly to the suppliers instead of the buyers. There is empirical evidence (in automobile industry, Busse et al., 2006) that the direct subsidy receiver tend to have information advantage and capture a larger proportion of the subsidy benefits than the case that he or she is not the direct subsidy receiver. Therefore, this assumption may underestimate the subsidies captured by sellers.

$$\frac{\text{Subsidies captured by top 1\% firms}}{\text{Subsidies captured by firms}}$$
$$= \frac{\$243.446\text{M}}{\$243.446\text{M} + \$9.457\text{M} + \$15.262\text{M}} = \frac{\$243.446\text{M}}{\$268.165\text{M}} \approx 90.8\%$$

Next, I compare the quality of a typical solar system sold by a smaller firm with the quality of the one sold by a top firm. I find that a typical solar system sold by a smaller supplier is likely to produce 2908 kWh more electricity over a system lifetime of 30 years. The assumptions of this calculation are as follow. The average performance metric of solar systems sold by large suppliers and medium/small suppliers are 0.8097 and 0.8225, respectively. The marginal effect of performance metric on daily electricity output normalized by DC-rated system capacity is 3.458. Then the average performance metric difference between large and smaller suppliers implies a difference in the total energy output of a 6 DC-rated kW solar system, which is expected to work for 30 years, to be[28]

$$(0.8225 - 0.8097) \times 3.458 \times 6 \times 365 \times 30 \approx 2908 \text{ kWh}.$$

To sum up, the high percentage of green consumers essentially transforms a large portion of the solar incentives that is over 50% of the total subsidy into rents enjoyed by the few large suppliers that provide lower quality products.

Alternative Explanations

Warranty

The price of a solar system includes the warranty. If large firms provide better or longer-lasting warranty choices than smaller firms, this may also explain some of the findings above. However, in California it is the industry standard for all solar systems to be sold with similar warranties.

A solar system is a durable good and normally has two warranties: the performance warranty and the equipment warranty. A solar panel's performance warranty typically guarantees the production of the solar system over time. The common practice in the solar energy industry is that solar companies generally guarantee that their panels will produce electricity at 80%–90% of their power output rating at the end of 25 years. A solar panel's product warranty covers the integrity of the panel itself, and protects the consumers against problems such as manufacturing defects, environmental issues, premature wear and tear etc. Most solar companies offer a 10-year product

[28] I assume here that the efficiency of a solar system does not change over time.

warranty from the date of installation. Therefore, there is little variation in warranties.[29]

Branding Effect

I have argued that consumers who buy from a top company are very likely to be motivated by the "environmental" nature of the solar system. However, it is possible that a consumer who buys from a top company may simply choose that firm because of the brand.[30] My argument against this alternative explanation is that if "greenness" is an important feature of the branding effect, then essentially consumers are still largely motivated by the environmental "warm glow". For example, consumers put value on the brand "Tesla" because they perceive that Tesla represents a company that advances the green technology. In other words, if Tesla produces conventional vehicles or conventional power generation systems, it is harder for it to package itself into a brand valued by consumers. Therefore, in the context of this paper, I argue that the branding effect, if consumers do put value on certain brands, is still a result of the environmental "warm glow" effect.

9 Conclusion

Using micro-level data, this paper documents a previously unrecognized but important consequence of green technology subsidies. The intended result of a green subsidy is normally to reduce the price consumers pay and to stimulate adoption of the green good. However, in the California solar market, I illustrate that over 50% of the subsidy benefits have transferred into firm rents, and most are enjoyed by a few large solar suppliers. In addition, environmentally focused consumers purchase from the top solar sellers whose solar systems are of lower quality. I calculate that compared to the quality of small company systems, the large companies' systems on average produce 2908 kWh less electricity over their lifetimes. This unintended result is due to the fact that there is a significant proportion of consumers who primarily pursue the environmental "warm glow" that comes with the green good. These consumers may also care about the economic performance of the product. However, since it is a new product with uncertain future outputs and it is a very infrequent purchase decision made by most consumers, it is costly for the general buyers to identify the product quality. The product quality is so hard to observe or so noisy that it is likely that consumers have to put less weight on it and relatively more weight on the environmental attribute in

[29] Another concern is that consumers may pay a premium to large firms because they believe that large firms are unlikely to go bankrupt. This logic does not fit the California market. SolarCity was the largest solar supplier in California. However, it went bankrupt in 2016, although it was later acquired by Tesla. On the other hand, if consumers buy from a local contractor, that contractor is more likely than a leading supplier to still be around in the neighborhood for maintenance service provision for a longer time period.

[30] For example, consumers buy Tesla solar panels may simply because they want to buy a Tesla product.

their purchase decision. Therefore, the steep learning curve and high search cost for consumers in this market is another reason behind this undesirable consequence of green subsidies.

A green good is subsidized because of its green nature. On the other hand, the undesirable consequence of subsidy as discussed in this paper also stems from its green nature. Therefore, it is difficult to fully eliminate this undesirable result. However, mitigation is possible. To mitigate this unintended consequence, efforts can be made by policy makers to educate consumers and make it easier for them to learn and search in this market. For instance, it may be desirable to standardize the quality metric for solar systems so that it is less costly for consumers to differentiate between good quality and bad quality systems. Another way to reduce consumer search cost is to build bidding platforms to assist consumers compare quotas and to encourage firm competition. Third, it may also be desirable to share some of the administrative data with potential buyers on the local solar market. For example, it is easier for consumers to find the most cost-effective product if consumers know the cost, energy outputs, and seller of the solar systems in their neighborhood. All these actions can effectively make the product quality less noisy especially for the environmentally focused consumers so that they can make a more informed purchase decision that puts more considerations on the product economic performance and the economic returns. It is likely that this change in consumers' demand function can incentivize firms to provide better quality products and increase the subsidy benefits consumers can capture.

The conceptual analysis framework of this paper applies to all green goods. Therefore, the implications of the findings in this paper can possibly be extended to subsidies for other green goods or other products that have the multi-attributes property. Examples include subsidies for electric or hybrid vehicles; rebates for Energy Star home appliances; and incentives for healthy food, such as organic or non-GMO products. Further empirical work in these settings will help gauge the generality of the implications of this paper. However, this paper highlights the concerns about market segmentation resulting from the environmental motivations of consumers in evaluating green subsidy programs, a topic that, to the best of my knowledge, has not been documented previously. Continuing to evaluate and address this concern discussed in this paper can be an important component of future environmental policy-making.

References

(2013). Solar photovoltaic cell/module shipments report. Technical report, US Energy Information Administration.

(2016). *California Solar Initiative Program Handbook*.

Aguirre, I., Cowan, S., and Vickers, J. (2010). Monopoly price discrimination and demand curvature. *American Economic Review*, 100(4).

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115.

Allen, J., Clark, R., and Houde, J. F. (2014). Price dispersion in mortgage markets. *The Journal of Industrial Economics*, 62(3).

Bollingery, B. and Gillingham, K. (2014). Learning-by-doing in solar photovoltaic installations. *Working paper*.

Borenstein, S. (2017). Private net benefits of residential solar pv: the role of electricity tariffs, tax incentives, and rebates. *Journal of the Association of Environmental and Resource Economists*, 4(S1).

Busse, M., Silva-Risso, J., and Zettelmeyer, F. (2006). $1,000 cash back: The pass-through of auto manufacturer promotions. *American Economic Review*, 96(4).

Cabral, M. (2017). Claim timing and ex post adverse selection. *The Review of Economic Studies*, 84(1).

Cachon, G. P. and Swinney, R. (2009). Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science*, 55(3).

Chandra, A., Gulati, S., and Kandlikar, M. (2010). Green drivers or free riders? an analysis of tax rebates for hybrid vehicles. *Journal of Environmental Economics and Management*, 60(2).

Cowan, S. (2012). Third-degree price discrimination and consumer surplus. *The Journal of Industrial Economics*, 60(2).

Diamond, D. (2009). The impact of government incentives for hybrid-electric vehicles: Evidence from us states. *Energy Policy*, 37(3).

Dickson, P. and Ginter, J. (1987). Market segmentation, product differentiation, and marketing strategy. *Journal of Marketing*, 51(2).

Dong, C., Wiser, R., and Rai, V. (2018). Incentive pass-through for residential solar systems in california. *Energy Economics*, 72.

Dorsey, J. (2017). Harnessing the power of the sun through online platforms: evidence from the rooftop solar market.

Esteban, S. and Shum, M. (2007). Durable-goods oligopoly with secondary markets: the case of automobiles. *RAND Journal of Economics*, 38(2).

Fabra, N. and Reguant, M. (2014). Pass-through of emissions costs in electricity markets. *American Economic Review*, 104(9).

Gallagher, K. S. and Muehlegger, E. (2011). Giving green to get green? incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, 61(1).

Gillingham, K., Deng, H., Wiser, R. H., Darghouth, N., Nemet, G., Barbose, G. L., Rai, V., and Dong, C. (2014). Deconstructing solar photovoltaic pricing: the role of market structure, technology and policy. *Lawrence Berkeley National Lab Working Paper*.

Goldberg, P. K. and Hellerstein, R. (2008). A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market. *American Economic Review: Papers & Proceedings*, 98(2).

Gowrisankaran, G. and Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, 120(6).

Griffith, R., Leibtag, E., Leicester, A., and Nevo, A. (2009). Consumer shopping behavior: how much do consumers save? *Journal of Economic Perspectives*, 23(2).

Hendel, I. and Nevo, A. (2006a). Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74(6).

Hendel, I. and Nevo, A. (2006b). Sales and consumer inventory. *RAND Journal of Economics*, 37(3).

Hendel, I. and Nevo, A. (2013). Intertemporal price discrimination in storable goods markets. *American Economic Review*, 103(7).

Henwood, K. (2014). Subsidy pass-through in residential solar markets. *Working paper*.

Ito, K. (2014). Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(2).

Kahn, M. E. (2007). Do greens drive hummers or hybrids? environmental ideology as a determinant of consumer choice. *Journal of Environmental Economics and Management*, 54(2).

Kahn, M. E. and Matsusaka, J. G. (1997). Demand for environmental goods: Evidence from voting patterns on california initiatives. *The Journal of Law & Economics*, 40(1).

Mierendorff, K. (2016). Optimal dynamic mechanism design with deadlines. *Journal of Economic Theory*, 161.

Nair, H. (2007). Intertemporal price discrimination with forward-looking consumers: Application to the us market for console video-games. *Quantitative Marketing and Economics*, 5(3).

Nemet, G., O'Shaughnessy, E., Wiser, R. H., Darghouth, N., Barbose, G. L., Gillingham, K., and Rai, V. (2017a). Sources of price dispersion in u.s. residential solar installations. *Berkeley Lab*.

Nemet, G. F., O'Shaughnessy, E., Wiser, R., Darghouth, N., Barbose, G., Gillingham, K., and Rai, V. (2017b). Characteristics of low-priced solar pv systems in the u.s. *Applied Energy*, 187.

O'Shaughnessy, E. and Margolis, R. (2017). Using residential solar pv quote data to analyze the relationship between installer pricing and firm size. *National Renewable Energy Laboratory*.

Paco, A. and Raposo, M. (2010). Green consumer market segmentation: empirical findingsfrom portugal. *International Journal of Consumer Studies*, 34(4).

Pashigian, B. P. and Bowen, B. (1991). Why are products sold on sale?: explanations of pricing regularities. *The Quarterly Journal of Economics*, 106(4).

Pless, J. and van Benthem, A. A. (2017). The surprising pass-through of solar subsidies. *NEBR Working Paper*.

Sallee, J. M. (2011). The surprising incidence of tax credits for the toyota prius. *American Economic Journal: Economic Policy*, 3(2).

Schmalensee, R. (1981). Output and welfare implications of monopolistic third-degree price discrimination. *American Economic Review*, 71(1).

Shepard, A. (1991). Price discrimination and retail configuration. *Journal of Political Economy*, 99(1).

Shimomura, K.-I. and Thisse, J.-F. (2012). Competition among the big and the small. *RAND Journal of Economics*, 43(2).

Sorensen, A. T. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, 108(4).

Sterner, T. and Coria, J. (2003). *Policy Instruments for Environmental and Natural Resource Management*.

Stokey, N. L. (1979). Intertemporal price discrimination. *The Quarterly Journal of Economics*, 93(3).

Wang, E. Y. (2015). The impact of soda taxes on consumer welfare: implications of storability and taste heterogeneity. *RAND Journal of Economics*, 46(2).

Warner, E. J. and Barsky, R. B. (1995). The timing and magnitude of retail store markdowns: evidence from weekends and holidays. *The Quarterly Journal of Economics*, 110(2).

Weyl, E. G. and Fabinger, M. (2013). Pass-through as an economic tool: principles of incidence under imperfect competition. *Journal of Political Economy*, 121(3).

Williams, K. R. (2017). Dynamic airline pricing and seat availability. *Cowles Foundation Discussion paper No. 3003*.

Woodward, S. E. and Hall, R. E. (2012). Diagnosing consumer confusion and sub-optimal shopping effort: Theory and mortgage-market evidence. *American Economic Review*, 102(7).

Zhang, L. and Wu, Y. (2012). Market segmentation and willingness to pay for green electricity among urban residents in china: The case of jiangsu province. *Energy Policy*, 51.

Figures and Tables

Figure 1: Realized Rebate Levels

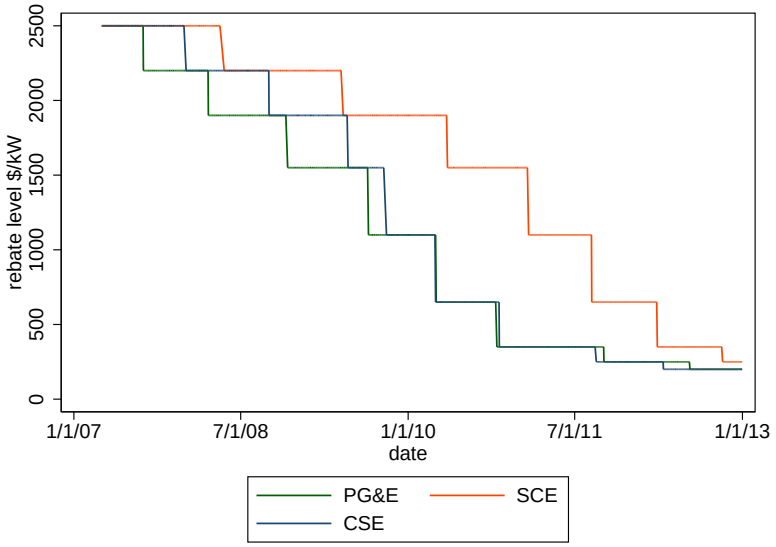
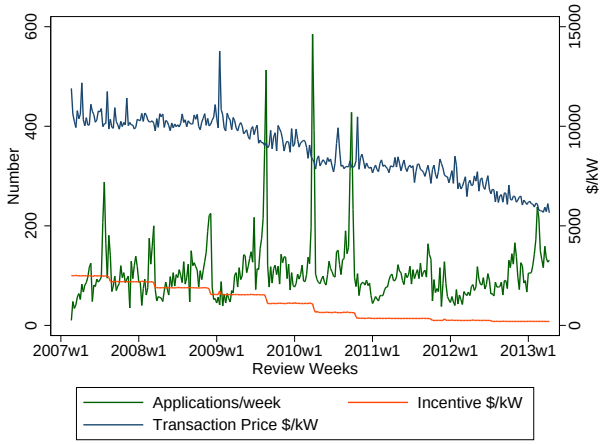
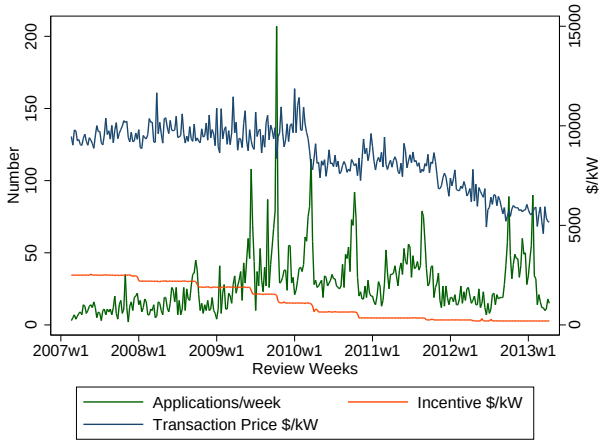


Figure 2: Application Dynamics in PG&E, CSE and SCE

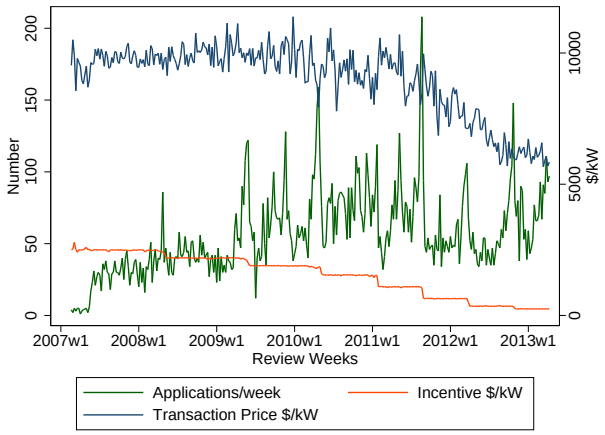
(a) PG&E



(b) CSE



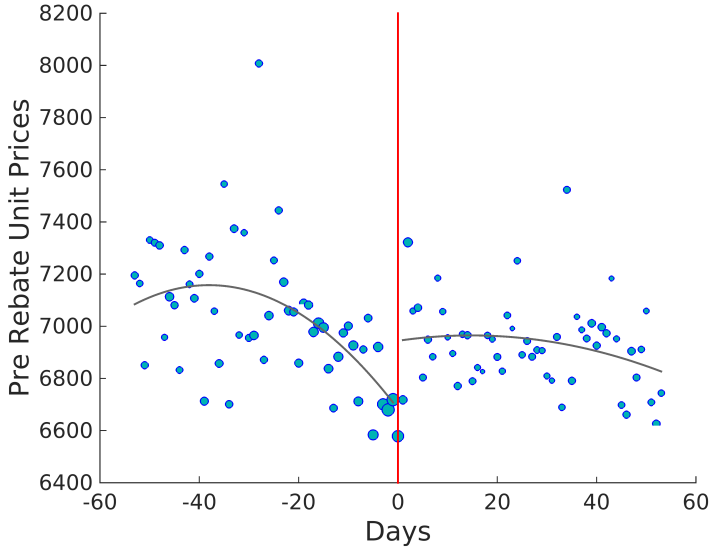
(c) SCE



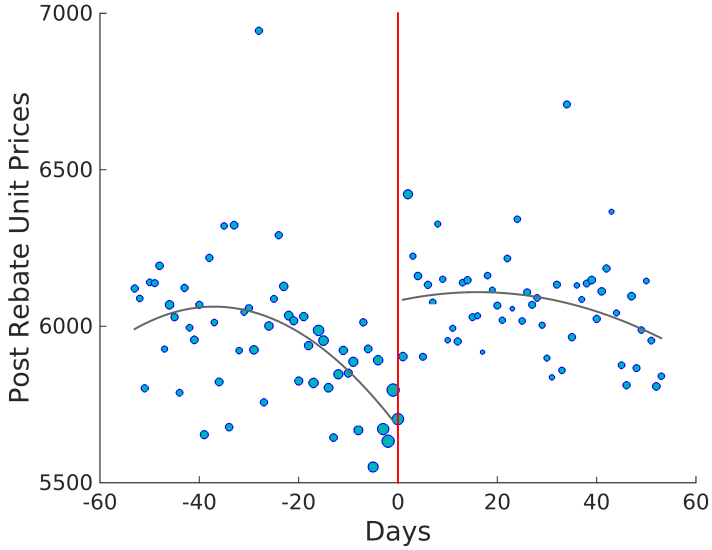
Remarks: Data is collapsed by week. All time series start at week 8 in 2007 and end at week 15 in 2013. Sample is limited to non-3rd-party owners, since price of 3rd-party system may not reflect real transaction price.

Figure 3: Transaction Prices and Quantities Before and After Rebate Changes in PG&E

(a) Pre-rebate Price



(b) Post-rebate Price



Remarks: For (a), the dots are computed in the following way. First, compute the unit price, i.e., price per AC-rated kW, for each solar system. Second, for each rebate change, demean the unit prices by the average price of 40 days around the rebate change date. Third, for each day, take the average of the demeaned prices by days before and after the change. For (b), dots are computed similarly as in (a), but are post-rebate prices (unit price-rebate level). The size of a dot is proportional to the number of transactions.

Figure 4: Overall HHI by Year-Month

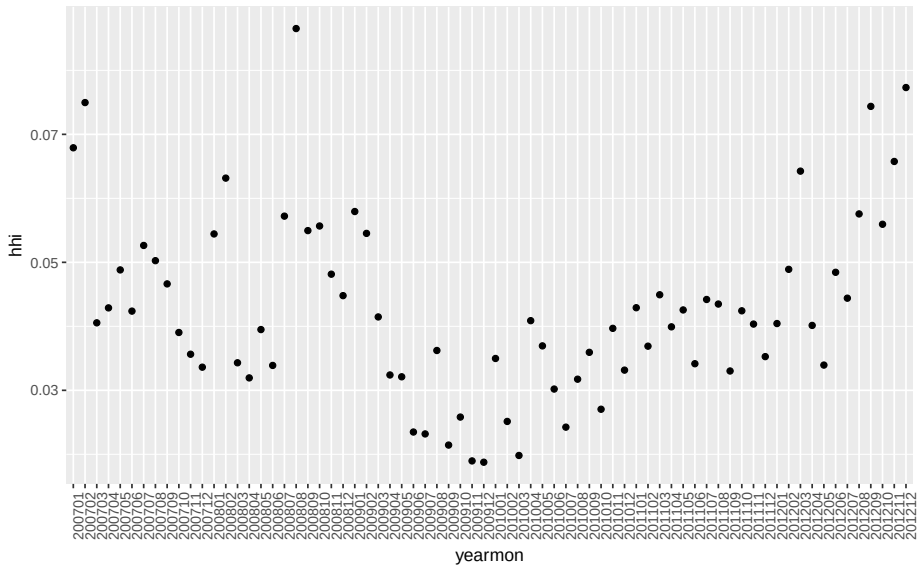
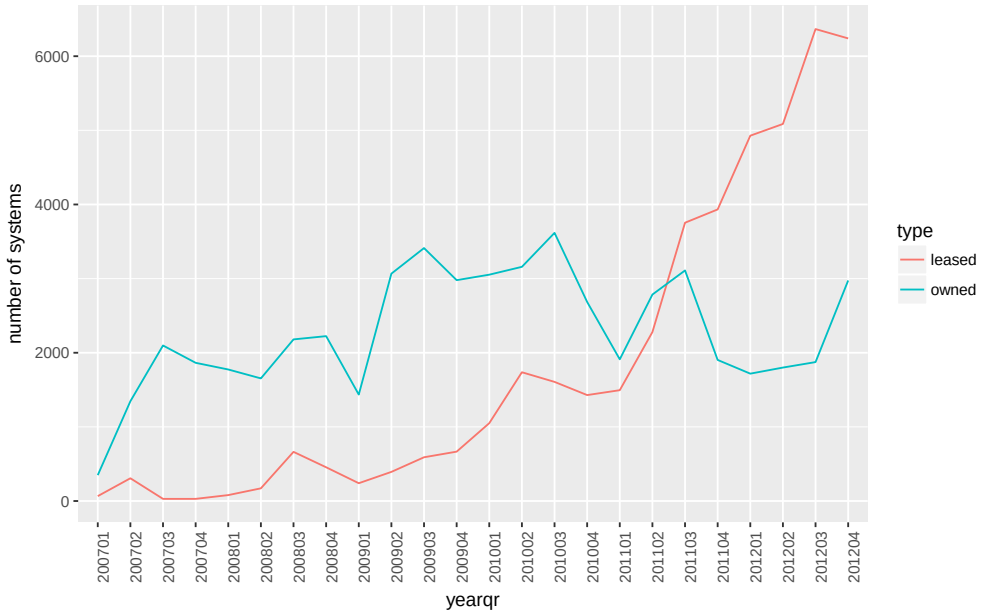
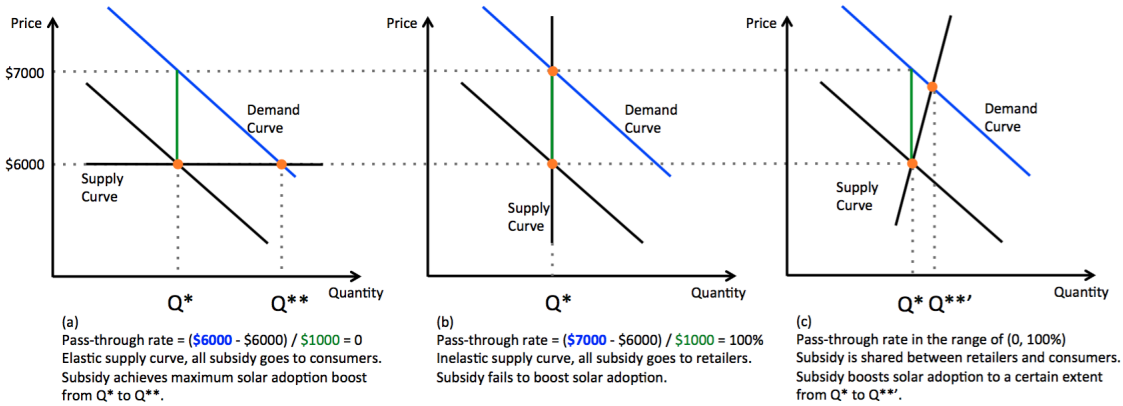


Figure 5: Non-TPO Sales versus TPO Sales



Remarks:
This chart shows the sales of consumer-owned solar systems (in green) against the sales of TPO solar systems (in red) over time.

Figure 6: Three Scenarios

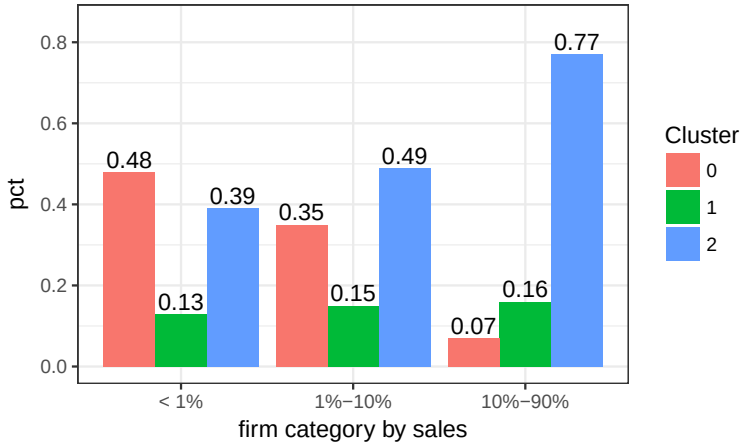


(a)
Pass-through rate = ($6000 - $6000) / $1000 = 0
Elastic supply curve, all subsidy goes to consumers.
Subsidy achieves maximum solar adoption boost from Q* to Q**.

(b)
Pass-through rate = ($7000 - $6000) / $1000 = 100%
Inelastic supply curve, all subsidy goes to retailers.
Subsidy fails to boost solar adoption.

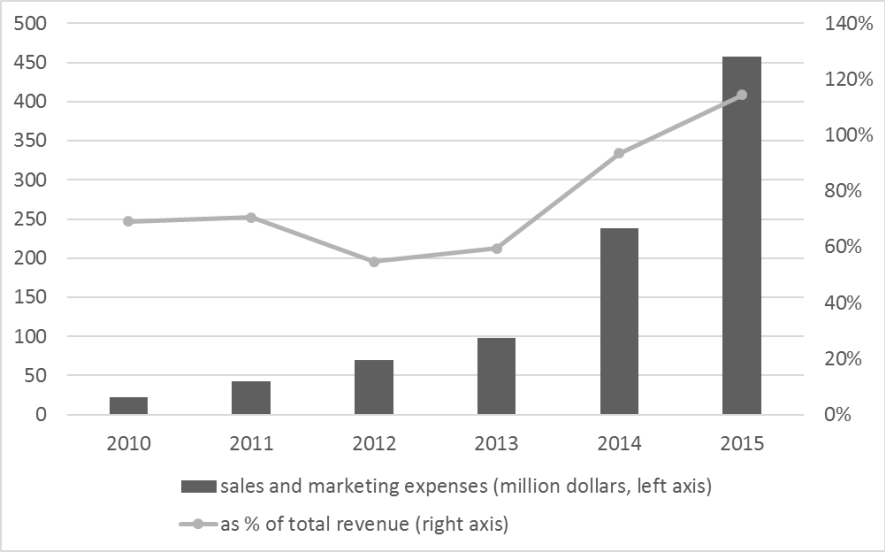
(c)
Pass-through rate in the range of (0, 100%)
Subsidy is shared between retailers and consumers.
Subsidy boosts solar adoption to a certain extent from Q* to Q**'.

Figure 7: Distribution of K-means Clusters by Firm Category



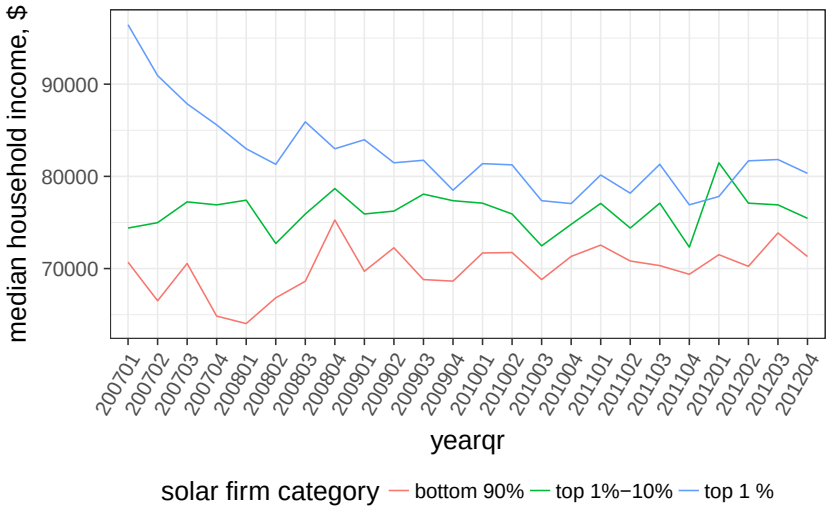
Remarks:
This chart shows the distribution of the k-means clusters in each firm category. Cluster 2 are all firms that do not have a website or I get access error when trying to scrape the website. Cluster 0 and cluster 1 are categorized by the k-means algorithm with two clusters on all the website texts.

Figure 8: Sales and Marketing Expenses of SolarCity

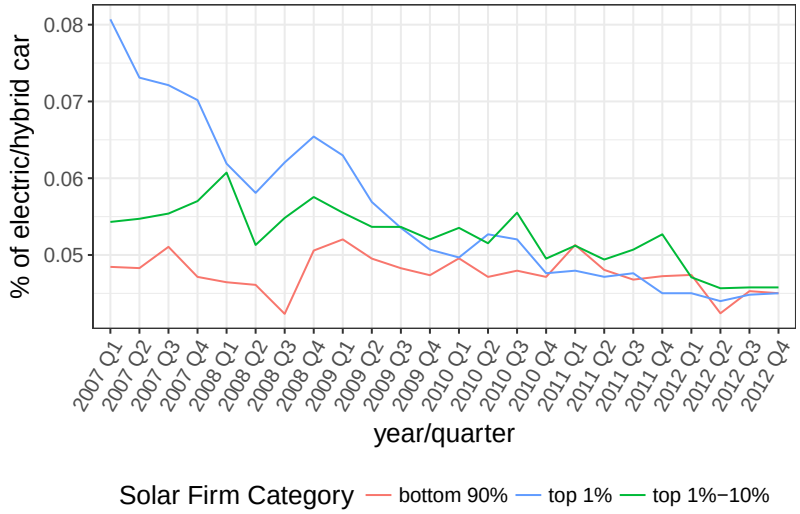


Source: SolarCity, 2013–2016, 10-K financial reports.

Figure 9: Evidence of Market Segmentation

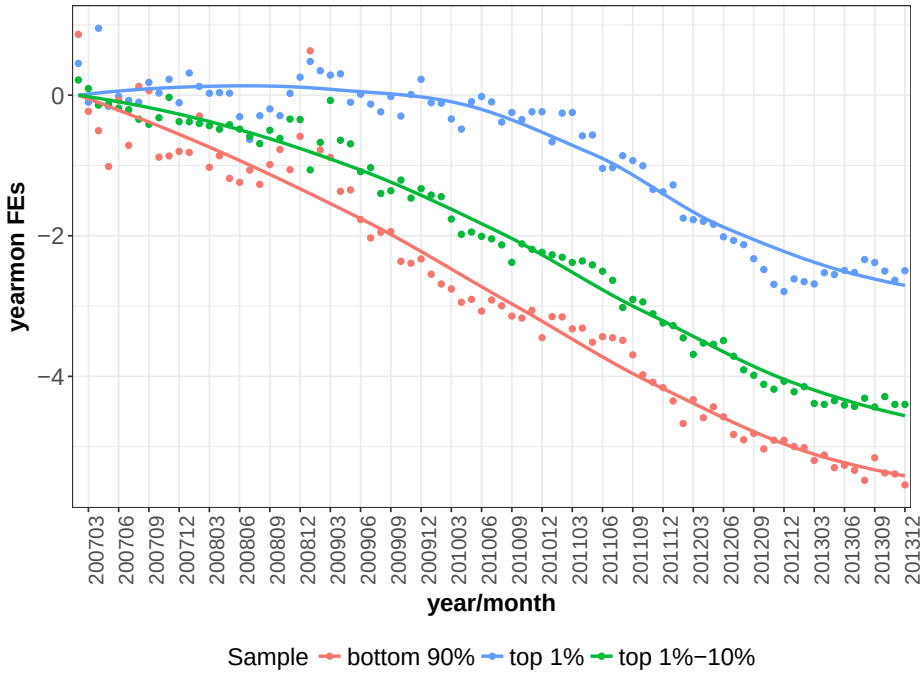


(a) Wealthiness of Buyers by Firm Category



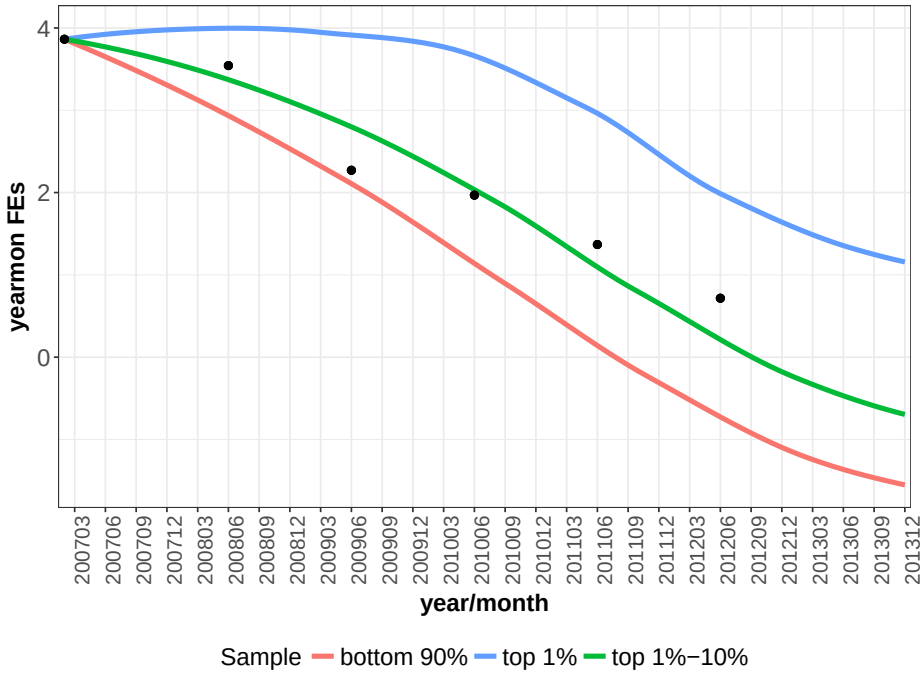
(b) Greenness of Buyers by Firm Category

Figure 10: Time FEs by Firm Category



(a) Time FEs by Firm Category

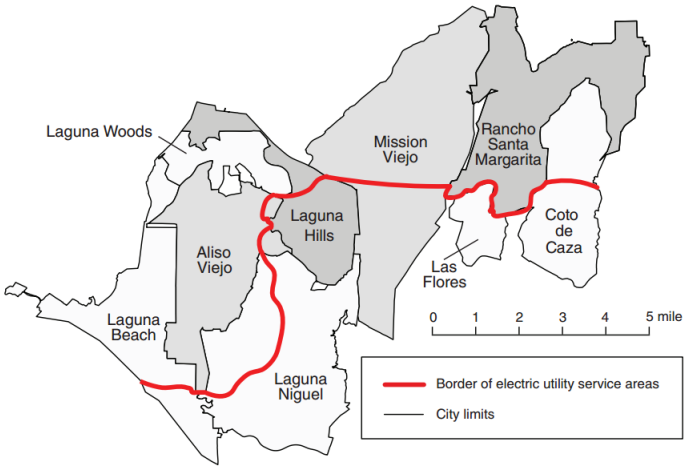
Remarks: First periods are aligned to be 0. Dots represent time fixed effect. Lines represent local polynomial regression fittings.



(b) Time FEs by Firm Category, against factory-gate prices of solar panels

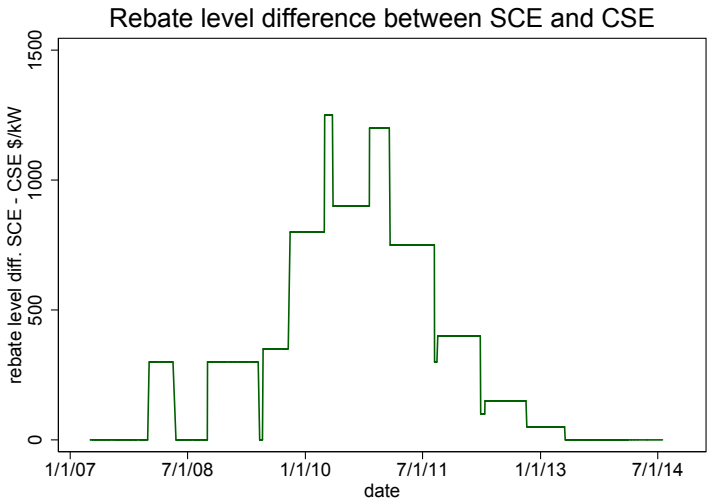
Remarks: Black dots represent average factory-gate price of solar modules by year. Data source: NPD Solarbuzz. First period of time fixed effects is adjusted (shifted downwards) to be aligned with the first period price of solar panels.

Figure 11: Border of Electricity Service Areas in Orange County, California



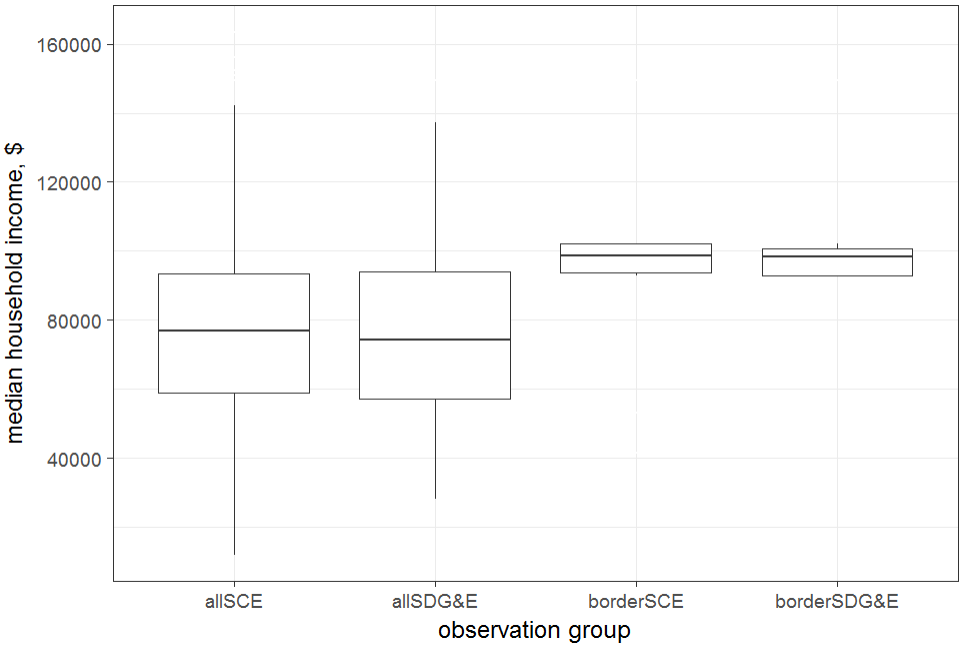
Remarks:
This figure shows the subsidy program service border between SCE (north of the border) and SDG&E (CSE, south of the border). In fact, this is also the utility service border between SCE and SDG&E. Source: Ito (2014).

Figure 12: Rebate Difference on the Border of SCE and SDG&E (SCE)



Remarks:
This figure shows subsidy level difference between SCE and SDG&E (CSE) over time. For all of the time, SCE subsidy level is higher than SDG&E (CSE) if not equal to, and such difference can be as high as over 1000 dollars per kW.

Figure 13: Demographic Comparison across the Service Border



Remarks:
This figure shows the range of median household income in the four regions: all SCE service territory, all SDG&E service territory, cities on the border but on the SCE side, cities on the border but on the SDG&E side. The black line is the range of median household income. The black box shows the 25%–75% percentile.

Figure 14: Difference-in-Differences in Capacity and Total Price

Figure 15: Performance Metric Gap between Top 1% Sellers and Other Sellers

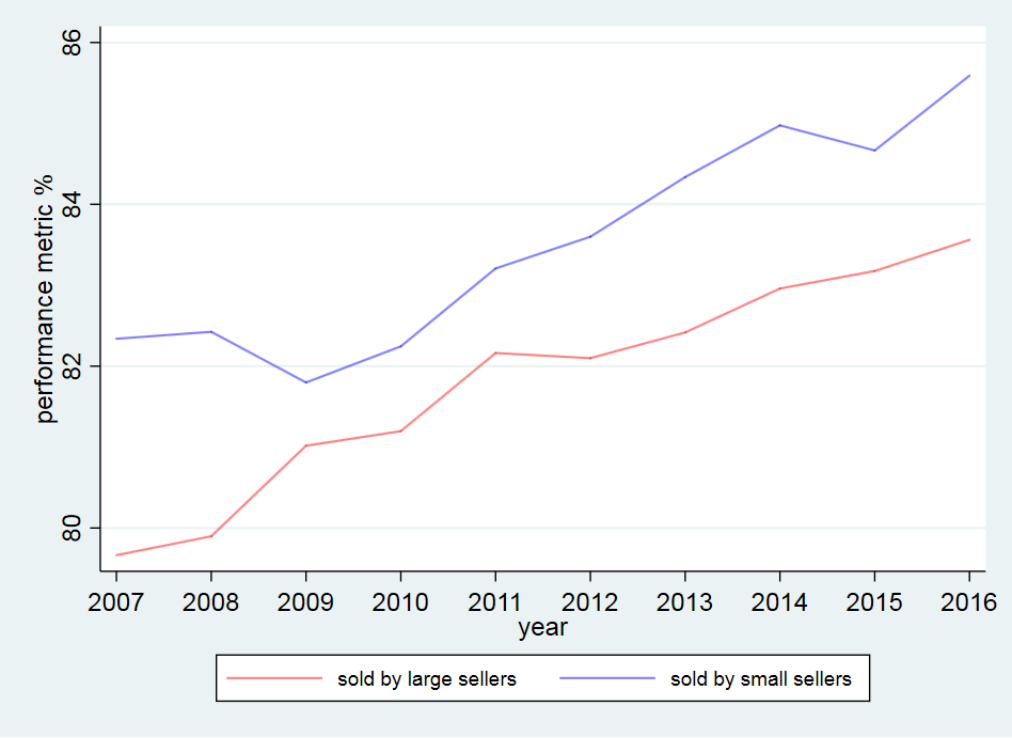


Table 1: Rebate Steps & Installed Capacity

Step	Rebate level L ($/W)	Capacity Trigger (MW)		
		PG&E	SCE	CSE
1	2.50	0	0	0
2	2.20	10.1	10.6	2.4
3	1.90	14.4	15.2	3.4
4	1.55	18.7	19.7	4.4
5	1.10	23.1	24.3	5.4
6	0.65	27.4	28.8	6.5
7	0.35	31.0	32.6	7.3
8	0.25	36.1	38.0	8.5
9	0.20	41.1	43.3	9.7
10	0.00	50.5	53.1	36.9

Table 2: Number of Systems Sold by Company Categories

Contractor Size	# of Companies	Mean	S.D.	Min	Max	Total (%)
Overall	2295	58.35	532.2	1	20071	133907
Top 1% companies	23	3144.6	4341.6	768	20071	72325 (54.1%)
Top 10% companies (excluding top 1%)	207	215.3	161.5	67	764	44558 (33.3%)
Top 50% companies (excluding top 10%)	918	16.7	15.6	3	67	15351 (11.5%)
Bottom 50% companies	1147	1.5	0.7	1	3	1673 (1.2%)

Remarks:
Data comes from California Solar Initiative administration. Solar companies are assigned into four categories in terms of the total sales during the sample period: top 1%, top 1–10%, top 10–50%, and bottom 50%.

Table 3: Impact of Rebate on Transaction Price: Overall Estimates

Regressors	Dependent Variable: price per AC-rated Watt						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rebate per Watt	0.790***	0.853***	0.826***	0.657***	0.640***	0.632***	0.671***
	(0.153)	(0.141)	(0.125)	(0.127)	(0.125)	(0.108)	(0.104)
Region FEs	Utility	Utility	County	County	County	Zip-code	Zip-code
Time FEs	Y-M	Date	Y-M	Y-M	Y-M	Y-M	Date
Component FEs	NO	NO	NO	YES	YES	YES	YES
Demographics	NO	NO	NO	NO	YES	NO	NO
Seller FEs	NO	NO	NO	NO	NO	NO	NO
No. of Obs.	67,057	67,057	67,057	67,057	67,002	67,002	67,002

Remarks:
Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the CSI-AC rating). Rebate level is defined as the rebate (dollar) divided by the CSI-AC-rated capacity. Component fixed effects are the model and brand of the solar panels and solar inverters. Demographics include median household income, median home value, and education levels at the zip-code level. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers).

Table 4: Impact of Rebate on Transaction Price: Separate Estimates I

Regressors	Dependent Variable: price per AC-rated Watt
	(1)
Rebate per Watt	0.454***
	(0.096)
Top 1% sellers×Rebate	0.456***
	(0.049)
Top 1%–10% sellers×Rebate	0.112***
	(0.037)
Region FEs	County FEs
Time FEs	Year-month FEs
Component FEs	YES
Demographics	YES
Seller FEs	NO
No. of Observations	67,002

Remarks:
Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the CSI-AC rating). Rebate level is defined as the rebate (dollar) divided by the CSI-AC-rated capacity. Component fixed effects are the model and brand of the solar panels and solar inverters. Demographics include median household income, median home value, and education levels at the zip-code level. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers).

Table 6: Demographic Characteristics of Households before and after Subsidy Changes

variable	means		p value of t-test
	10 days before	10 days after	
median hh income ($)	82916.76	81221.8	0.0254
median home value ($)	528213.9	510972.2	0.0420
pct of bachelors or higher (%)	24.27052	23.81045	0.0063

Remarks:
This table compares the household characteristics for households before the subsidy change versus those after the subsidy change. The t test rejects the null hypothesis that household characteristics are the same before and after.

Table 5: Impact of Rebate on Transaction Price: Separate Estimates II

Regressors	(1)	(1-g)	(2)	(2-g)	(3)	(3-g)	(4)	(4-g)
Rebate per kW	0.364***	0.339***	1.016***	0.920***	0.046	0.050	0.188*	0.094
	(0.050)	(0.059)	(0.090)	(0.096)	(0.057)	(0.072)	(0.097)	(0.135)
Green_car_pct×Rebate		0.556		1.706***		0.039		1.936
		(0.496)		(0.554)		(0.679)		(1.394)
Region FEs	County FEs	County FEs	County FEs	County FEs	County FEs	County FEs	County FEs	County FEs
Time FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs
Component FEs	YES	YES	YES	YES	YES	YES	YES	YES
Demographics	YES	YES	YES	YES	YES	YES	YES	YES
Seller FEs	YES	YES	YES	YES	YES	YES	YES	YES
Sample	ALL	ALL	Top 1% Sellers	Top 1% Sellers	1%--10% Sellers	1%--10% Sellers	Bottom 90% Sellers	Bottom 90% Sellers
No. of Observations	67,002	65,312	19,563	19,055	32,362	31,531	15,077	14,726

Remarks:
Clustered standard errors at city level are reported in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the CSI-AC rating). Rebate level is defined as the rebate (dollar) divided by the CSI-AC-rated capacity. Component fixed effects are the model and brand of the solar panels and solar inverters. Demographics include median household income, median home value, and education levels at the zip-code level. Green_car_pct is the percentage of electric or hybrid car adoption among all car registration at the zip-code level, with a mean of 0.057 and a standard deviation of 0.032. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers).

Table 7: Robustness Check I

Sample	(1)	(2)	(3)	(4)
Exclude x days within rebate changes				
x =	5 days	10 days	20 days	30 days
Top 1% Seller Sample	1.143***	1.159***	1.193***	1.290***
	(0.113)	(0.119)	(0.135)	(0.149)
No of Obs.	18,288	17,400	15,874	14,295
Top 1%–10% Seller Sample	0.047	0.013	-0.047	-0.050
	(0.067)	(0.072)	(0.079)	(0.086)
No of Obs.	30,224	28,812	26,154	23,800
Bottom 90% Seller Sample	0.272**	0.207	0.266*	0.312*
	(0.128)	(0.134)	(0.146)	(0.167)
No of Obs.	14,079	13,443	12,308	11,283
Component FEs	YES	YES	YES	YES
Demographics	YES	YES	YES	YES
Seller FEs	YES	YES	YES	YES

Remarks: Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the CSI-AC rating). Rebate level is defined as the rebate (dollar) divided by the CSI-AC-rated capacity. Component fixed effects are the model and brand of the solar panels and solar inverters. Demographics include median household income, median home value, and education levels at the zip-code level. Year-Month FEs and Utility territory FEs are included in all models.

Table 8: # of Solar Systems Sold on the Border of SCE/CSE

City	Service Area	
	SCE	CSE
Aliso Viejo	44	2
Coto De Caza	24	25
Laguna Beach	88	12
Laguna Hills	16	83
Laguna Niguel	33	78
Mission Viejo	203	56
Rancho Santa Margarita	75	12
Trabuco Canyon	89	8
Total	572	276

Table 9: Robustness Check II: Limit to Cities on the Border between SCE and CSE

Regressors	Price per kW					
	(1)	(2)	(3)	(4)	(5)	(6)
Rebate per kW	0.450	-0.097	0.413	-0.135	0.419	-0.122
	(0.360)	(0.331)	(0.380)	(0.374)	(0.387)	(0.372)
Top 1% Sellers×Rebate		1.006***		1.007***		1.005***
		(0.271)		(0.274)		(0.273)
Region FEs	Utility FEs	Utility FEs	Utility FEs	Utility FEs	Utility FEs	Utility FEs
Time FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs	Y-M FEs
Component FEs	YES	YES	YES	YES	YES	YES
Demographics	YES	YES	YES	YES	YES	YES
Seller FEs	YES	YES	YES	YES	YES	YES
Electricity Prices	NO	NO	Tier 5 Price	Tier 5 Price	Avg Price	Avg Price
No. of Observations	810	810	806	806	806	806

Notes: Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample limits to non-TPO home solar systems on the SCE and SDG&E border. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the CSI-AC rating). Rebate level is defined as the rebate (dollar) divided by the CSI-AC-rated capacity. Component fixed effects are the model and brand of the solar panels and solar inverters. Demographics include median household income, median home value, and education levels at the zip-code level. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers). Tier 5 Price is the highest marginal electricity price faced by a single-family home household (weighted average by year). Avg Price is the average electricity price faced by a single-family home household (also weighted average by year).

Table 10: Subsidy Pass-through Estimates by Product Quality Quantile I

Regressors	Dependent Variable: price per kW (nameplate-rating)		
	(1)	(2)	(3)
Perf_q0_25×Rebate	0.980*** (0.107)	0.203*** (0.063)	0.172 (0.171)
Perf_q25_50×Rebate	0.898*** (0.102)	0.091 (0.064)	0.273* (0.151)
Perf_q50_75×Rebate	0.887*** (0.099)	0.074 (0.045)	0.175 (0.165)
Perf_q75_100×Rebate	0.854*** (0.089)	0.054 (0.051)	0.044 (0.158)
Control for firm FEs	YES	YES	YES
Control for demographics	YES	YES	YES
Sample	Top 1% Sellers	1%–10% Sellers	Bottom 90% Sellers
No. of Observations	19,563	32,362	15,077

Remarks: Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the name-plate DC rating). Rebate level is defined as the total rebate (dollar) divided by the name-plate DC capacity. Demographics include median household income, median home value, and education levels at the zip-code level. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers).

Table 11: Subsidy Pass-through Estimates by Product Quality Quantile II

Regressors	Dependent Variable: price per kW (nameplate-rating)		
	(1)	(2)	(3)
Perf_q0_25×Rebate	0.880*** (0.121)	0.231*** (0.080)	0.100 (0.182)
Perf_q25_50×Rebate	0.912*** (0.100)	0.166* (0.095)	0.205 (0.260)
Perf_q50_75×Rebate	0.924*** (0.115)	0.125** (0.055)	0.102 (0.175)
Perf_q75_100×Rebate	0.841*** (0.096)	0.099 (0.067)	0.028 (0.155)
Perf_q0_25×Rebate×Green_car_pct	1.368** (0.680)	-0.439 (0.830)	1.735 (1.118)
Perf_q25_50×Rebate×Green_car_pct	-0.110 (0.500)	-1.159 (1.056)	1.709 (3.418)
Perf_q50_75×Rebate×Green_car_pct	-0.591 (0.696)	-0.831 (0.560)	1.608 (1.019)
Perf_q75_100×Rebate×Green_car_pct	0.205 (1.064)	-0.740 (0.648)	0.495 (1.057)
Control for firm FEs	YES	YES	YES
Control for demographics	YES	YES	YES
Sample	Top 1% Sellers	1%-10% Sellers	Bottom 90% Sellers
No. of Observations	19,055	31,531	14,726

Remarks: Clustered standard errors at city level are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. The sample consists of all non-TPO home solar systems in the California Solar Initiative Program. The dependent variable is unit transaction price, defined as the transaction price (dollar) divided by the capacity (kilowatt, measured by the name-plate DC rating). Rebate level is defined as the total rebate (dollar) divided by the name-plate DC capacity. Demographics include median household income, median home value, and education levels at the zip-code level. Green_car_pct is the percentage of electric or hybrid car adoption among all car registration at the zip-code level, with a mean of 0.057 and a standard deviation of 0.032. Prices and rebates are in Jan 2007 dollars (adjusted by California weighted average CPI for urban consumers).

Table 12: Numbers and Estimates in the Back-of-the-Envelope Calculation

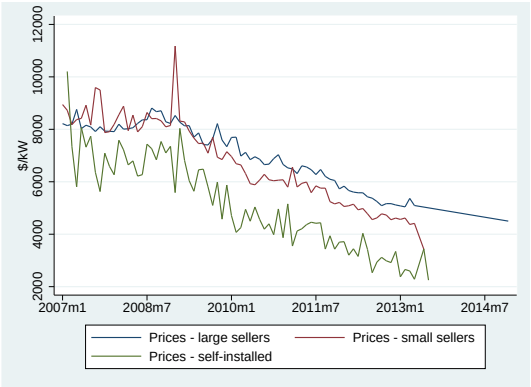
	Number of Projects	Average Capacity (DC-rated kW)	Average Rebate ($1,000/DC kW)	Average Performance Metric	Estimated Subsidy Pass-through Rate
	(1)	(2)	(3)	(4)	(5)
All	133,907	5.85	0.73	0.817	
from Top 1% Suppliers (L)	72,325	5.61	0.60	0.809	1.016
from Top 1–10% Suppliers (M)	44,491	6.08	0.76	0.823	0.046
from Bottom 90% Suppliers (S)	17,091	6.25	0.76	0.820	0.188

Remarks: This table shows the numbers and estimates used in the back-of-the-envelope calculation in Section 8. The sample includes all customer-owned systems and TPO systems. Numbers in columns (1)–(4) are computed from the data sample. Estimated subsidy pass-through rate for each segment is from columns (2), (3) and (4) in Table 5.

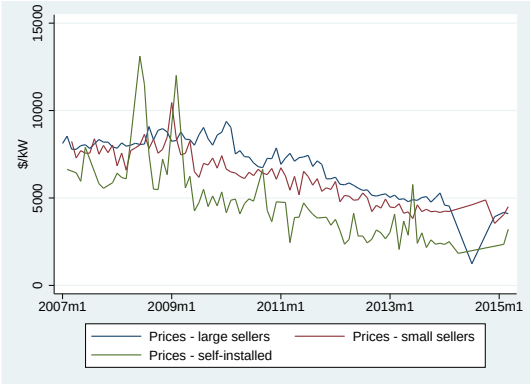
Appendix A: Figures

Figure A.1: Transaction Price by Firm Group

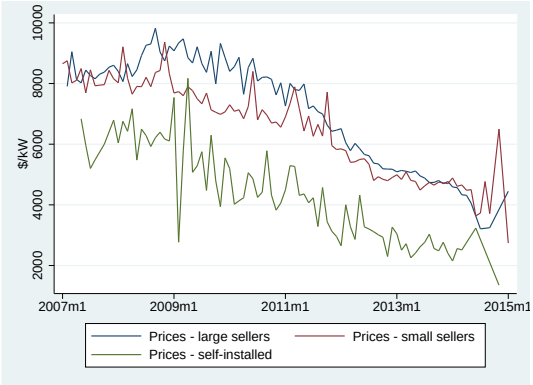
(a) PG&E



(b) CSE



(c) SCE



Remarks:
This set of charts shows the average price level set by each firm group. Large sellers are the top 1% sellers in terms of total sales. Small sellers are all the other sellers. The self-installed group consists of all solar systems that are self-installed by the consumers themselves (or other contractors so the installation fees are not included in the price).

Figure A.2: Rooftop Solar System Components

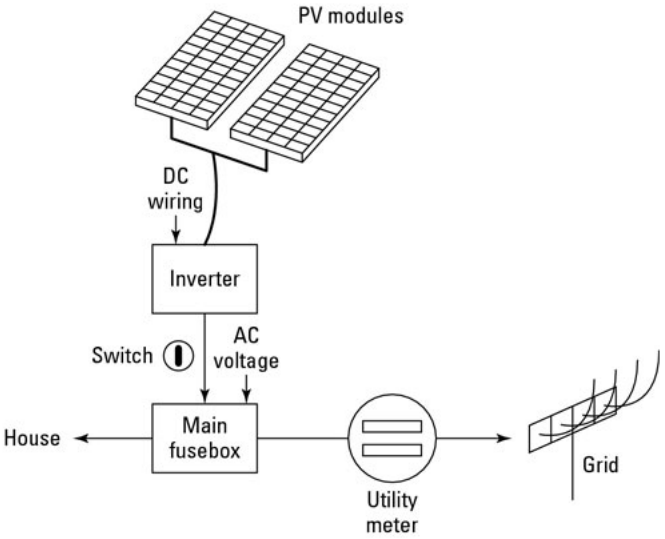
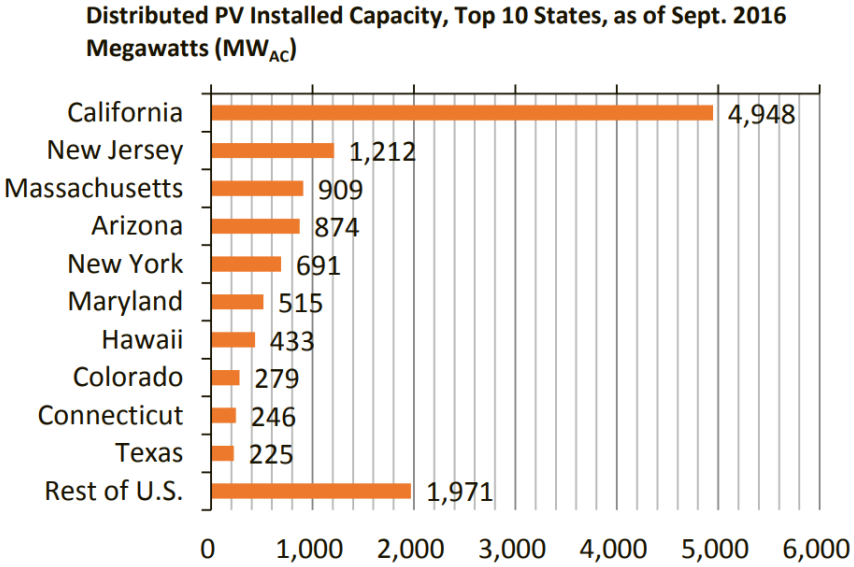
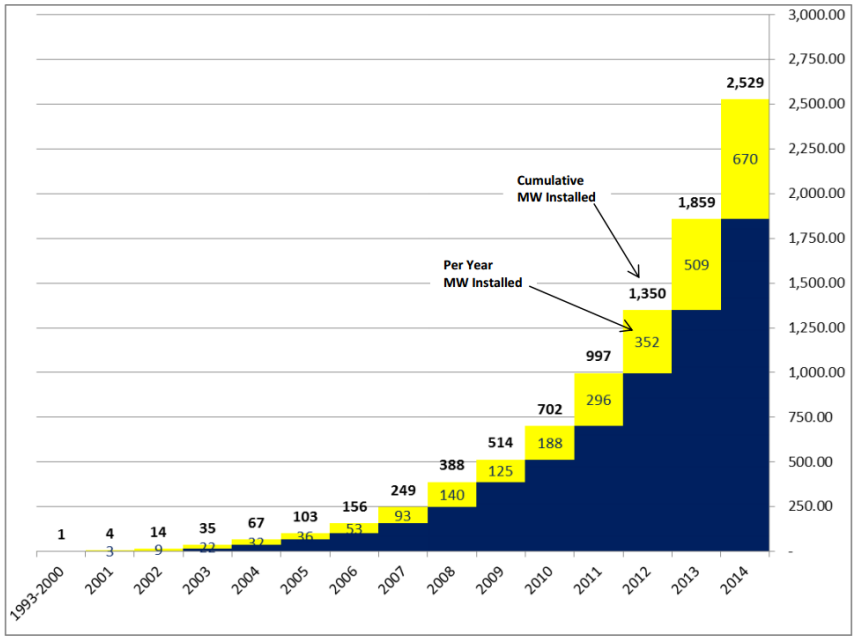


Figure A.3: Distributed Solar Capacity in the US and CA



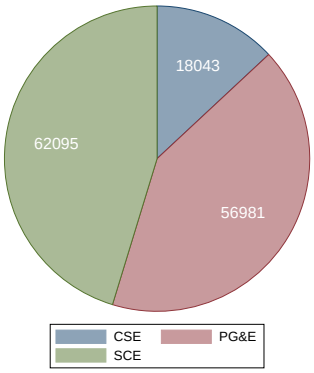
(a) Top 10 States of Distributed Solar in the US



Data is through December 31, 2014 and includes CSI, NSHP, ERP and SGIP data, but not POU or RPS data.

(b) Distributed Solar Capacity (MW) by year in CA's IOU Territories
Source: EIA, CSI 2015 Annual Report

Figure A.4: Number of Solar Projects in each Utility Territory



Notes: Sample includes both host-owned systems and 3rd-party-owned systems in the rebate program.

Figure A.5: Composition of Leased/PPA Systems

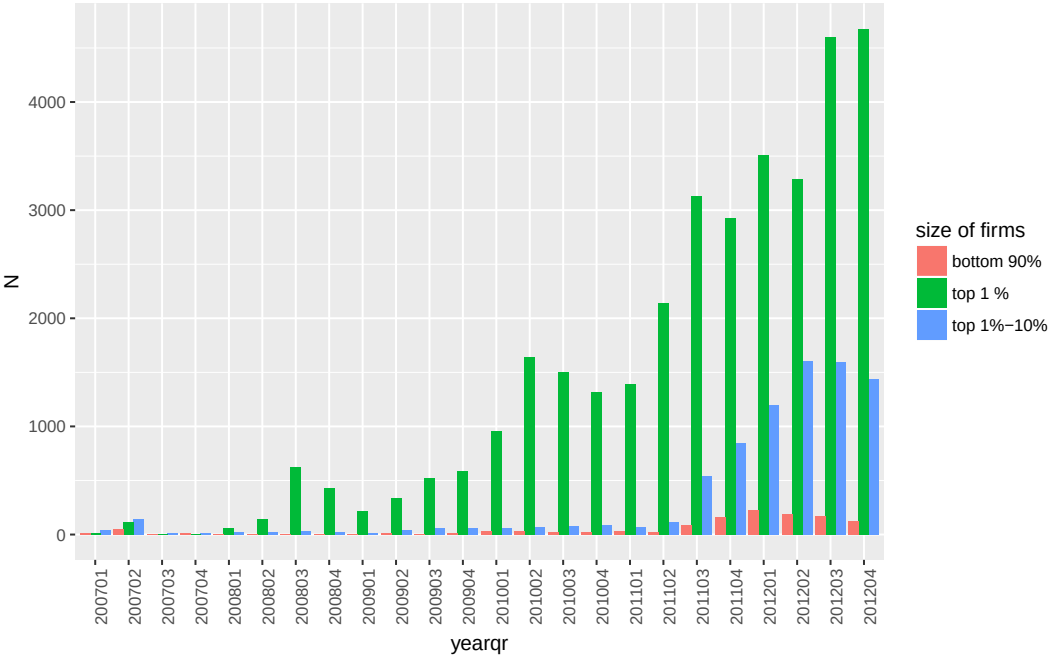


Figure A.6: Composition of Customer-owned Systems

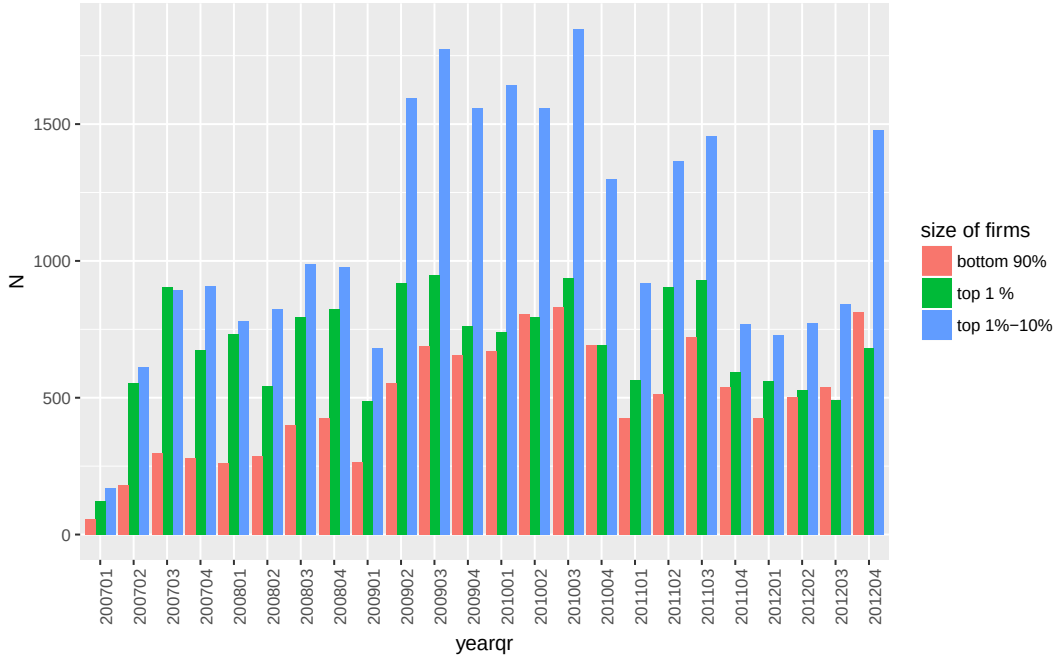
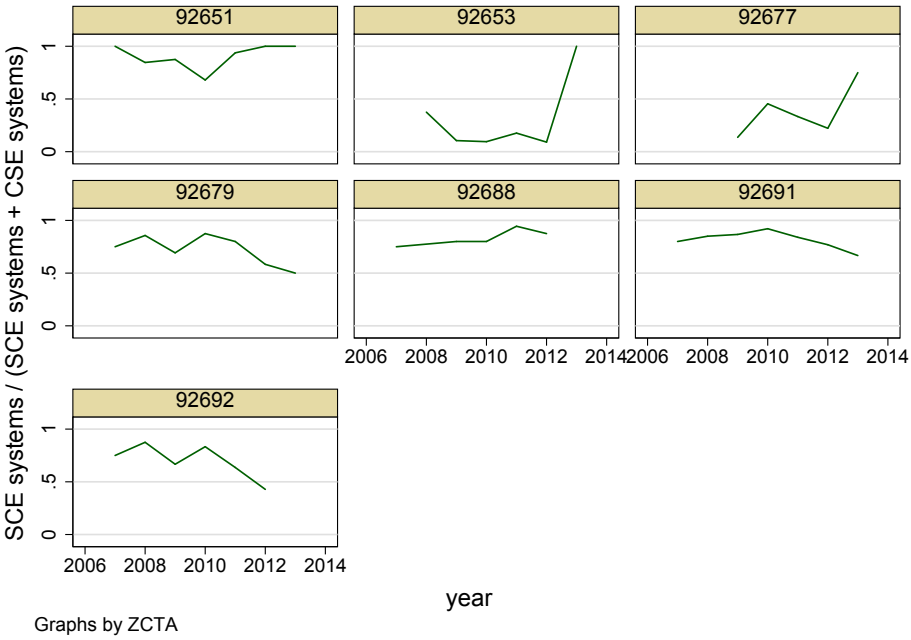


Figure A.7: The Fraction of SCE Systems by Year



Appendix A: Tables

Table A.1: Summary Statistics

VARIABLE	Obs	Mean	S.D.	Min	Max
Transaction Prices ($)	137119	34916.9	20965.3	1400	1028017
Rebate Amount ($)	137119	3464.1	4479.0	86	137895
System Capacity (kW)	137119	5.862	3.230	0.92	86.4

Notes: Transaction prices are pre-rebate prices.

Table A.2: Performance Metric and Returns of Systems

Regressors	Dependent Variable			
	log(daily kWh / total system cost)		log(daily kWh / system nameplate size)	
	(1)	(2)	(3)	(4)
log(performance metric)	0.652***	0.691***	0.776***	0.766***
	(0.020)	(0.021)	(0.017)	(0.020)
IOU FEs	YES	YES	YES	YES
County FEs	YES	YES	YES	YES
Installation Year FEs	YES	YES	YES	YES
Billing month FEs	YES	YES	YES	YES
Billing year FEs	YES	YES	YES	YES
TPO Indicator	YES	YES	YES	YES
Seller FEs	NO	YES	NO	YES
No. of Observations	29,289	28,840	29,244	28,733

Remarks:
OLS standard errors are reported in parentheses. * p<0.10, ** p<0.05, *** p<0.01. One observation is a monthly electricity output of a solar system. The sample consists of 543 solar systems. For each solar system, 60 consecutive monthly electricity outputs since its installation are observed. The dependent variable is monthly electricity output (kWh) divided by total system cost ($) in model specification (1) and (2), and monthly electricity output (kWh) divided by total system size in terms of nameplate rating (DC rated capacity) in model specification (3) and (4). Both dependent variable and independent variable (the performance metric), are log transformed. Control variables are utility territory fixed effects, county fixed effects, installation year fixed effects, billing month and year fixed effects, TPO system indicator and solar seller fixed effects.

Table A.3: Transactions by Zip-code on the Border of SCE/SDG&E(CSE)

Program administrator	Zip-code areas							
	92651	92653	92677	92679	92688	92691	92692	Total
CSE	14	82	78	33	16	24	29	276
SDG&E (SCE)	88	18	33	117	74	137	66	533

Notes: The number in each cell is the total number of non-TPO transactions in that zipcode-utility region.

Zip-codes on/around the utility territory borders (Table A.3 and Figure A.7)

- On the border of SCE/SDG&E(CSE): 92651, 92653, 92677, 92679, 92688, 92691, and 92692.
- On the border of SCE/PG&E: 93230 and 93286.
- Around the border of SCE/SDG&E(CSE): 92536, 92618, 92637, 92651, 92653, 92656, 92677, 92679, 92688, 92691, and 92692.
- Around the border of SCE/PG&E: 93110, 93230, 93274, 93286, 93292, 93401, 93657, 93664, and 95117.

Counties on/around the utility territory borders

- On the border of SCE/PG&E: Santa Barbara, Kern (city: Mc Farland), Kings (city: Hanford), Tulare (Woodlake), Los Angeles, Fresno
- On the border of SCE/SDG&E(CSE): Orange, Riverside

Appendix B: Text Analysis on Solar Company Web-pages

I explain the text analysis on solar companies' websites by two examples. One example is Tesla Solar, the largest solar supplier in California. The other example is a local bay area solar supplier, Slingshot Power. The screenshots of their websites are in Figure A.8.

Web scraping. I use Python to scrape the websites of solar companies, given the website URLs I have compiled. I scraped 611 solar companies' web-pages.

Text preparation. The Scraped texts contain special characters and punctuation. I clean the texts by removing these special characters and converting all words to lowercase. I next tokenize the texts. This means that I split sencenses based on white space and punctuation. For example, commas and periods are taken as separate tokens. Contractions are split apart (e.g. "What's" becomes "What" and "'s"). Quotes are kept. Since a quote does not convey much information, I also filter out these tokens (such as all standalone punctuation) that I am not interested in. Next, I filter out tokens that are stop words such as "a" and "to". I also stem words to their roots or bases so that "cheaper" becomes "cheap". The benefit of stemming is to both reduce the vocabulary and to focus on the sentiment of a document.

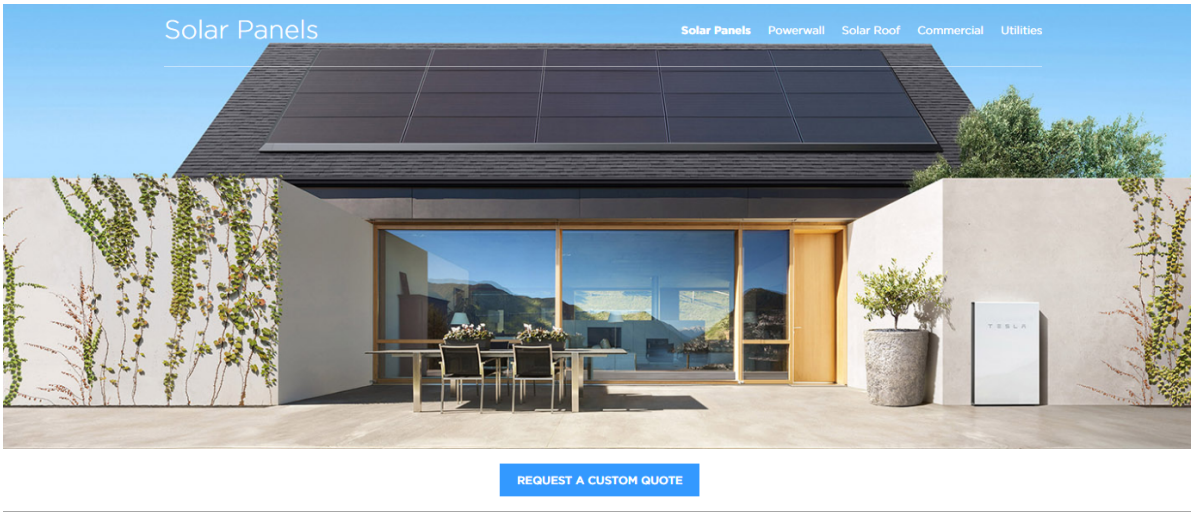
I next **convert words into TF-IDF metrics.** TF-IDF (short for term frequency-inverse document frequency) is to reflect how important a word is to a document in a collection of documents. A word with a high weight in TF-IDF means that this word has a high term frequency in the given document and a low document frequency in the whole collection of documents. This step is important as all websites are likely to contain a significant amount of the word "sun". Therefore, the word "sun" does not convey a lot of information to help categorize firm advertising strategies. By converting words into TF-IDF metrics, "sun" has a lower weight in the TF-IDF matrix if it shows up in every website.

Finally, I fit **k-means clustering** on this TF-IDF matrix.[31] The number of clusters I set is three. K-means clustering successfully identifies those websites with request errors and I call these websites Group 2. These websites contain words such as "domain", "com", and "web". The rest of the websites are categorized into two groups. Tesla Solar is categorized into the group (I call it Group 0) with relevant words: "sunpower", "best", and "repower". Whereas Slingshot Power is categorized into the group (I call it Group 1) with relevant words to be "construction", "service", and "projects".

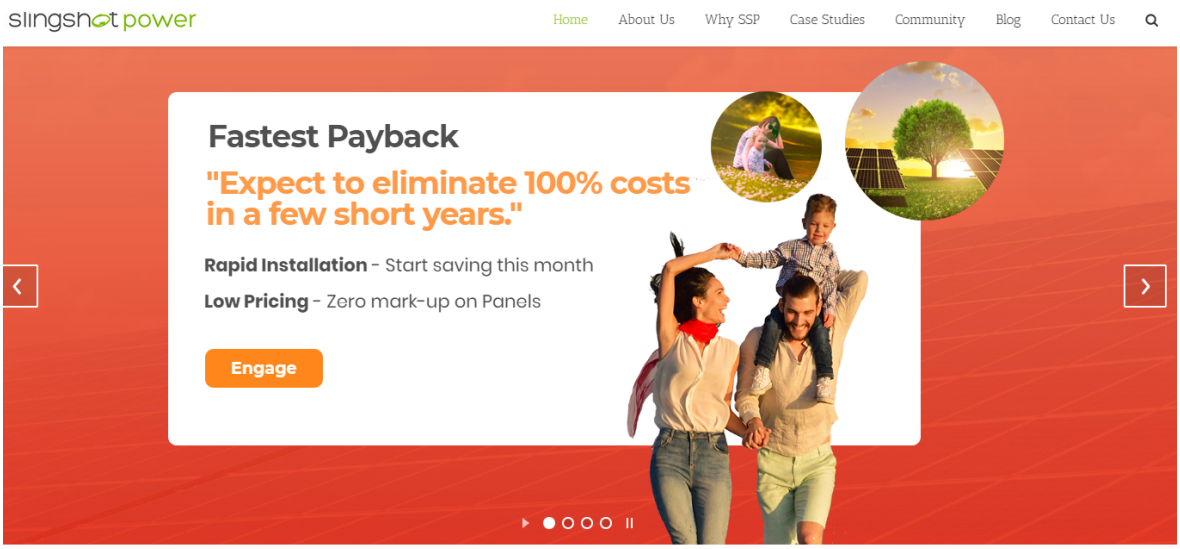
[31] The TF-IDF matrix is sparse. I reduce its dimension using truncated SVD.

Figure A.8: Two Web-page Examples

(a) A Screenshot of the Website of Tesla Solar (the Largest CA Solar Supplier)



(b) A Screenshot of the Website of A Local Solar Company in Bay Area Regions



Website urls: (a) https://www.tesla.com/solarpanels and (b) http://slingshotpower.com/

Scraped texts

https://www.tesla.com/solarpanels:
['\n Solar panels produce renewable clean energy while securing low utility rates. Our panels exceed industry standards for durability and lifespan. ', 'Our solar panels blend into your roof with integrated front skirts and no visible mounting hardware. The result is a clean, streamlined look.', 'Powerwall charges with energy produced by solar panels, making that energy available when needed, day or night. Powerwall also enables your solar panels to produce energy during grid outages. \n Learn more about our solar panels. ', 'By clicking Submit, I agree to be contacted at the number provided with more information or offers about Tesla products. I understand these calls or texts may use computer-assisted dialing or pre-recorded messages. This consent is not a condition of purchase.', '\nTesla Â© 2018\nAbout Elon Musk\nPrivacy & Legal\nContact\nCareers\nGet Newsletter\nForums\nLocations\nSign Out\n\nâ–¾\n\n', '\n\n Tesla, Inc \n', 'Solar Panels', 'Produce Your Own Clean Energy', 'Request a Custom Quote', '\n Solar Panels ', '\n Solar Panels ', 'Panels with a Sleek, Low-Profile Design', 'Seamless Integration with Powerwall', 'North America', 'Europe', 'Middle-East', 'Asia/Pacific', 'Become a Tesla Insider']

http://slingshotpower.com/:
['Fastest Payback ', 'Rapid Installation - Start saving this month Low Pricing - Zero mark-up on Panels\n', 'Engage', 'Easiest Service', ' Sizing Made Easy - Sizing to eliminate 100% electric costs.Next Week Schedule Guarantee - All customers select best install date with us. ', 'Engage', 'Best Products', 'Engage', 'Extra Production - Lowest degradation provides maximum production Extra Protections - Extraordinary 25 Year Production Warranty\n', 'Slingshot Power', 'â€œA social business enterprise delivering \na better solar experienceâ€\n', '1. Size the System\n2. Home Survey\n3. Sign and Save', 'Engage', 'Learn More >>>', 'Learn More >>>', 'Learn More >>>', 'We offer the most affordable solar power at the lowest prices, backed up by the best service and fastest installs. Why work with anyone else?\nLearn More >>>', 'Our team of experts will advise you on the best products to maximize returns on your investment.\nLearn More >>>', 'PHONE:\n1.650.260.7655', 'EMAIL:\nsupport@slingshotpower.com', 'VISIT:\n1718 Stone Ave. Suite B\nSan Jose, CA 92125', 'Slingshot Power services\nthe entire San Francisco Bay Area.', 'San Jose Solar Power Installation', 'Launch your Clean Energy future', 'California\xa0# 984488', 'Privacy Policy', 'Fastest Payback', 'Easiest Service', 'Best Products', 'Residential Solar Panel Installation', 'Hotel Solar Panel Installation', 'Happy Customers', 'Happy Workers', 'Happy Planet', 'â€œExcited to Recommend Slingshot Powerâ€', 'â€œProud to install solar for Slingshot Powerâ€', 'Our company receives 90% Referrals', 'We will match any other offer on the market guaranteeing you the best packaged value', 'Local Crews helping your community', 'â€œFeels good to save money doing goodâ€', '"Expect to eliminate 100% costs in a few short years."', '"100% Satisfaction"', '"Only best products are offered to avoid low production."', 'Start Today!']

Appendix C: Identification of Subsidy Pass-through Rate

To interpret β_1 as the subsidy pass-through, it requires that variations in the rebate level across utility territories are exogenous. Let's see whether this is true by examining how equilibrium price is determined. Suppose that the transaction price P_{ikt} is determined by the rebate level R_{kt}, demand shifters, supply shifters, and an independent and identically distributed error term u_{ikt}. Then further suppose that we have the following linear functional form that determines the equilibrium price (The current functional form implies that the demand and supply shifters do not affect the subsidy pass-through rate).

$$P_{ikt} = \beta_0 + \beta_1 \cdot R_{kt} + \beta_2 \cdot \text{DEMAND\_SHIFTERS}_{ikt} + \beta_3 \cdot \text{SUPPLY\_SHIFTERS}_{ikt} + u_{ikt} \tag{3}$$

The demand shifters can be decomposed into three components: the system-level relative bargaining power of buyer that picks up the buyer characteristics, the value of solar in utility region k that picks up the geographical variation in solar value, and the value of solar at time t that picks up the varying expectation/awareness of the value of solar system over time.

$$\begin{aligned}\text{DEMAND\_SHIFTERS}_{ikt} = {} & \text{BARGAIN\_POWER}_i \\ & + \text{VALUE\_OF\_SOLAR}_k + \text{VALUE\_OF\_SOLAR}_t\end{aligned}$$

On the other hand, the supply shifters can be decomposed into the system-level cost and the market competitiveness of market m in utility region k at time t. The former component picks up the consumer-specific acquisition costs and site-specific design and installation costs.

$$\text{SUPPLY\_SHIFTERS}_{ikt} = \text{COST}_i + \text{MKT\_COMPETITION}_{kmt}$$

Therefore, by replacing and rearranging the terms, Equation (3) can be expressed as the following equation

$$\begin{aligned}P_{ikt} = {} & \beta_0 + \beta_1 \cdot R_{kt} \\ & + \beta_2 \cdot (\text{VALUE\_OF\_SOLAR}_k + \text{VALUE\_OF\_SOLAR}_t) \\ & + \beta_3 \cdot \text{MKT\_COMPETITION}_{kmt} \\ & + (\beta_2 \cdot \text{BARGAIN\_POWER}_i + \beta_3 \cdot \text{COST}_i + u_{ikt})\end{aligned} \tag{4}$$

where VALUE_OF_SOLAR_k and VALUE_OF_SOLAR_t can be controlled by the utility region

fixed effects and time fixed effects respectively. MKT_COMPETITION can be controlled by including measurement of local market conditions, for example HHI. BARGAIN_POWER$_i$ and COST$_i$ are unobserved and cannot be controlled for. However, under certain conditions, they are idiosyncratic shocks and independent of the rebate level R_{kt}.

To see this, consider the determination of R_{kt}. Rebate level R_{kt} can be expressed as a conditional function of the pre-specified rebate design that specifies the trigger values (cumulative capacity) in Table 1. Given the rebate design, the rebate level in region k at time t is determined by the cumulative capacity installed up to time t, i.e., $\sum_{\tau=1}^{t} Q_{k\tau}$. The equilibrium $Q_{k\tau}$ in each period can be further expressed as the function ($\mathbb{G}(\cdot)$) of the local demand shifters and supply shifters. Under some regularity conditions (for instance, $\mathbb{G}(\cdot)$ is linear), the order of summation and the G function are interchangeable. Finally, rebate level of region k at t is a function of the average demand shifters and the average supply shifters over all the idiosyncratic shocks of the system-level bargaining power and system-level site-specific cost up to time t in region k, given the pre-specified rebate program design.

$$
\begin{aligned}
R_{kt} &= \mathbb{F}\left(\sum_{\tau=1}^{t} Q_{k\tau} \mid \text{rebate\_design}\right) \\
&= \mathbb{F}\left(\sum_{\tau=1}^{t} \mathbb{G}\left(\text{DEMAND\_SHIFTERS}_{k\tau}, \text{SUPPLY\_SHIFTERS}_{k\tau}\right) \mid \text{rebate\_design}\right) \\
&\overset{cond}{=} \mathbb{F}\left(\mathbb{G}\left(\sum_{\tau=1}^{t} \text{DEMAND\_SHIFTERS}_{k\tau}, \sum_{\tau=1}^{t} \text{SUPPLY\_SHIFTERS}_{k\tau}\right) \mid \text{rebate\_design}\right) \\
&= \tilde{\mathbb{F}}\left(\text{avg\_DEMAND\_SHIFTERS}_{kt}, \text{avg\_SUPPLY\_SHIFTERS}_{kt} \mid \text{rebate\_design}\right)
\end{aligned}
$$

Therefore under these conditions, as long as the system-level shocks are mean-zero, then they are independent of the rebate level R_{kt} and the coefficient β_1 in Equation (2) is identified as the subsidy pass-through rate. On the other hand, if $\mathbb{G}(\cdot)$ is non-linear and the system-level shocks are not i.i.d. across region and across time, then the estimation of β_1 could be biased. For example, suppose the supply curve remains linear but the demand curve is convex, then a positive supply shock that shifts the curve out would cause a much larger effect on the equilibrium Q (as well as R) than a negative supply shock. Imagine that the probability of a low consumer acquisition cost in PG&E region is much higher than in the other two utility territories, then this would lead to a lower P as well as a lower R, which ruins the identification. To deal with this threat, I exploit spatial discontinuities in electricity service areas with in the same city to control for these unobserved factors as a robustness check.