Access to Alternatives: Increasing Rooftop Solar Adoption with Online Platforms

Jackson Dorsey*

February 7, 2019

Abstract

As of 2016, only 1.1% of U.S. rooftops had solar panels despite government subsidy support and falling hardware costs. In this paper, I estimate a structural model of the residential solar PV market using new detailed data on seller bids and consumer choices. The results illustrate that search/informational frictions are an important barrier to solar PV adoption for two reasons: (1) installers can charge higher markups and (2) consumers fail to connect with high-quality sellers in the market. Counterfactuals simulations show that solar PV adoption increases by 84% when buyers can solicit additional price quotes through an online platform.

JEL Codes: D22, D44, L11, L15, Q40, Q58
Keywords: solar PV, renewable energy, multi-attribute auctions, search costs, online platforms, market power, quality, technology adoption

*Department of Business Economics and Public Policy, Kelley School of Business, Indiana University (e-mail: jfdorsey@iu.edu). This paper was previously circulated with the title: "Harnessing the Power of the Sun through Online Platforms: Evidence from the Rooftop Solar Market." I thank Ashley Langer, Derek Lemoine, Gautam Gowrisankaran, Stan Reynolds, Gary Solon, Derek Wolfson, Keith Meyers, Dev Bose, Alvin Murphy, Matt White, Stefan Weiergraeber, Dan Sacks, and seminar participants at Arizona State, University of Arizona, University of Maryland (AREC), Cornerstone Research, University of Mannheim, Federal Reserve Bank (Dallas), Charles River Associates, SUNY Albany, Indiana University (Kelley), ISO New England, International Industrial Organization Conference (IIOC), NBER (EEE) Summer Institute, and the CU Environmental Economics Workshop for helpful comments, discussions, and suggestions. I also thank Luke Tarbi and his team at EnergySage Inc. for providing instrumental help with data access.

Solar photovoltaics (PV) enable homeowners to generate carbon-free electricity by bolting PV panels to their rooftop. Therefore, both local and federal governments have enacted a multitude of policies to reduce electricity sector emissions by encouraging investment in residential solar PV. Indeed, these generous subsidy programs and plummeting PV hardware costs have bolstered a substantial increase in total U.S. solar PV capacity. Nonetheless, only one in every hundred rooftops had solar panels in 2016 (Margolis et al., 2017).

Despite policy support and recent technological advances, there remain several barriers inhibiting diffusion of rooftop solar PV. Currently, the vast majority of residential solar installations occur through a decentralized market where sellers do not post prices. Therefore, in addition to the cost of hardware (i.e., the panels themselves), households face substantial indirect costs of searching and comparing quotes between different sellers in the market. These search frictions hinder the growth of rooftop solar PV for a few reasons. First, if collecting price quotes is costly for buyers, any installer asked to give a quote can expect to be bidding against few or no other sellers, thereby giving that installer incentive to charge a higher markup.[1] Secondly, if there is vertical differentiation in installation services or hardware offerings, some potential buyers will fail to consider more preferable suppliers in the market.

One potential way to mitigate these issues is through an intermediary (e.g., an online platform) that connects buyers with additional sellers.[2] Providing consumers with additional bids could increase solar PV adoption by enhancing price competition and linking potential adopters with more experienced and reputable installers.

In this paper, I empirically evaluate how the introduction of an intermediary affects markups, solar panel adoption, and welfare in the rooftop solar market. To do so, I collect proprietary data from an online platform that facilitates auctions for residential solar PV installation projects. The data provides detailed information on purchase choices and seller bidding behavior, which I use to estimate a structural model of the solar PV market. Estimating the model involves combining discrete choice methods to estimate buyers' choice rule and adapting insights from the empirical auctions literature (Guerre et al., 2000; Krasnokutskaya and Seim, 2011) to infer sellers' marginal costs (costs of installation) and bid preparation costs. After recovering the primitives of the model, I simulate counterfactual market outcomes under different assumptions about competitive structure and subsidies in the market. In particular, I compare solar PV adoption rates and welfare when buyers obtain a limited set of quotes (decentralized market) to a case where buyers obtain additional bids through the online platform.

I find that access to the platform increases solar panel adoption by 84%. There are two

[1] An influential literature starting with Stigler (1961) and has discussed equilibrium pricing behavior of firms and consumers' optimal search effort in markets where consumers lack perfect information about prices. Diamond (1971) and Stahl (1989) show that even small consumer search costs can lead to imperfectly competitive pricing by firms.

[2] Previous work has also theoretically investigated the role of intermediaries in search markets (Gehrig, 1993; Hall and Rust, 2003; Spulber, 1996). More recently, several empirical studies examine the effect of introducing an intermediary or a technology that increases price transparency in other industries such as life insurance (Brown and Goolsbee, 2002), fisheries (Jensen, 2007), waste management (Salz, 2017), health care (Brown, 2017), and retail gasoline (Luco, 2016).

primary drivers of this result: 1) the market becomes more competitive which drives down the lowest quoted price by 7% on average and, 2) buyers connect to higher quality sellers and better hardware offerings. While price reductions are an important driver of adoption, I find that reduced prices alone can only explain 51% of the increase in adoptions caused by the platform. More surprisingly, improved access to high-quality sellers and equipment alone account for 35% of the rise in adoption. Due to heterogeneous preferences across buyers, obtaining more bids also increases adoption by linking households to better-matched sellers. In addition to the increase in adoption, the platform leads to an 84% rise in total welfare. Consumer surplus grows by 117% as consumers enjoy lower prices and access to more (and higher quality) sellers. Sellers benefit from a 72% increase in profits because the fall in prices is more than offset by the rise in the sales quantity.

To put the results in context, I compare the effects of the platform to an existing subsidy policy, the federal investment tax credit (ITC). In particular, I run an additional counterfactual simulation where instead of giving consumers access to the platform (additional bids), I hold the average number of bids constant and provide a subsidy equal to 30% of the system price, equivalent to the ITC. I find that the subsidy policy causes a 56% increase in the number of solar systems installed. Surprisingly, the 56% increase generated by the subsidy policy is smaller than the 84% increase induced by giving consumers additional quotes through the platform. The effect of the platform is larger because it not only decreases prices but also gives buyers access to additional high-quality installers and improved hardware choices.

The results have important implications for public policy. As of 2016, only 3% of solar customers used a platform to purchase a PV system (O'Shaughnessy and Margolis, 2017). The counterfactuals suggest that policymakers could increase solar PV adoption by developing their own platforms or by encouraging participation on existing platforms. As one example, the state of Connecticut recently introduced a state-sponsored platform to decrease residential solar prices and increase adoption. Platforms have also been used to promote competition in other industries such as healthcare.[3]

The majority of existing policies aimed at spurring the adoption of solar PV systems have used subsidies to reduce the cost of adoption.[4] Although these policies have been effective at increasing uptake of solar PV,[5] they require considerable public expenditures and often are at risk of being removed. For instance, the federal production tax credit for wind energy has expired and been renewed several times over the past decade. The federal solar ITC is now scheduled to be eliminated by 2023. The counterfactual results imply clear benefits of channeling more potential solar PV buyers towards a centralized marketplace where they would receive more

[3]A primary aim of the Affordable Care Act was to reduce premiums and expand health insurance coverage for Americans. One principal mechanism for meeting that goal was the establishment of new individual health insurance marketplaces (platforms) where consumers can shop for, compare, and purchase plans.

[4]See Borenstein (2017) for a discussion of the role of electricity rate structure and government subsidies on consumer adoption of solar PV, as well as the distributional effects of subsidy programs.

[5]Hughes and Podolefsky (2015) and Burr (2014) find that one subsidy program, the California Solar Initiative, doubled the amount residential solar PV purchases in between 2007 and 2013.

bids. Moreover, because there are already existing platforms to connect buyers and sellers in the solar PV market, the cost of implementing a platform (or marketing the use of a platform) would likely come at a much lower price compared to a subsidy program.

While the literature on residential solar adoption is rapidly expanding, there remains a significant gap in empirical work to understand how non-price factors such as consideration sets influence adoption choices. The majority of existing work has exclusively focused on estimating the adoption response to price or subsidy changes (Burr, 2014; De Groote and Verboven, 2016; Gillingham et al., 2016; Gillingham and Tsvetanov, 2014; Hughes and Podolefsky, 2015; Langer and Lemoine, 2017; Pless and van Benthem, 2017; Reddix II, 2015).[6] To my knowledge, there are no existing estimates in the literature that quantify the importance of non-price attributes on consumers' solar PV adoption decision. To fill this gap, I exploit rich data on consumers' choice sets and model the adoption decision as a discrete choice between differentiated products. An advantage of my approach is that I can jointly estimate consumers' sensitivity to price as well as other attributes such as panel quality and seller quality that could also be instrumental in consumers' decision to adopt solar. For instance, an installer's warranty package, years of experience, customer service, and star rating are all non-price factors that likely influence consumer purchase choices.

The demand estimates show that non-price factors are indeed critical to consumer choices. Namely, the average buyer would be willing to pay 27% more for a high-quality seller compared to a median quality installer.[7] I also find that the average buyer is willing to pay a significant premium for higher quality equipment (i.e., more efficient panels).

Supply-side behavior is also crucial in determining equilibrium prices and adoption rates. However, most previous work has not explicitly modeled or discussed the supply-side of the solar PV market. A few notable exceptions include Bollinger and Gillingham (2014), Pless and van Benthem (2017), O'Shaughnessy and Margolis (2017), and Gillingham et al. (2016) who all provide evidence of market power in the residential solar industry.[8] In general, modeling the supply-side of the market has proved challenging because public solar PV datasets do not provide information on which installers are operating in specific geographic areas or which installers are in each consumer's consideration set. Therefore, researchers have either been unable to estimate the exact size of markups or needed to make strong assumptions about how

[6] Hughes and Podolefsky (2015) and Gillingham and Tsvetanov (2014) both use reduced-form approaches to estimate the elasticity of demand for residential solar systems and to quantify the adoption response to subsidy programs. Burr (2014), Reddix II (2015), De Groote and Verboven (2016), and Langer and Lemoine (2017) all develop dynamic discrete choice models to estimate demand for solar PV systems and to assess the welfare effects of different subsidy policies.

[7] I estimate firm fixed effects to determine the quality of sellers. Here, I define high-quality sellers as those with firm fixed effects in the 9th decile.

[8] In a related paper, Gerarden (2017) studies the impact of subsidies on solar panel manufacturers' investments to reduce costs. Gowrisankaran et al. (2016) investigate how intermittency from solar power affects utility investments and grid operations.

to measure local market structure (which sellers are in a buyer's choice set).[9] In this paper, I improve on the existing literature by collecting new data from an online platform which allows me to observe the choice set of each buyer, including non-selected sellers' bids. I use the data to develop and estimate a model of seller bidding and participation in multi-attribute auctions. A multi-attribute auction is a procurement mechanism where bidders submit multi-dimensional bids (i.e., bid price, panel efficiency, and star-rating). I build on methods recently developed by Krasnokutskaya et al. (2019) and Yoganarasimhan (2015) to recover the model parameters. To infer markups, I use each seller's first-order condition for an optimal price bid to decompose observed bid prices into a marginal cost and a markup component. I also estimate sellers' bid preparation costs using firms' observed participation decisions in the auctions. The estimated supply model reveals substantial market power. Specifically, I find that average gross markups account for nearly 40% of system prices.[10] High markups are particularly problematic in this market because they counteract government subsidy programs that aim to expand solar PV adoption.

This is the first paper to specify a structural model of the supply and demand side of the residential solar market in a unified framework. The full model can provide a better understanding of how welfare and equilibrium adoption may change under different policy environments. In the next section, I discuss the details of the online platform, provide descriptive statistics, and show regression evidence that buyers using the online platform pay lower solar PV prices on average compared to buyers that do not. In Section 3, I develop a model of buyer and seller behavior in the solar PV market and then discuss the methods used to pair the model to the data. Section 4 presents the results, and Section 5 concludes.

2 Data and Setting

Shopping for a rooftop solar system can be a time-consuming endeavor. Installers typically do not post prices because installation costs can vary depending on input costs, location, rooftop characteristics, size of the system, the efficiency of the modules, and other factors. Therefore, inquiring buyers often need to make phone calls to individual installers and then schedule a site visit just to obtain a project proposal and a price quote. Because search is costly in this market, many buyers only receive a limited number of price quotes, which increases the incentive for sellers to exercise market power. A 2015 survey found that the median potential consumer obtained just two quotes before making a decision (Aggarwal, 2015). This market environment

[9] Bollinger and Gillingham (2014) develop a dynamic model of installer pricing under imperfect competition to motivate a reduced-form regression equation that allows them to estimate static and dynamic markups; the authors use county HHI to proxy for the level of competition in a local market. Pless and van Benthem (2017) measure the pass-through rate of solar subsidies to prices and find a pass-through rate of over 100% and demonstrate this is only possible in the presence of market power. O'Shaughnessy and Margolis (2017) find that more competition results in lower prices in the residential solar market.

[10] Gross markup is the price minus the marginal cost of the installation and does not include installer overhead costs. After accounting for overhead costs, my estimates imply that average net markups make up 23% of prices.

also poses a challenge for researchers aiming to document market power because it is hard to observe how many installers are actually in a buyer's choice set. Most publicly available data on solar installations only includes the chosen seller, so it is impossible to know which other installers the customer was considering.

In this study, I directly observe buyers' choice sets by collecting proprietary installer price quote and consumer purchase data from EnergySage Inc. EnergySage operates an online platform that connects potential solar customers to a network of solar PV installers. The platform was the first online medium that allowed buyers to shop and compare solar system offers from different sellers in one place.

The EnergySage platform allows households interested in installing solar PV to conduct their own multi-attribute auctions to select an installer for their project. Each EnergySage "auction" includes several stages. First, consumers create an account with the website and provide necessary information such as the physical address of the potential installation and a monthly electricity bill.[11] Second, registered installers[12] receive a notification of the project which includes details such as a Google Maps photo of the buyer's roof (depicted in Figure 8 of the Appendix), as well as the monthly electricity usage of the buyer. Installers are then able to submit a project quote to the buyer which includes the system price, panel brand, inverter brand, and details about the seller such as a customer rating and a description of their solar installation experience. Finally, after installers have submitted their bids, the potential consumers can select one of the quotes and move forward with the transaction, or they can opt not to purchase any of the offers.[13] Figure 9 in the appendix shows an example of the purchaser's comparison tool on the platform.

A distinguishing feature of this environment is that buyers can base their selection off any criteria they choose and are not obligated to purchase the quote with the lowest price. Although many installers offer leases or power purchase agreements, 97% of buyers on EnergySage choose to purchase a system with cash or through a loan. The large skew towards purchases (instead of leases) is likely because EnergySage provides a calculation of the net present value of each offer and purchased systems nearly always offer higher overall value to the buyer. Additionally, overall market shares for leased systems have been declining considerably in recent years. Finally, most bids on the EnergySage marketplace are from local installers or regional installers. SolarCity, the nation's largest installer, leases out the majority of their systems but is not active on the platform.

[11] Alternatively, users can submit an estimate of their monthly electricity use.

[12] In order to submit bids, installers must first be pre-screened by EnergySage to ensure they are licensed, insured, and experienced.

[13] Buyers and sellers can communicate with each other via private messaging or phone calls before a selection is made. However, sellers cannot call buyer unless they are requested to do so by the buyer.

2.1 Descriptive Statistics

The analysis includes potential solar PV projects that occurred in 2015 and 2016 within the states of California, Connecticut, Massachusetts, and New York. I only consider residential projects and drop any projects that are smaller than 2KW or larger than 20KW in capacity.[14] I also do not include projects where a lease or power purchase agreement was selected.[15]

Table 1: Project Summary Statistics

	Mean	SD	10-%tile	50-%tile	90-%tile
Realized Number of Bids	3.95	1.90	1.00	4.00	6.00
Project Size (Watts)	7155.92	3138.63	3631.67	6560.00	11515.00
Distinct Panel Brands	2.80	1.28	1.00	3.00	4.00
# of Bids w/ Premium Panel	1.49	1.40	0.00	1.00	3.00
# Bids w/ Premium Plus Panel	0.16	0.41	0.00	0.00	1.00
# of Bids w/ Microinverter	2.93	1.85	1.00	3.00	5.00
# of Bids from Permanent Sellers	3.24	1.73	1.00	3.00	6.00
Observations	10545				

After dropping commercial projects and projects over 20KW in size, the EnergySage data set includes 10,545 potential projects or "auctions". Figure 1 maps the locations of all of the projects in the sample and Table 1 provides descriptive statistics of the projects.[16] The average project received just under four bids. However, there is noticeable variation in the number of bids across projects, which can be seen in Figure 10 of the appendix. The size of potential projects also varies widely. The median project had a 6,560-watt capacity, but the standard deviation was over 3000 watts, with 80% of projects between 3,631 watts and 11,515 watts. Buyers also often obtain quotes for several different types of panel (module) brands.[17] EnergySage also shows buyers a rating classification of each panel brand, premium panels have higher efficiency and better warranties.[18] More efficient panels are attractive because for a given physical system size[19] a more efficient panel will create more electricity. The average buyer receives 1.49 bids that are "premium" panels.[20] EnergySage also identifies the very highest quality panels as "premium plus",[21] however, these panels are much more expensive and less than 15% of households receive a bid offering a "premium plus" panel. Another vital component of a solar system is the inverter.

[14] I define each project's size as the mean size quote for that project.

[15] I drop these projects because comparing per-watt prices for leases vs. purchases is not straightforward. Furthermore, these projects compose less than 4% of choices and thus discluding them is unlikely to have major effects on the analysis. I also drop a handful of price quotes below $2/watt that appear to be miscoded.

[16] Also see Appendix Figure 14 for a map of the installers' imputed locations.

[17] A PV module consists of many PV cells wired in parallel. A panel can consist of one or more modules and is the largest hardware component of a system in terms of size and cost.

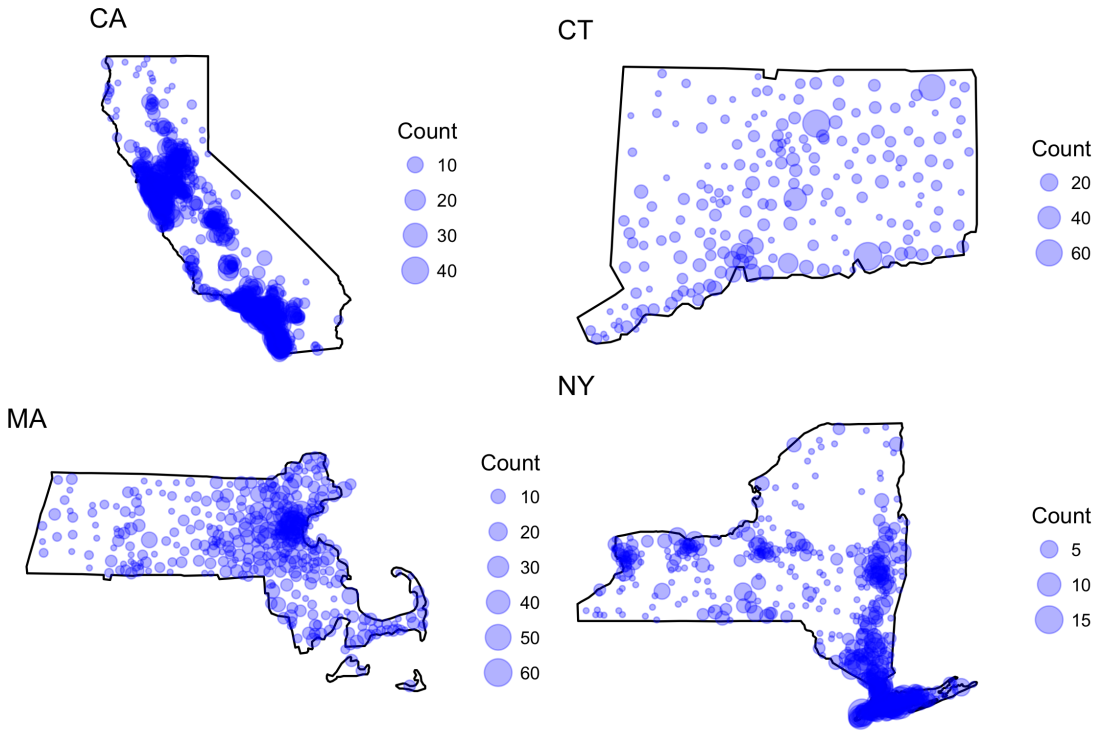
[18] EnergySage designates LG Electronics panels as "premium".

[19] Physical size is distinct from capacity, if two panels with the same capacity but one is more efficient, the more efficient panel will be physically smaller.

[20] Each seller is only allowed to place a single bid. For example, a seller cannot place two different bids for different panel qualities.

[21] Panels from SunPower Corporation are "premium plus".

Figure 1: Potential Project Locations



Notes: Count is the total number of potential projects within a ZIP code during the full sample.

The inverter converts the direct current (DC) output of the PV panel into alternating current (AC). String inverters are the cheapest and most commonly deployed inverter technology and can perform well if there is no shade at the project location at any time during the day. However, a system with a string inverter will only produce as much as the least productive panel in an array, leading to reduced output if shade covers part of the roof. Microinverter and power optimizer technologies can help the system to perform better in partial shade conditions but typically are more expensive. Some quotes in the data include microinverter technologies (in the analysis, I include power optimizers in the microinverter category), the average household received just under three bids with microinverters.[22]

Projects also receive quotes from installers with varying levels of experience within the platform. The average buyer received 3.35 bids from "permanent" sellers, which I define as an installer that made at least 100 bids and at least one sale on the platform during the sample. Potential customers can see reviews and star ratings of each installer, how many years the installer has been installing solar, and how many total projects they have completed. Given this information, buyers can form perceptions about a seller's quality. Sellers' quality is thus an important component to account for in the empirical analysis. Unfortunately, the installer

[22]The data does not distinguish explicitly between microinverter and string-inverter bids but does list the inverter brand. I define a bid as having a microinverter if the inverter brand is Enphase Energy or SolarEdge Technologies. These two companies together controlled 95 percent of the module-level power electronics market in 2015.

identities and star ratings over time are not available in the data. However, I do observe a unique installer identification number that can be used to track the behavior and performance of the same installer over time. I also observe a cross-section of each seller's star ratings and total installation experience. The seller ratings and experience information was collected in September 2017, nine months after the sample period ended. I report summary statistics for the installers in Appendix Table 10. As of September 2017, 76% of sellers had ratings on EnergySage, and of them, 84% had a 5-star rating. The average seller had seven ratings, had completed 1,579 total residential installations, and had been installing solar for nine years. In section 3.4, I discuss my approach used to account for seller quality in the analysis.

The number of active installers participating in auctions in each state increases over time (see Appendix Figure 11). At the beginning of the sample, each state had between 17 and 28 distinct installers, and by the end of 2016 each state had between 25 to 58 different sellers submitting bids.[23] A substantial portion, 70%, of the 220 installers are transient sellers that made fewer than 100 total bids. However, each permanent seller bids much more often, so bids from transient sellers only compose about 13% of all price quotes.

Table 2: Summary Statistics - Bid Characteristics

Panel A: Full Sample

Price ($/watt)	3.611
	(0.494)
Premium Panel (0,1)	0.386
	(0.487)
Premium Plus Panel (0,1)	0.0412
	(0.199)
Microinverter (0,1)	0.761
	(0.426)
Permanent Seller (0,1)	0.841
	(0.365)
Observations	40575

Panel B: CA South - 2016H1

	Selected Bid (0,1)	
	0	1
Price ($/watt)	3.554	3.442
	(0.351)	(0.213)
Premium Panel (0,1)	0.542	0.714
	(0.498)	(0.454)
Premium Plus Panel (0,1)	0.0153	0.0357
	(0.123)	(0.187)
Microinverter (0,1)	0.792	0.881
	(0.406)	(0.326)
Permanent Seller (0,1)	0.822	0.976
	(0.382)	(0.153)
Observations	4273	

Across the 10,545 auctions, 40,575 bids were submitted during the two-year sample. Table 2 provides summary statistics of the submitted bid characteristics. Panel A shows that the mean price bid was $3.61 per watt,[24] additionally, 38% of bids included premium panels, and 76% of bids included microinverters. Panel B shows bid characteristics separately for selected and non-selected bids for a single market, Southern California in the first half of 2016. Not-surprisingly,

[23] Some installers participate in auctions in multiple states.

[24] This is the average gross bid price and does not include the 30% ITC or any state rebates.

quotes that are selected are lower in price on average. However, buyers also seem to care about other attributes. Selected quotes are more likely to be from permanent installers, more likely to have premium or premium plus panels, and are more likely to include microinverters. Indeed, many buyers chose a quote that was not the lowest priced.[25] Because consumers appear to value non-price features of solar systems, I cannot model the consumer's decision as a first-price auction. A solar PV installation is a differentiated product, and therefore firms' optimal bidding strategy will depend on customers' preferences for different attributes of the installation.

One feature that is apparent in the EnergySage price quotes that is also salient in other solar PV sales data is the presence of significant price variation. System prices vary widely across both space and time.[26] Figure 12 in the appendix displays a kernel density plots of price quotes for each state. In Southern California, price quotes range from as low as $3/watt to over $4/watt. The average system size is 7,155 watts; so this can mean differences in overall system prices of over $7,000. These price differences likely result from differences in costs across installers, differences in unobserved installer quality, differences in unobserved project characteristics (i.e., different roof types), as well as differences in the amount of competition (number of bids) for specific projects. In the next subsection, I investigate the relationship between the amount of competition and both market prices and quantities.

2.2 Effects of Competition on Quantities and Prices

Before moving to the structural model, I first show descriptive evidence of the impacts of competition on both solar PV adoption (quantity sold) and prices. To begin, I discuss the relationship between the number of bidders on a project and the probability that a bid is selected. Next, I illustrate the relationship between competition and prices by comparing purchase prices of solar systems on EnergySage to prices paid for comparable installations off of the platform in the same geographic area.

2.2.1 Number of Bids and the Solar Adoption Decision

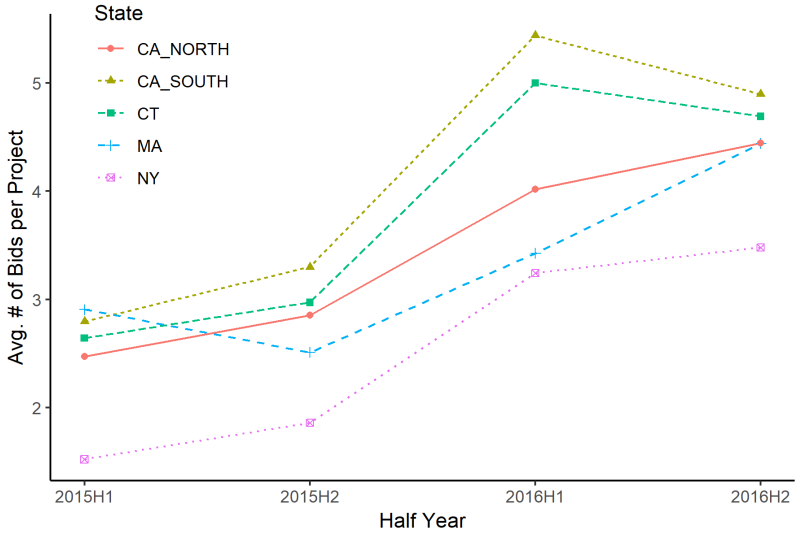
The competitive landscape for auctions held through the EnergySage platform evolved throughout the sample as shown in Figure 2. Over time, projects receive more bids on average. Also, there is substantial variation in the number of bids that projects receive across states. Auctions in Southern California and Connecticut were the most competitive, averaging around five bids per project in 2016.

The number of bids could impact the total solar PV sales through two channels. First, increasing competition should drive down prices and, therefore, increase the equilibrium quantity purchased. Secondly, if sellers are differentiated, additional bids can provide consumers with

[25] This pattern is also apparent in almost any given market (not just Southern CA in 2016H1). However, the pattern is not as clear when looking at selected versus non-selected bids for the full sample. This is because average prices and attributes are changing over time and so is the overall probability of purchase.

[26] See Tables 8 and 9 in the appendix for the project and bid summary statistics by state.

Figure 2: Average Number of Bids per Project Over Time



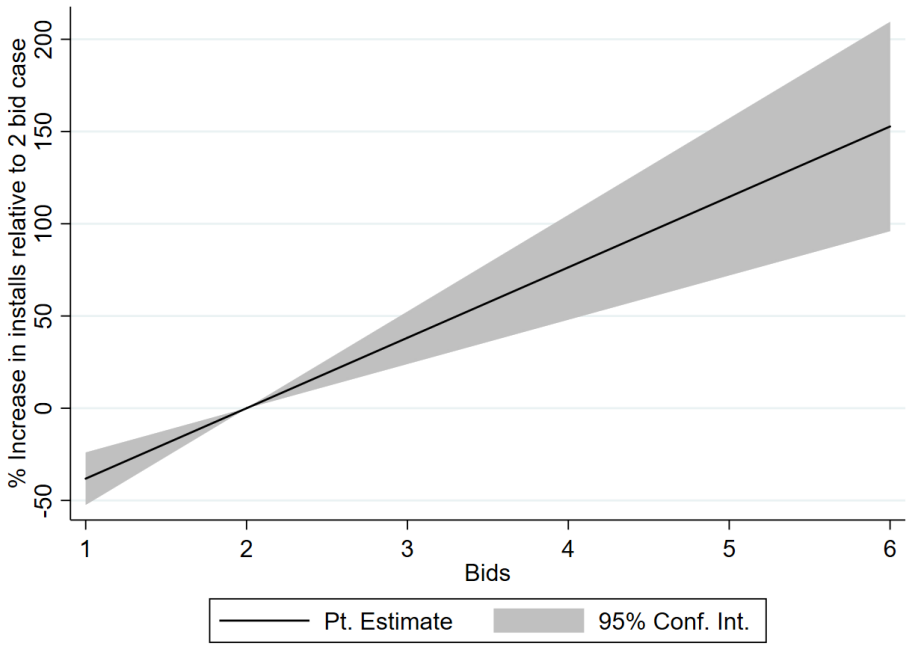
more desirable options. For example, if installers differ regarding customer service, warranties, experience, or hardware offerings then adding bids can increase the likelihood that buyers are aware of higher-quality and better-matched installers for their specific project. I estimate the following regression to better measure the effect of increased auction participation on consumers' adoption choice:

$$\mathbb{1}[Purchase]_i = \alpha Bids_i + \nu_{st} + \lambda_z + \varepsilon_i \tag{1}$$

Where $\mathbb{1}[Purchase]_i$ is an indicator function that takes the value of one if consumer i purchased a solar PV system and otherwise equals zero. $Bids_i$ is the number of bids that were submitted for i's project. Finally, ν_{st} and λ_z are state-time, and ZIP code fixed effects respectively, and ε_i is an idiosyncratic error. This regression exploits variation in the number of bids within a ZIP code controlling for demand shocks affecting each state-quarter. The regression model provides predictions about the expected probability that each household purchases a solar system conditional on the number of bids received. EnergySage's auction close rates (purchase probabilities) are proprietary and cannot be reported directly. Instead of reporting conditional purchase probabilities, I use the regression estimates from equation 1 to calculate how expected purchase probabilities change *relative* to a case where the buyer receives two bids. More specifically, I use the estimates to obtain the percentage increase in adoptions from receiving n bids as $\frac{\mathbb{E}[1[Purchase_i]|bids=n]}{\mathbb{E}[1[Purchase_i]|bids=2]} * 100$.[27] Two bids was chosen as the benchmark because a 2015 installer survey found that the median shopper in the residential solar market obtained two quotes before making a purchase decision.

[27]For example, if the regression predicted that projects with two bids had a 20% purchase probability and three-bid projects had a 30% purchase probability then the increase in adoption from receiving three bids would be 50%

Figure 3: Effect of Additional Bids on Solar PV Purchases



Notes: The graph plots the percentage change in expected sales relative to the case where buyers obtain two bids. The increase in probability is calculated using estimates from regression equation 1. Delta method standard errors are used to form the 95% confidence interval.

Figure 3 shows that each additional bid is associated with a 38% increase in the expected number of sales relative to the case that buyers obtain two bids. This effect is relatively large indicating that doubling the number of bids nearly doubles the number of solar system purchases. The results are quite robust to different sets of fixed effects as shown in Table 11 in the appendix.

Can lower bid prices explain the link between the number of bids and sales? Figure 13 in the appendix shows that mean bid prices consistently fell through the sample period as auctions became more competitive. Therefore, it's plausible that lower bid prices are driving the relationship between bid quantity and purchase probability. However, the decline in prices over time was also due to falling hardware input costs, so it is not immediately clear how much of the price decline can be attributed to changes in competition. In the next subsection, I provide additional evidence about the relationship between competition and prices by comparing solar PV prices on the EnergySage platform compared to prices paid by buyers off of the platform.

2.2.2 Prices on the Platform vs. Prices off the Platform

The online platform gives consumers access to more potential installers. Likewise, sellers are aware that customers on the platform likely see more competing price bids, which could cause themselves to place lower bid prices. To investigate the effect of the platform on equilibrium

prices paid, I collect additional data from the Lawrence Berkeley National Lab and the National Renewable Energy Laboratory's Open PV dataset. The Open PV Project is a collaborative effort between government, industry, and the public to assemble data on Solar PV installations across the United States. The data comes from solar incentive programs, utilities, installers, and other groups.

To compare equilibrium prices paid on the platform to prices off the platform, I append the Open PV data for California, Connecticut, Massachusetts, and New York with the EnergySage data for selected bids (EnergySage bids that were chosen by a consumer).[28] I estimate the following model:

$$P_i = \alpha \mathbb{1}[Platform]_i + \beta X_i + \nu_{st} + \lambda_z + \varepsilon_i \tag{2}$$

Where P_i is the price paid for system i in dollars per watt (or its logarithm), $\mathbb{1}[Platform]_i$ is an indicator function that takes the value of one if the system was purchased via EnergySage and otherwise equals zero, X_i is a vector of system characteristics such as system size, panel quality dummies, and a microinverter dummy. Finally, ν_{st} and λ_z are state-time, and ZIP code fixed effects respectively, and ε_i is an idiosyncratic error.

Table 3: Effect of the EnergySage Platform on Installed Prices ($/watt)

	(1) Price	(2) Price	(3) ln(Price)	(4) ln(Price)
Platform (0,1)	-0.767***	-0.666***	-0.169***	-0.146***
	(0.0366)	(0.0350)	(0.00793)	(0.00754)
Controls	No	Yes	No	Yes
State-Time FE	Yes	Yes	Yes	Yes
Zipcode FE	Yes	Yes	Yes	Yes
N	135341	135341	135341	135341
R^2	0.163	0.238	0.159	0.243

Notes: Controls include system size, panel quality (rated by EnergySage), and a dummy for if the system includes a microinverter or power optimizer. All standard errors are listed in parenthesis. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

The regression results, displayed in Table 3, indicate a strong correlation between the use of the platform and prices paid. In column 2, we see that average prices paid were over $ 0.66/watt lower on EnergySage than comparable systems off the platform. Column 4 gives estimates of the same model with the natural logarithm of price as the outcome variable; the results suggest that systems on EnergySage were sold at 15% lower prices.

There are several possible explanations for the difference in prices between EnergySage and

[28] I only consider purchased residential systems between 2KW to 20 KW in size in both data sets. Also, it is likely that each of the EnergySage observations will also appear in the Open PV data. I attempt to deal with this issue by using a matching procedure to pair each observation in the EnergySage data with an observation that has similar observables in the Open PV data (same ZIP code, same time-period, similar price, similar size) and dropping the redundant observations. Since the EnergySage dataset is small relative to the Open PV dataset, the regression results are similar even if I do not drop the redundant observations.

off-line transactions. For one, the degree of competition on EnergySage is likely to be much higher. The median buyer outside the platform obtained two price quotes (Aggarwal, 2015), in contrast to the median buyer on EnergySage who collected four bids. Therefore, installers should be able to exercise market power and charge higher prices off of the platform. However, the difference in prices could also be caused by variation in demand between EnergySage buyers and other buyers, perhaps EnergySage users are more price-elastic than the typical buyer. Another possibility is that EnergySage attracts installers that have lower costs. While the above results are suggestive, it is difficult to isolate the effect of competition on prices precisely.

In Section 3, I develop a structural model to measure the direct impact of competition on equilibrium prices. Formulating an explicit model of the supplier bidding makes it possible to compute how prices would change as competition changes while holding consumer preferences and seller costs fixed. The model of consumer choice also enables us to decompose the importance of bid price relative to non-price factors that drive the purchase decision. Finally, by combining the demand and supply models, we can assess welfare under different counterfactual market environments and policies.

3 Structural Model

Each buyer i seeks to procure installation services for a single indivisible project using a multi-attribute auction. Throughout the paper, i references both an individual buyer and their respective project. Buyer i's project is distinguished by its project type τ_i, which depends on the project's location, size, and time period. For each project of type τ, there is a set $\mathcal{N}(\tau)$ of potential sellers that choose whether or not to submit a bid for the project.

Each seller j is differentiated by their quality group index which belongs to a discrete set of $k+1$ values, $\mathcal{Q} = \{t, q^1, .., q^k\}$. A seller that only submits a limited number of total bids is considered to be a "transient" seller and belongs to the quality group t. All other sellers that submitted a sufficiently large number of total bids and made at least one sale are considered to be "permanent" sellers and belong to one of k quality groups $q^1, .., q^k$, with q^k denoting the highest quality. Each seller's quality group index is observable to both the buyer and the other potential sellers. If a seller chooses to participate in the auction for project i he then also selects a price bid B_{ij}. Each seller's bid is also characterized by a vector of non-price characteristics $\mathbf{x}_{ij}$ such as panel quality and inverter type. $\mathbf{x}_{ij}$ is allowed to vary across projects for a given seller.

3.1 Demand: Buyer's Choice Problem

The allocation rule in a multi-attribute auction comes from the buyer's choice problem. Let $\mathcal{K}_i \subset \mathcal{N}(\tau_i)$ be the set of sellers that decide to participate in the auction for project i. Buyer i then chooses between the project bids and an unspecified outside option (k^0) to maximize their utility. Household i's utility from selecting option j is given by:

$$u_{ij} = \alpha B_{ij} + \mathbf{x}'_{ij}\beta + \delta_\tau + \zeta_{ig} + (1-\lambda)\varepsilon_{ij} \tag{3}$$

Here B_{ij} is the bid price for option j, $\mathbf{x}_{ij}$ contains observable characteristics of the system, such as the panel brand quality, inverter type, the size of the system, and the installer quality group index. δ_τ is a demand shifter for projects of type τ that allows utility for all of the "inside options" to vary depending on location, time-period, and project size. ε_{ij} is an independent and identically distributed random term that is assumed to follow a type-one extreme value distribution. ζ_{ig} is also an idiosyncratic term but is assumed to be constant for each buyer across all the "inside options". ζ_{ig} follows the unique distribution distributed such that $\zeta_{ig} + (1-\lambda)\varepsilon_{ij}$ is also an extreme value random variable. This utility specification gives rise to the nested logit model (Cardell, 1997). The nested logit model allows for more flexible substitution patterns in comparison to the standard logit model because it accommodates correlation in preferences for products within pre-specified groups. Here, I specify one group to be the "outside option," and the other group to contain all of the project bids. Some households may register for the platform just out of curiosity about solar PV prices and may not be serious about making a purchase. Likewise, there may be customers that are very adamant about buying a solar PV system. Therefore, these consumers would be unlikely to select the outside option even if some of the options in their choice set were removed. The nested logit model allows for these types of individuals. As λ approaches zero, each buyer has no correlation in preferences for each "inside option", and the model reduces to the standard logit model. As λ goes to one, the random component of buyers' preferences for each "inside option" become perfectly correlated. Finally, the overall level of utility is not identified, so I normalize the utility of the outside option to equal zero plus an error term, this normalization is standard in the literature.

Given the utility specification in equation 3, the probability that household i chooses option j is:

$$Prob_{ij} = \frac{exp\left(\frac{\alpha B_{ij} + \mathbf{x}'_{ij}\beta + \delta_\tau}{1-\lambda}\right)\left(\sum_{k \neq k^0} exp\left(\frac{\alpha B_{ik} + \mathbf{x}'_{ik}\beta + \delta_\tau}{1-\lambda}\right)\right)^{-\lambda}}{1 + \left(\sum_{k \neq k^0} exp\left(\frac{\alpha B_{ik} + \mathbf{x}'_{ik}\beta + \delta_\tau}{1-\lambda}\right)\right)^{1-\lambda}} \tag{4}$$

Where the sum is over all bids $k \in \mathcal{K}_i$ in the individual's choice set except the outside good k^0.

3.2 Supply: Seller Bidding and Participation Decision

The supply-side model has several fundamental differences from a standard differentiated products model. First, firms must make an explicit decision about whether to submit a price quote to each potential buyer. Second, sellers do not have information about exactly how many competing suppliers will make bids to the customer. Moreover, the suppliers do not have perfect information about the identity and characteristics of the competitors they will face, nor about the price quotes those competitors will submit. Firms cannot see the exact identity of com-

petitors that offer bids for a particular project. However, they can observe the total number of bids that were submitted to an auction ex-post. They also see which other firms participate on the platform in their area. Therefore, it is reasonable to assume that the suppliers know the distribution of possible competition they are likely to face for a given project. The model of firm behavior outlined below accounts for these features of the online platform.

I model suppliers bidding behavior as a two-stage process. In the first stage, each potential bidder $j \in \mathcal{N}(\tau_i)$ must decide whether or not to enter the auction for the project i. At the time of entry, firms do not know their exact marginal cost of completing the project, but they know the distribution of possible costs they could incur. They also know the probabilities of each of their opponents entering the auction, the characteristics of those opponents and the distribution of possible prices those opponents would submit conditional on entry. Additionally, they know the mean utility of the buyer (but not the random component of utility).[29] Therefore, each firm can form an expectation about their profits conditional on the decision to enter the auction. If firm j decides to enter the auction for the project, they incur a bid preparation cost $S(\tau_i, \mathcal{Q}_j) + \eta_{ij}$, where $\eta \sim Lognormal(0, \sigma^2(\mathcal{Q}_j))$. The bid preparation cost contains a deterministic component that depends on the project type and the seller's quality group, and a random component that's variance depends on the seller's quality group. I assume that the random component is i.i.d. across projects and firms and is private information of each potential bidder. If a firm decides to enter auction i they learn the non-price characteristics of their bid $\mathbf{x}_{ij}$ and the marginal cost of completing the project c_{ij}.

To make the model empirically tractable, I assume that the non-price characteristics, $\mathbf{x}_{ij}$ including the size of the project bid are not strategic choices of bidders. This assumption means that firms are not choosing non-price characteristics such as panel quality and inverter type strategically when placing a bid. While this assumption is made primarily for tractability, the assumption also finds support in the data. For example, installers will typically use the same equipment for many consecutive projects. They may change module brands occasionally, but the hardware available to them to complete a given project is likely predetermined by their existing inventory. The practical interpretation of this assumption is that sellers need to check their existing product stock (which is predetermined) before knowing the exact non-price characteristics of their bid. They learn the non-price components of the bid by incurring the bid preparation costs.

The second part of the assumption means that the size of the project is pre-determined by the buyer and is not a choice variable for the seller. On the EnergySage platform, buyers submit a monthly electricity bill and sellers then choose the exact size of the system. In practice, the system size quotes are very similar across installers for a given project because almost all sellers size the system so that it will cover 100% of the buyer's annual electricity use. A regression of installers' system size quote on project dummies has an R^2 of 0.87, which means almost all of the system size choice can be explained by the project characteristics and is not likely a critical

[29] In particular, I assume that sellers know all of the parameters of the buyer's utility function, α, β, δ, and λ.

strategic variable.[30] Analogous to the other non-price characteristics, each firm learns the size of their project bid after incurring the bid preparation cost. Sellers know consumers' taste for size and can optimize their bid price based on the size draw that they receive.

When firms make their entry decision they do not know their marginal cost or non-price characteristics, but they do know the joint distribution from which their marginal cost and non-price characteristics will be drawn, $F_{CX|\mathcal{Q}_j,\tau_i}(c, \mathbf{x}|\mathcal{Q}_j, \tau_i)$. The distribution depends on both the sellers quality group and the project type. After the firms make their entry decisions in stage one, each firm's marginal cost and non-price characteristics are drawn from $F_{CX|\mathcal{Q}_j,\tau_i}$ and the installer then decides on a price bid during the second stage.

3.2.1 Sellers' Bid Pricing Problem

It will be helpful to first consider the firm's problem in the second stage after marginal costs and non-price characteristics are realized. Conditional on entering an auction, the firm j solves the following problem when setting a bid price for project i:

$$\max_{B_{ij}} \quad [B_{ij} - c_{ij}]\mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i) \tag{5}$$

Where B_{ij} is firm j's price bid, c_{ij} is firm j's marginal cost, and $\mathbf{x}_{ij}$ are the firm's non-price characteristics for project i. $\mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i)$ is the equilibrium probability of winning the auction conditional on placing placing a bid price of B_{ij}, having non-price characteristics $\mathbf{x}_{ij}$, and belonging to quality group $\mathcal{Q}_j$. The equilibrium expected probability of being selected is a function of the type of project τ_i. We work with expected probabilities because the seller does not know exactly which competitors he will face nor the bids of those competitors.

When formulating firms expectations, I assume that all sellers submit bids simultaneously. Therefore, the installers do not know the exact number of bidders they will be competing against nor the identities of their competitors. Thus, firms' expectations (about the probability of winning) will only be a function of the project type, conditional on the price and non-price characteristics of their bid. In practice, firms on the platform submit bids at slightly different times. Although the identities of competing bidders are not visible to auction participants, firms can see how many bids have already been submitted for a given auction. Therefore, it is possible that firms could update their expectations based on the number of bids that have already been submitted. The assumption of simultaneous bidding is made primarily to simplify computation in the empirical exercise. However, I provide evidence that the assumption is a reasonable approximation of firms' behavior. In Appendix Table 12, I regress bid price on the order that a bid was submitted, controlling for the total number of bids, installer fixed effects, state fixed effects, and time-period fixed effects. The coefficient on "order of bid" is small and not significant. This suggests firms are not making significant changes in bidding strategy based on the order they

[30] In theory, it would be possible to include size as an additional choice variable of the seller. Then the seller would have two first-order conditions for an optimal bid (size and price). However, this would likely lead to multiple equilibria for sellers in the bidding stage as numerous price-size pairs may satisfy the first-order-conditions.

submitted a bid.

Under the assumption of simultaneous bidding, a firm's expected probability of winning $\mathcal{P}_{ij}$ can be expanded as follows:

$$\begin{aligned}\mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i) = \mathbb{E}[Prob_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \boldsymbol{\mathcal{Q}}_{-j}|\tau_i)] = \\ \int Prob_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \boldsymbol{\mathcal{Q}}_{-j}|\tau_i) \cdot dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \boldsymbol{\mathcal{Q}}_{-j}|\tau_i)\end{aligned} \tag{6}$$

Recall that $Prob_{ij}$ is the probability that buyer i selects firm j's bid conditional on realized set of competitors submitting a vector of price bids $\mathbf{B}_{i,-j}$, having a stacked vector of non-price characteristics $\mathbf{X}_{i,-j}$, and having quality indices $\boldsymbol{\mathcal{Q}}_{-j}$. G represents the joint distribution function of $\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}$, and $\boldsymbol{\mathcal{Q}}_{-j}$ occurring in equilibrium, conditional on the project being of type τ_i. Since each firm's entry draw and marginal cost draw is assumed to be i.i.d., we can express dG as the product of the probabilities that each competing firm l decides to enter the auction and then bids B_{il} and has non-price characteristics x_{il}.

I define the optimal bid function as $B^*_{il}(c_{il}|\mathbf{x}_{il}, \mathcal{Q}_l, \tau_i)$ and $H(\mathcal{Q}_l, \tau_i)$ as the probability that a potential seller l that is of quality $\mathcal{Q}_l$ enters an auction of type τ_i. Then we have:

$$dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \boldsymbol{\mathcal{Q}}_{-j}|\tau_i) = \prod_{l \in \mathcal{N}(\tau_i)\setminus\{j\}} H(\mathcal{Q}_l, \tau_i) \cdot dF_{CX|\mathcal{Q}_l,\tau_i}\big(B^{*-1}(B_{il}|\mathbf{x}_{il}, \mathcal{Q}_l, \tau_i), \mathbf{x}_{il}|\mathcal{Q}_l, \tau_i\big) \tag{7}$$

Where B^{*-1} represents the inverse bid function. The expression inside the product is the probability that firm l enters the auction multiplied by the probability that firm l bids B_{il} and has non-price characteristics x_{il}.

Firm i's first-order condition for an optimal bid is given by:

$$(B_{ij} - c_{ij})\frac{\partial \mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i)}{\partial B_{ij}} + \mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i) = 0 \tag{8}$$

Given a vector of non-price characteristics, the optimal bid function $B^*_{il}(c_{il}|\mathbf{x}_{il}, \mathcal{Q}_l, \tau_i)$ is defined implicitly by equation 8.

3.2.2 Sellers' Participation Decision

Now consider the firm's decision of whether or not to enter an auction. Each firm will enter if the expected marginal profits conditional on entering are larger than the fixed cost of bid preparation $S(\tau_i, \mathcal{Q}_j) + \eta_{ij}$. When firm j decides to enter auction i, they only know the project type and their own quality index, and their private entry cost draw. Firm j's expected profits conditional on entering the auction for project i can be expressed as follows:

$$\mathbb{E}[\pi_{ij}|\mathcal{Q}_j,\tau_i] = \int \Big[(B^*_{ij}(c_{ij}|\mathbf{x}_{ij},\mathcal{Q}_j,\tau_i) - c_{ij})\mathcal{P}_{ij}(B^*_{ij},\mathbf{x}_{ij},\mathcal{Q}_j|\tau_i)\Big] dF_{CX|\mathcal{Q}_j,\tau_i}(c_{ij},\mathbf{x}_{ij}|\mathcal{Q}_j,\tau_i) \tag{9}$$

Recall that $F_{CX|\mathcal{Q}_j,\tau_i}(c,\mathbf{x}|\mathcal{Q}_j,\tau_i)$ is the joint distribution of unknown marginal shocks and non-price characteristics whose realization is not known to the firm at the time of entry. Therefore, the firm will enter the auction as long as:

$$\mathbb{E}[\pi_{ij}|\mathcal{Q}_j,\tau_i] \geq S(\tau_i,\mathcal{Q}_j) + \eta_{ij} \tag{10}$$

Under the assumption that η_{ij} follows a lognormal distribution, the probability that firm j enters the auction for project i is:

$$H(\mathcal{Q}_j,\tau_i) = \Phi\left(\frac{ln\Big(\mathbb{E}[\pi_{ij}|\mathcal{Q}_j,\tau_i]\Big) - S(\tau_i,\mathcal{Q}_j)}{\sigma(\mathcal{Q}_j)}\right) \tag{11}$$

Where Φ represents the cumulative distribution function for a standard normal random variable.

Summary: Timing of the game

1. A potential buyer i initiates a multi-attribute auction by announcing the project type τ_i to all potential entrants $\mathcal{N}(\tau_i)$.

2. Each potential seller $j \in \mathcal{N}(\tau_i)$ receives a private entry cost shock η_{ij}. Each potential entrant then compares their entry cost $S(\tau_i,\mathcal{Q}_j) + \eta_{ij}$ to the expected marginal profit conditional on entering the auction $\mathbb{E}[\pi_{ij}|\tau_i,\mathcal{Q}_j]$. Each potential bidder chooses to enter if and only if expected marginal profits are larger than their entry cost.

3. Each seller that enters auction i receives a private marginal cost draw c_{ij} and also learns their non-price characteristics $\mathbf{x}_{ij}$. Sellers do not observe which other competitors have entered the auction. Each entrant then chooses a bid price B_{ij}.

4. Buyer i chooses from each of the project bids or the outside option.

3.3 Equilibrium

For each seller j, a strategy consists of two functions: a participation strategy $\mathcal{Q} \times \tau \times \mathbb{R}_+ \to \{0,1\}$, and a bidding strategy $\mathcal{Q} \times \tau \times \mathbf{x} \times \mathbb{R}_+ \to \mathbb{R}_+$. Specifically, sellers use information about the project type, their quality group, and their entry cost shock to determine the binary choice of whether or not to enter. In the bidding stage, firms consider the project type, their quality group, their marginal cost draw, and their non-price characteristics to form a price bid. I follow the convention in the literature by focusing on type-symmetric pure strategy Bayesian equilibrium (Krasnokutskaya et al., 2019). That is, all sellers of the same quality index use

the same participation strategy in equilibrium, and all sellers of the same quality index and same non-price-characteristics use the same bidding strategy in equilibrium. An equilibrium in the participation stage is a strategy profile such that all sellers satisfy Inequality 10, given the strategies of other firms. An equilibrium in the bidding stage requires that all firms satisfy equation 8 given the other installer's strategies. Krasnokutskaya et al. (2019) prove the existence of a type-symmetric pure strategy Bayesian equilibrium of this game. However, there is no guarantee of a unique equilibrium in the participation stage. Multiple equilibria will not impact the parameter estimates since I am not solving the model in estimation. However, multiple equilibria may affect the counterfactual analysis. To address this issue, I resolve the model using different starting points and do not find any indication of multiple equilibria. The next subsection describes the estimation procedure in detail.

3.4 Estimation and Implementation Details

I estimate the structural parameters in three steps. First, I solve for the demand parameters using a two-stage grouped fixed effects approach. Second, I use the estimated demand parameters to simulate firms' first-order conditions for each bid in the data and recover bid-specific markups. Finally, I use the estimates from the first two steps to calculate each bidders' expected marginal profit from entering each auction and estimate the entry cost parameters using observed entry decisions. I discuss the details of each step in the following subsections.

3.4.1 Demand Estimation and Seller Quality Groups

I estimate the demand parameters via maximum likelihood. Each buyer's utility is allowed to depend on bid price ($/watt), panel quality (dummies for premium and premium plus modules), inverter type (microinverter dummy), the capacity of the system, state fixed effects and quarter fixed effects. Here, the location and time fixed effects are demand shifters for all the inside options (δ_t). The price that enters the buyer's utility is equal to 70% of the installer's quoted price to account for the 30% investment tax credit (ITC).[31] I also allow for heterogeneous seller quality by including installer fixed effects for each of the 65 installers that placed over 100 bids and made at least one sale.

Let M be the total number potential buyers in the sample, then the log likelihood function is:

[31]None of the states besides New York had changes in subsidy rates during the sample period so any state rebate should be accounted for by the state fixed effects. In a later section, I show that the results are robust to adding explicit controls for New York's state-level incentives.

$$LL(\alpha, \beta, \delta, \lambda) = \sum_{i}^{M} \sum_{j \in \mathcal{K}_i} \mathbb{1}[i \; choose \; j] \cdot ln(Prob_{ij}) \tag{12}$$

$$= \sum_{i}^{M} \sum_{j \in \mathcal{K}_i} \mathbb{1}[i \; choose \; j] \cdot ln \left(\frac{exp\left(\frac{\alpha B_{ij} + \mathbf{x}'_{ij}\beta + \delta_\tau}{1-\lambda}\right)\left(\sum_{k \neq k^0} exp\left(\frac{\alpha B_{ik} + \mathbf{x}'_{ik}\beta + \delta_\tau}{1-\lambda}\right)\right)^{-\lambda}}{1 + \left(\sum_{k \neq k^0} exp\left(\frac{\alpha B_{ik} + \mathbf{x}'_{ik}\beta + \delta_\tau}{1-\lambda}\right)\right)^{1-\lambda}} \right)$$

After solving for the maximum likelihood estimates, I sort each of the sellers into quality groups. If a seller submitted less than 100 total bids during the sample, they are placed in the "transient" seller group. If a seller made more than 100 bids during the sample, they are considered a "permanent" seller and are placed into one of ten quality groups based on their fixed effects. In particular, I sort each firm into a group based off of the decile of its fixed effect estimate. For instance, if we ordered each firm by the fixed effect estimates and firm j was in the lth decile, the seller would be placed into l, "Permanent - Quality Group l".

Each seller's fixed effect will be higher if he systematically wins auctions more often than transient sellers conditional on price and non-price characteristics of his bids. I interpret sellers that consistently perform better (conditional on price bids) as having higher unobserved quality. Unobserved quality will include things like star rating, installation experience, customer service, and warranty services offered.

After sorting the installers into the eleven quality groups, I re-estimate the demand model but now including fixed effects for each seller quality group in the buyers' utility function instead of the full set of firm fixed effects. Figure 4 summarizes the non-price characteristics of each installer's bid.

Figure 4: Non-Price Characteristics of Bids

$$\overset{\text{Quality Group}}{\begin{bmatrix} \text{Transient} & \text{Quality Grp. 3} \\ \text{Quality Grp. 1} & \ldots \\ \text{Quality Grp. 2} & \text{Quality Grp. 10} \end{bmatrix}} \times \overset{\text{Panel Type}}{\begin{bmatrix} \text{Standard} \\ \text{Premium} \\ \text{Premium Plus} \end{bmatrix}} \times \overset{\text{Inverter}}{\begin{bmatrix} \text{String (Standard)} \\ \text{Microinverter} \end{bmatrix}}$$

Notes: When each seller places a bid in an auction, their quote will also include a vector of seller characteristics that is observable to the buyer. The panel quality and inverter type are observed in the data and the seller's quality group is determined by sorting firms by the first-stage installer fixed effects, The system size is also included in the non-price characteristics of the bid.

There are several reasons for sorting the sellers into quality groups. For one, it allows me to

group several installers to calculate each seller's expected marginal profit from entering a given auction. Computing expected marginal profits is required to estimate the entry cost parameters. If I did not sort the sellers into groups, then every permanent seller would have a different expected marginal profit from entering a given auction (because they each would have different probabilities of winning conditional on entry). This is problematic because some sellers may only participate in a few auctions of a particular type (i.e., Installer X may just bid on three large projects in Connecticut in 2016Q1) which means that the estimated marginal profit would have a very high variance due to being calculated using only a handful of observations. Similarly, some of the seller fixed effects may be estimated imprecisely, which would lead to high variability in the markup estimates. Finally, the quality groups allow for a more straightforward economic interpretation of results compared to listing all of the individual firm fixed effects estimates. In a later section, I show that the demand estimates after grouping the sellers are very similar to the estimates with the full set of fixed effects. I also discuss different estimation approaches such as using a K-means clustering algorithm to sort the sellers into groups and an iterated estimator that updates the quality groups at each stage until convergence.

3.4.2 Inferring Markups and Marginal Costs

In the next step, I recover a markup estimate for each bid in the data. To do so, I use the final demand estimates to form each firm's first-order condition for an optimal bid from equation 8. Notice that the FOC does not have a closed-form since it contains two expectations $\frac{\partial \mathcal{P}_{ij}(B_{ij},\mathbf{x}_{ij},\mathcal{Q}_j|\tau_i)}{\partial B_{ij}}$ and $\mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathcal{Q}_j|\tau_i)$. Therefore, we have to integrate the firm's probability of winning over different realizations of competitor sets and competitor bid prices that are unknown to the installer at the time of bidding. To understand the evaluation of each firm's FOC, consider an auction i that is of type τ_i.

1. First, obtain non-parametric estimates of the entry probabilities for each project type and each seller quality group. For example, the probability that "Permanent - Quality Group 7" sellers enter New York auctions in 2015Q2. This estimate is just the ratio of auctions entered divided by total auctions of that type. I assume a seller is a potential entrant for auction i if they entered at least one auction of type τ_i.

2. Next, use the probabilities from the previous step to simulate the entry decisions into auction i for each potential entrant in $\mathcal{N}(\tau_i)$.

3. Draw price bids and non-price characteristics for each of the simulated entrants using the empirical joint distribution of bids and non-price characteristics in the data. For example, if a type q^l seller enters a simulated auction of type τ_i; then randomly draw a bid (both bid price and non-price characteristics together) from the pool of all bids placed by type q^l sellers in auctions of type τ_i.

4. Evaluate the choice probabilities $Prob_{ij}$ and demand semi-elasticities $\frac{\partial Prob_{ij}}{\partial B_{ij}}$ inside the integrals given the bid prices and the competitors observed characteristics.

5. Repeat this process S times[32] and take the average of all the simulated choice probabilities, and simulated demand semi-elasticities to obtain estimates for the two expectations. Let s denote the simulation iteration, then the expressions are:

$$\widehat{\mathcal{P}_{ij}} = \frac{1}{S}\sum_{s=1}^{S} Prob_{ij}^{s}, \qquad \widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}} = \frac{1}{S}\sum_{s=1}^{S} \frac{\partial Prob_{ij}^{s}}{\partial B_{ij}} \tag{13}$$

6. Finally, use the average choice probabilities, and average demand semi-elasticities to calculate the markup portion of each bid. The markup term for firm j in auction i is equal to $-\frac{\widehat{\mathcal{P}_{ij}}}{\widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}}}$. Once we have an estimate of the markup term, the firm's FOC provides a one-to-one mapping that we can use to recover the marginal cost of each project in the data:

$$\widehat{c_{ij}} = B_{ij} + \frac{\widehat{\mathcal{P}_{ij}}}{\widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}}} \tag{14}$$

This process allows me to infer a project-specific marginal cost for every bid in the data. I then use the estimated marginal costs to form a non-parametric cost distribution for each seller-project-type pair (e.g., a cost distribution "Permanent - Quality Group 7" sellers in New York 2015Q2 auctions).

The choice of project categories is crucial for obtaining credible estimates of markups. Defining project type categories exemplifies a trade-off between bias and variance. On the one hand, defining too few project types can bias markup estimates if projects are heterogeneous. For example, installations in New York will be different ex-ante than projects in Northern California because of differences in labor costs, permitting requirements, and differences in the set of possible competing bidders. Likewise, a New York project in 2015 will be different from a New York project in 2016 because of differences in hardware input costs, differences in consumer preferences, and differences in potential bidders. For this reason, we would not want to use bids placed in Northern California 2015Q1 when simulating a New York 2016Q3 auction because New York installers are not likely to be using these bids to form their expectations. However, if I define too many project categories, (i.e., each state-week has its own category), then markup estimates for each bid will have high variance because there is only a handful of projects to use to simulate realizations of each auction. Figure 5 displays the project type definitions used in the primary analysis. I specify 80 different project types determined by the location, time period, and the size of the project. I allow for project type to vary across the

[32] I simulate 1000 iterations of each auction type.

quarter of sample within a state to account for the rapidly changing competitive environment on the platform and evolving input costs over time. In later sections, I discuss the robustness of the results to changes in the definition of project types.

Figure 5: Project Types

$$\begin{array}{ccccc} \text{Location} & & \text{Time Period} & & \text{Project Size} \\ \begin{bmatrix} \text{CA North} & \text{CT} & \text{NY} \\ \text{CA South} & \text{MA} & \end{bmatrix} & \times & \begin{bmatrix} \text{2015Q1} & \ldots \\ \text{2015Q2} & \text{2016Q4} \end{bmatrix} & \times & \begin{bmatrix} \text{Small } (\leq 6.5 \text{ KW}) \\ \text{Large } (> 6.5 \text{ KW}) \end{bmatrix} \end{array}$$

Notes: Each auction is categorized based of the location, time period, project size. There are $5 \times 8 \times 2 = 80$ project types. For example, {CT, 2016 Q1, Large} defines a single project type.

3.4.3 Entry Cost Parameters

In the final step. I use the estimated marginal costs to form each firms' *pre-entry* expected marginal profit from entering an auction i. For each bid in the data, I can calculate the firms' *post-entry* expected profit (before the buyer makes a choice) using the bid price, marginal cost, and probability of winning. The *post-entry* expected profit for seller j in auction i is equal to $(B_{ij} - \widehat{c_{ij}}) \cdot Prob_{ij}$. To calculate a seller's *pre-entry* expected profit $\widehat{\mathbb{E}[\pi_{ij}]}$ from entering an auction i; I take the average over all of the *post-entry* expected profits that are realized by the seller's quality group $\mathcal{Q}_j$ for projects of type τ_i. Define $N(\tau_i, \mathcal{Q}_j)$ as the total number of bids placed by type $\mathcal{Q}_j$ sellers in auctions of type τ_i, then *pre-entry* expected profit are estimated as:

$$\widehat{\mathbb{E}[\pi_{ij}]} = \frac{1}{N(\tau_i, \mathcal{Q}_j)} \sum_{i \in \tau_i} \sum_{j \in \mathcal{Q}_j} (B_{ij} - \widehat{c_{ij}}) \cdot Prob_{ij} \tag{15}$$

Here, I use $\sum_{i \in \tau_i}$ to mean the sum over all auctions of type τ_i and $\sum_{j \in \mathcal{Q}_j}$ to indicate the sum over all bids submitted by sellers of type $\mathcal{Q}_j$. Next, I use the *pre-entry* expected profits $\widehat{\mathbb{E}[\pi_{ij}]}$ to maximize the following entry log likelihood function:

$$EntryLL(\mu, \sigma) = \sum_{i}^{M} \sum_{j \in \mathcal{N}(\tau_i)} \left\{ \mathbb{1}[j \ enters \ i] \cdot ln \left(\Phi \left(\frac{ln\left(\widehat{\mathbb{E}[\pi_{ij}]}\right) - \mu(\tau_i, \mathcal{Q}_j)}{\sigma(\mathcal{Q}_j)} \right) \right) \right. \tag{16}$$

$$\left. + \left(1 - \mathbb{1}[j \ enters \ i]\right) \cdot ln \left(1 - \Phi \left(\frac{ln\left(\widehat{\mathbb{E}[\pi_{ij}]}\right) - \mu(\tau_i, \mathcal{Q}_j)}{\sigma(\mathcal{Q}_j)} \right) \right) \right\}$$

Where $\mathbb{1}[j \ enters \ i]$ is an indicator function that equals one if seller j enters auction i and is

zero otherwise. I assume that μ is a linear function of project location (state dummies), seller type (quality group dummies), and year of the sample (2016 dummy). I also allow σ to vary across permanent and transient sellers.

3.5 Identification

Identification of the demand parameters follows standard arguments. The utility specification includes quality-group fixed effects, state fixed effects, and time-period fixed effects. Therefore, identification of the coefficient on price, α, comes from variation in the prices quoted to different households by a given seller quality group, while also controlling for hardware type and common demand shocks across time and across states. Namely, the price coefficient is identified by the extent that choice probabilities change as sellers of a given type $\mathcal{Q}_j$ submit higher prices. Controlling for the (unobserved) quality of sellers is vital for identifying the price coefficient because the quality of sellers is likely to be correlated with price and also with the probability that individuals choose the bid. For example, more experienced sellers may have lower costs on average and also be more likely to be selected. Identification of the other parameters in the utility function follows similar arguments.

In addition to the price parameter, λ is another crucial parameter for the counterfactual analysis. Recall that λ (the nesting parameter) determines the extent that buyers will substitute away from the "outside option" if they receive additional bids. λ is identified by exploiting variation in the number of bids that households receive within a specific location and period. For example, in 2016 Q1 some households in Connecticut receive two bids while others may receive three or more. To consistently estimate λ, the variation in the number of bids should be coming from supply-side factors. For example, say Household A solicits bids the week after Household B and receives one fewer bid because one of the suppliers is now busy installing many other rooftop systems that week. On the other hand, the estimate of λ will be biased if variation in bids is driven endogenously by demand-side factors. For instance, if some buyers are ex-ante more likely to go through with buying solar (i.e., wealthier and more educated households) and therefore more suppliers choose to bid on the project. If this is the case, it will appear that the additional bids are causing more solar purchases, but in fact, the likelihood of buying solar is causing an increased number of bids. To alleviate these identification concerns, I collect additional demographic data on household income, education, the democratic vote share, and other factors.[33] I show that the estimates of the price parameter and the λ parameter are robust to including a wide range of demographic controls.

Once the demand parameters are identified, understanding the identification of marginal costs is also straightforward. With demand parameters in hand, we can compute firms' optimal markups. Then using the markups, we can employ Equation 14 to create a one-to-one mapping between bid prices and costs.

[33] I collect the data from the 2016 American Community Survey which reports zip code level demographic information.

Finally, the entry cost parameters are identified by variation in expected marginal profits, holding seller type constant.[34] In particular, μ and σ are pinned down by the extent the probability of entry of type $\mathcal{Q}_j$ sellers changes as expected marginal profits increase. In theory, we could trace out the entire entry cost distribution for each seller type non-parametrically if we observed the probability of entry at every level of expected marginal profit. To generate variation in expected marginal across auctions, we need exogenous variation in an observed variable, which does not affect the entry cost distribution but enters the seller's ex-ante payoff (expected marginal profit) before the entry decision. I assume that a project's size (large or small) does not affect the firm's entry cost, but does affect expected marginal profit. For this assumption to hold, it must be true that bid preparation time and effort is not different for larger residential projects compared to smaller residential projects on average. I also assume that a firm's average entry cost does not vary within a year.[35] Since expected marginal profits will vary each quarter, I am also using variation in expected profits across quarters but within a year to identify the entry cost parameters.

4 Results

4.1 Demand Parameter Estimates

The first column of Table 4 contains the estimates for the baseline demand specification using the two-step grouped fixed effects estimator. λ, the nesting correlation parameter is significant, indicating that it is important to account for buyer heterogeneity in the likelihood of adopting solar. The price coefficient is negative and statistically significant as expected.

Table 4 also shows the estimates for three of the quality group dummies, Quality Groups 2, 5, and 9. I only report the three groups in the interest of space, and I report the rest of the quality group fixed effects in Figure 15 in the appendix. The omitted quality group is "transient" sellers, so the coefficients should be interpreted as changes in utility relative to choosing a similar bid from a "transient" installer. The magnitude of the coefficients on the quality group dummies are large relative to the price coefficient. In particular, the estimates mean buyers would be willing to pay 27% more for a seller in the 9th quality decile relative to a firm in the 5th decile. The magnitude of the coefficients on the "Premium" and "Premium Plus" dummies are also substantial. Namely, the average consumer would be indifferent to paying 23% more for premium panels relative to standard panels. For a 7 kW system these numbers imply differences in willingness to pay of over $3000 depending on hardware and installer characteristics, which illustrates the importance of accounting for non-price factors in buyers' adoption decision.

Columns 2-5 of Table 4 show estimates under alternative specifications of the buyer utility. The second column shows how the coefficients change if I do not include any controls for seller quality. The price coefficient is substantially more negative, indicating that there is a

[34] While also controlling for common entry cost shocks across states and years.

[35] I allow the entry cost to change between 2015 to 2016 but not every quarter.

Table 4: Demand Estimates

	(1)	(2)	(3)	(4)	(5)
λ	0.294***	0.271***	0.275***	0.281***	0.387***
	(0.058)	(0.067)	(0.063)	(0.058)	(0.059)
Price ($/watt)	−0.794***	−0.990***	−0.811***	−0.827***	−0.834***
	(0.108)	(0.114)	(0.124)	(0.112)	(0.101)
Premium Panel	0.482***	0.643***	0.496***	0.473***	0.568***
	(0.072)	(0.079)	(0.085)	(0.098)	(0.069)
Premium Plus Panel	1.164***	1.743***	1.164***	1.085***	1.391***
	(0.143)	(0.165)	(0.175)	(0.159)	(0.147)
Microinverter	0.250***	0.313***	0.247***	0.227***	0.207***
	(0.078)	(0.077)	(0.091)	(0.081)	(0.066)
Seller Quality Group 2	−0.667***			−0.722***	
	(0.206)			(0.205)	
Seller Quality Group 5	0.153			0.114	
	(0.130)			(0.120)	
Seller Quality Group 9	0.697***			0.653***	
	(0.148)			(0.136)	
Star Rating $<$ 5					−0.285**
					(0.119)
Star Rating = 5					0.385***
					(0.097)
$>$ 6 Months on Platform					0.135**
					(0.063)
$>$ 1000 Total Installs					0.113**
					(0.056)
State FE	Yes	Yes	Yes	Yes	Yes
Time-Period FE	Yes	Yes	Yes	Yes	Yes
Quality Group FE	Yes	No	No	Yes	No
Size Control, Large-Project FE	Yes	Yes	Yes	Yes	Yes
Firm FE	No	No	Yes	No	No
Panel Brand FE	No	No	No	Yes	No
Observations	10,545	10,545	10,545	10,545	10,545

Notes: The first column contains estimates for the baseline model. Time-period fixed effects are included for each quarter-year of the sample. Each specification that contains quality-group fixed effects uses 10 quality groups in estimation, only the fixed effects for Groups 2, 5, and 9 are reported in the table, estimates for the other quality groups can be found in Figure 15. Star ratings and total installation experience were extracted from the EnergySage site in September 2017, 8 months after the sample-period was completed.

negative correlation between installers' unobserved quality and their cost.[36] The third column shows results with the seller fixed effects for each permanent seller, 65 in total (this model is also the first stage used to sort the sellers into quality groups). All of the estimates are very similar to the baseline estimates; which indicates that the specification of quality groups is not driving the results. The fourth column shows that results are also robust to adding panel brand fixed effects.[37] In Section 4.2.1, I discuss the robustness of the estimates to more alternative specifications including different numbers of quality groups.

The last column of Table 4 does not include quality group fixed effects but instead, adds observable measures of installer quality. The model consists of a dummy for if the seller had a five-star rating and another indicator for if the sellers rating was below five stars. Sellers without any reviews are the excluded group. I also include a dummy for if the installer had been active on the platform for at least six months and an indicator for if the firm had completed over 1000 total residential installations (both on and off the platform). These observable quality measures are imperfect because they were obtained from the platform database in September 2017, nine months after the sample period ended. Therefore, these variables do not correspond to the exact information that buyers observed when making purchase decisions. Nonetheless, the data can still be helpful for understanding which factors could be contributing to firms' unobserved quality (firm fixed effects). I find that buyers are much more likely to pick sellers that have five-star ratings, and installers with a sub-five star rating were chosen with a lower probability than installers with no ratings. Buyers also prefer sellers with more experience on the platform and with more overall installation experience. The magnitudes of the coefficient estimates also seem to be of similar magnitudes to the quality group fixed effects estimates in the baseline specification.

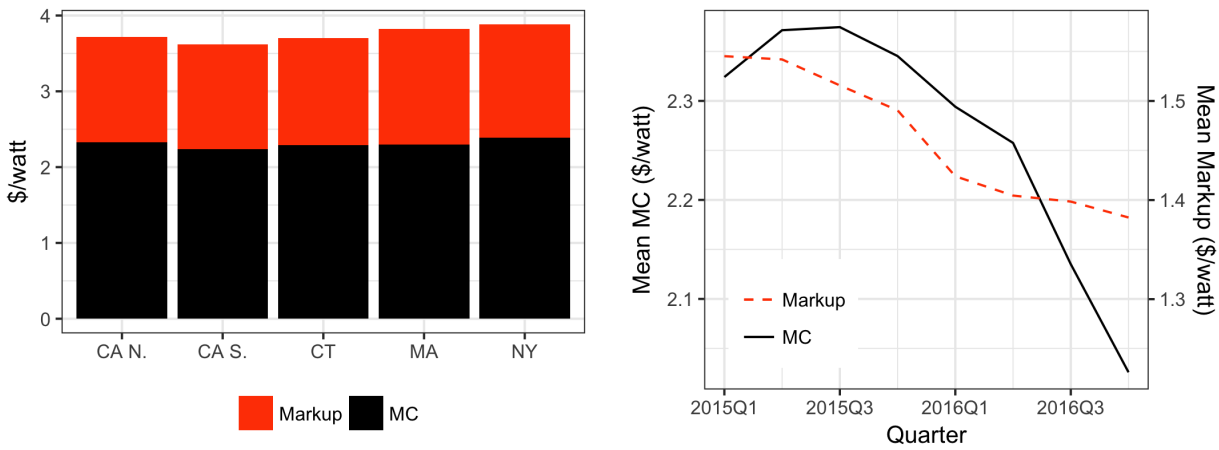
4.2 Marginal Cost, Markup, and Entry Cost Estimates

After estimating the demand parameters, I use the estimates to infer markups, and marginal costs for each bid in the data using the procedure explained in the previous section. Figure 6 reports summary statistics of the marginal cost and markup results. In the left panel, I decompose the average bid price for each state into an average marginal cost and an average markup component. The full sample average markup was \$1.42/watt, and the average marginal cost was \$2.19/watt. Therefore, markups accounted for 39% of prices on average. This implies that markups are a major portion of solar PV prices. However, these are gross markup estimates that do not include installers' overhead or marketing costs. Fu et al. (2016) report that installers' benchmark marketing cost should be \$0.31/watt and other overhead expenses should account for \$0.28/watt. Using the overhead cost estimates from Fu et al. (2016), I calculate that installers' mean net markups were \$0.83/watt, 23% of total system prices on average. The right panel of Figure 6 shows how gross markups and marginal costs changed over the sample. Markups fell

[36] In the next subsection, I provide evidence of this negative relationship.

[37] I include panel brand dummies for any brand that composed over 5% of bids in the data.

Figure 6: Decomposition of Solar PV Prices



(a) Price Decomposition by State - 2016 Q1

(b) MC & Markups Over Time

Notes: Panel (a) decomposes 2016Q1 price bids into a markups and a marginal cost component for each state on average. The markup component represents the difference between price and marginal cost and does not account for entry costs or other overhead costs. Panel (b) shows how average markups and marginal costs changed over the sample.

steadily as the platform became more competitive, accounting for over $0.15/watt decrease bid prices over the sample period. Marginal costs remained steady for the first half of the sample and then sharply declined in the second half of 2016. Figure 18 in the appendix shows that the fall in marginal costs can be explained by decreasing hardware costs during the second half of the sample period.[38] In Appendix Figure 17, I plot the average marginal costs and markups for each state over the sample. I also report the state-level estimates in Appendix Table 19.

To gain a better understanding of the main drivers of differences in marginal costs across bids, I run a regression with the implied marginal cost for each bid as the dependent variable and the characteristics of the bid as the explanatory variables. I include three dummy variables for installer quality in the regression: quality groups 1-3 (low quality), quality groups 4-7 (medium quality), and quality groups 8-10 (high quality). The results are shown in Appendix Table 17. Note that sellers with higher quality tend to have lower costs. Failing to control for sellers unobserved quality could lead to over-estimates of buyers' price elasticities. If we do not control for the seller's quality, it would appear that customers are picking sellers that have lower prices, but we would also be picking up the negative correlation between price and quality. The estimates from the demand specification without installer quality group controls (see Column 2 of Table 4) support this claim. The price coefficient is indeed lower (more negative). The negative relationship between quality and cost can also be seen in Figure 21 in the appendix. The figure plots the marginal cost distribution for each of the quality groups for one market, and it is clear that the cost distribution for the highest quality sellers is shifted to the left.

[38]Hardware cost estimates were obtained from Bloomberg New Energy Finance. Bloomberg Reports Module Spot Prices and an inverter price index monthly.

The other coefficients in the marginal cost regression also have the expected signs. Systems with high-quality panels and microinverters carry higher marginal costs on average. Also, the negative estimate on project size indicates there are economies of scale within an installation. Economies of scale arise because some parts of the marginal cost of a project such as permitting fees are fixed and do not vary with the size of the installation.

The entry cost parameter estimates can be found in Appendix Table 20. In appendix Figure 19, I have also used the estimated parameters to plot the predicted entry probabilities for Massachusetts in 2016 for three seller types as a function of expected marginal profits. Permanent sellers are more likely to enter for any level of expected profits. Also notice that even if expected profits get very large, the entry probabilities do not approach 100%. This finding could be explained by installers occasionally having obligations to complete other projects outside the platform and may be capacity constrained and unable to submit a bid regardless of how profitable the project appears to be.

4.2.1 Comparison of Estimates to Reported Costs and Robustness Checks

In this section, I compare the marginal cost estimates to existing survey-based estimates and reported costs. I also discuss the robustness of the results to changes in the specification of the model.

Figure 20 in the appendix compares the estimated marginal costs and markups from the model to existing estimates from National Renewable Energy Laboratory[39] (NREL) and reported marginal costs from installers' earnings statements. The estimated marginal costs are similar to the marginal costs reported in quarterly reports by large publicly traded installers, Solar City, Vivint, and Sunrun. In the first quarter of 2016, Solar City reported that installation costs were \$1.90/watt. Vivint and Sunrun reported costs of \$2.34/watt and \$2.73/watt, respectively. My estimate for the average marginal cost in 2016Q1 falls right between these reports at \$2.29/watt. My cost estimate is slightly higher than NREL's benchmark estimate of \$1.89/watt. This result is not surprising considering the NREL analysis quantifies a lower bound for costs (i.e., industry best practices). Overall, the model estimates appear to be consistent with existing studies and reports.

Before turning to the counterfactual exercises, I also perform a series of robustness checks to ensure that the demand parameters are not sensitive to small changes in modeling choices. As discussed in Section 3.5, it is critical to ensure that endogenous supplier participation choices are not driving the demand parameter estimates. In particular, λ, the nesting parameter and the price parameter may be biased if suppliers are selectively bidding more often (within a location and period) on households that they believe would be more likely to purchase solar. To alleviate some of these concerns, I collect zip code demographic data from the American Community Survey. I re-estimate the demand model including additional controls in the utility function such as income, race, educational attainment, political preference, and household size.

[39] See Fu et al. (2016).

Table 13 in the appendix shows that the nesting parameter and the price sensitivity parameter are almost unchanged after including these additional controls. Ideally, we would control for demographics at the household level. Unfortunately, the EnergySage data only include project zip code but not project address or household characteristics (I do observe the size of each solar system in the bid data which provides a proxy for household size and income). Nonetheless, it is comforting to know that the demand estimates hardly change after incorporating zip code level demographic shifters in the utility function.

Table 15 shows some additional robustness checks for the demand estimates. For example, I add controls for financing offers (loans), state-level incentives, other demand shifters (state-year fixed effects), and time-varying controls for installer experience on the platform.[40] The price coefficient or λ do not change much under any of these alternative specifications.

Table 16 explores alternative methods for calculating the quality-group fixed effects. In column two, instead of sorting the firms into ten equally sized groups, I alternatively define the quality groups by using a k-means clustering algorithm. The k-means algorithm assigns each installer into ten clusters to fit the data best; each cluster can contain a different number of firms. In the third column, I again use ten equally sized quality groups, but I use an iterated estimator for the groupings. At each iteration, the installers are assigned to new groups based off of the firms average residuals from the previous iteration. I continue to update the groups until convergence. The motivation for this estimator is the possibility of misclassifying the quality groups under the two-step estimation approach. Namely, if the installer fixed effects are imprecisely estimated, some sellers may be sorted into the "incorrect" group which could bias the other parameter estimates. Fortunately, I find the price coefficient and λ are very similar to the baseline estimate using both alternative estimators. Finally, Appendix Figure 16 shows how the ratio of the price coefficient to the nesting parameter $\alpha/(1-\lambda)$ changes if I use a different number of quality groups. Under the nested logit demand structure, the term $\alpha/(1-\lambda)$ enters multiplicatively into each firm's optimal markup, so this ratio determines how large the markup estimates will be. The ratio largely fluctuates if I use fewer than five quality groups in estimation. The estimated fraction $\alpha/(1-\lambda)$ stabilizes at around six quality groups. Any specification with over six groups, including the firm fixed effects model (65 groups) yields similar estimates.

I also test the robustness of the markup estimates to changes in the definition of auction type categories. One concern is that the project size categories do not allow for sufficient heterogeneity across projects of different sizes. Column 2 of Appendix Table 18 shows that markup estimates barely change if I use three project size categories instead of the two categories used in the baseline model. Another possibility is that I am not sufficiently controlling for geographic heterogeneity in competition, costs, and preferences. Appendix Figure 22 shows how

[40] Three of the four states did not have state incentives that varied over the sample so any state incentive should be picked up through the state fixed effects. New York is the only state that had subsidies that changed over the sample period. The estimated coefficient on "loan offered" is negative, likely because this variable is endogenous (consumers are allowed to request financing when they create an account).

the average number of bids changes across counties and also how average bid prices vary across counties. Massachusetts and Connecticut are smaller states, so the variation in prices across counties is minimal. However, in New York and California, there is more geographic variation in prices. Specifically, prices are higher in major urban areas like New York City and San Francisco. To account for this heterogeneity, I separate New York into two different regions, NYC Metro and Upstate (rest of state). I also divide California into three geographic areas (instead of two), Bay Area, Urban Southern California, and Rural California (rest of state). Column 3 of Appendix Table 18 shows that the markup results are robust to these changes. Furthermore, Columns 4 and 5 show that the results are also robust to changes in the length of "time-period" used to sort the projects into types.

An additional modeling assumption that could impact the results is that both consumers and suppliers are modeled as static decision makers. On the supply side, sellers do not consider the impact of their current price bid on the profitability of subsequent projects. There are a few potential channels through which dynamics could affect the seller's pricing problem. First, if sellers anticipate learning-by-doing, they may set lower prices today so that they can gain experience and reduce their marginal cost of future projects. Bollinger and Gillingham (2014) show evidence of appropriable learning effects in solar installations when a contractor has little experience but that the learning effect declines as an installer completes more projects.[41] The average installer on EnergySage had installed more than 1000 residential systems by 2016. Therefore, learning is unlikely to be a significant factor in pricing decisions within these relatively mature installation markets. The static pricing assumption may be also be violated if firms place lower prices today so that they can increase their reputation (through ratings and reviews) to increase the probability of future sales. To show evidence that reputation dynamics are not a first-order concern, Table 14 provides demand estimates that control for the number of cumulative sales that each seller has completed on the platform at the time a bid was submitted. The coefficient on the cumulative sales variable is close to zero and is not statistically significant. For these reasons, dynamics should not cause a substantial bias in estimating marginal costs.

Several studies have shown the importance of dynamics in the consumer's solar installation decision (Burr, 2014; De Groote and Verboven, 2016; Langer and Lemoine, 2017; Reddix II, 2015). In this paper, consumers are not explicitly modeled dynamically. However, the framework does not rule out forward-looking behavior. The previous literature has specified consumer adoption as a function of market average prices and abstracted from each consumer's decision about which installer to choose. In the current framework, each consumer decides from a menu of prices and an outside option. Here, the outside option could include waiting to solicit bids again at a later time. The value of the "outside option" is allowed to shift across space and time in the through the state and time-period fixed effects in the utility function.

[41] Bollinger and Gillingham (2014) find evidence of learning-by-doing in the California solar market using data from 2002-2012.

4.3 Counterfactual Analysis

In this section, I analyze counterfactual scenarios in which I change the competitive environment and subsidies available in the solar PV market. Specifically, the counterfactuals allow us to see how prices and solar adoption decisions vary when if we increase the number of bids that potential buyers obtain while holding preferences fixed.

In the first counterfactual, I simulate prices and adoption rates in the case that buyers were not able to use the platform to solicit bids and there is also no federal subsidy available for solar PV. In the absence of the platform, I assume that buyers receive two quotes on average. This assumption is consistent with Aggarwal (2015) who found that 58% of buyers, obtained two or fewer quotes and only 3% received four or more quotes. When simulating each auction in this counterfactual, I scale down each seller's entry probability by 50% relative to the fitted model. This simulation leads to buyers getting 2.08 bids on average. Installers that enter each auction have rational expectations about the number of competing bidders and adjust their prices accordingly. In Appendix B, I discuss the robustness of the assumption that buyers obtain two bids off the platform.

In the second counterfactual, I consider the case where buyers still cannot use the platform (and therefore, obtain approximately two bids on average) but have access to the 30% federal subsidy. Again, sellers can re-optimize their prices to capture some of the subsidy.

Next, I simulate a situation where buyers can access the platform but do not have access to the federal subsidy. Sellers in this counterfactual scenario can update their entry strategies to account for the removal of the subsidy and also update bidding strategies accordingly. I can then measure how giving consumers access to the platform compares to giving them the 30% subsidy but no access to more potential sellers (the second counterfactual). Finally, I solve for outcomes for when buyers use the platform and exploit the 30% subsidy for solar PV. This is the case observed in the data and used to fit the model.

In all the counterfactuals, not only can we evaluate the change in prices and demand, but we can determine seller profits under the different scenarios. Additionally, we can compute consumer surplus under the preferences revealed by the utility model. Of course, it should be noted that the revealed preferences are those of potential consumers that actually visited the platform during the sample period. Also, I estimate the utility model with decisions made by consumers while using the platform interface. If the platform interface provides better information about the options, then this should give better estimates of ex-post consumer surplus. However, a limitation of my approach is that the model may not predict how choices between installers would change if buyers have different perceptions about seller quality or hardware quality off of the platform.

4.3.1 Solving for Counterfactual Equilibria

For each counterfactual, I numerically solve for the new equilibrium following a policy change. To obtain the counterfactuals, I use a backward solution strategy. In particular, I first derive the equilibrium bidding strategies of sellers given competing sellers entry behavior. Next, I go back and solve for optimal seller entry given the equilibrium bidding strategies. I then combine these two steps to compute the equilibrium outcome for the entire system.

To derive a seller's equilibrium bid under a counterfactual scenario, we need information on the equilibrium distribution of bids (to calculate each firm's markup). However, the equilibrium bid distribution will change under the counterfactual, and we need to estimate the new distribution. I begin by using the original distribution of bids (and entry probabilities) observed in the data and then use a root-finding algorithm to calculate a new optimal bid price for each quote in the data. Solving for the optimal prices provides a new updated distribution of bids. I then update each firm's equilibrium entry probability given this updated distribution of bids. I repeat the process of updating the bid distribution and entry probabilities until convergence. In Appendix A, I provide a more detailed discussion of the algorithm used for computing the counterfactuals.

4.3.2 Model Fit

Table 5 compares the structural model predictions to the observed data. Relative adoption is the percentage of total solar system sales compared to the observed data (100% relative adoption would mean that the model predicts the same number of sales as in the observed data). The model slightly overpredicts the number of adoptions compared to the observed data, but the bid prices and entry predictions match very closely with the actual observation. Overall, the model seems to fit quite well.

Table 5: Model Fit

	Avg. Bid Price	Installs Relative to Observed	Avg. Number of Bids
Observed	\$3.61/watt	-	3.95
Model Prediction	\$3.64/watt	107.6%	3.96

4.3.3 Counterfactual Results

I report results from the counterfactual exercises in Table 6. The first column shows the outcomes for the baseline simulation where each buyer cannot access the platform. In the baseline simulation, buyers obtain 2.08 quotes on average, and there is also no 30% investment tax credit (ITC) for solar purchases. Without the platform or the ITC, the lowest bid that households receive is \$3.13/watt on average.[42]

[42] The average net bid price is also \$3.13/watt because there is no subsidy in this case.

In the second column, we see that providing a 30% subsidy for solar results in 15% higher bid prices (holding the average number of bids fixed at two). Sellers can capture part of the subsidy by modestly increasing prices without reducing quantity sold since net prices will be lower for consumers. At $2.53, the average net price for prospective consumers is 19% lower under the subsidy policy. In this scenario, adoption increases by 56% relative to the baseline. This finding is similar to other estimates in the literature. For example, Hughes and Podolefsky (2015) find that the California Solar Initiative (CSI) subsidies doubled the number of total installations that occurred in California before 2013. Here, the predicted increase in installs from the ITC is roughly half of what Hughes and Podolefsky (2015) find for the CSI, but the price decrease is also smaller at $0.60/watt compared to the $1.50/watt change that was induced by the CSI. In addition to the increase in solar installs, we also see that both consumer and producer surplus increase by 93% and 126% respectively. However, total welfare only increases by 4% after we account for the cost of the subsidy.[43]

Table 6: Counterfactual Results

	(1) Baseline - No Platform, No ITC	(2) No Platform, with ITC	(3) Platform Access, No ITC	(4) Platform Access, with ITC
Low Bid (Mean)	$3.13	$3.61	$2.92	$3.32
Net Low Bid (Mean)	$3.13	$2.53	$2.92	$2.32
Number of Bids (Mean)	2.08	2.08	3.52	3.96
Adoption (% Increase)	-	56%	84%	213%
CS (% Increase)	-	93%	117%	378%
PS (% Increase)	-	126%	72%	314%
Welfare (% Increase) = CS+PS-Subsidy Cost	-	4%	84%	114%

Notes: "Net low bid" is the nominal bid price minus any subsidy. The last four rows show the percentage increase in solar adoption, consumer surplus, producer surplus, and welfare compared to the baseline (Column 1). The first column shows results for the baseline simulation where buyers cannot access the platform, and there is no 30% investment tax credit (ITC). Column 2 displays the results under the scenario where buyers can't use the platform but are given the 30% ITC. Columns 3 and 4 show market outcomes when consumers use the platform without and with the ITC respectively.

The third column shows how the market prices and adoption rates change if customers can use the platform to solicit bids, but no 30% subsidy is available. Competition on the platform induces a 7% fall in prices relative to the "no platform" baseline. Moreover, adoption rates are boosted by 84% when consumers can use the platform. The platform also provides substantial surplus gains, particularly for buyers who enjoy a 117% increase in consumer surplus. Despite lower equilibrium bid prices, sellers also benefit from a 72% jump in overall profits relative to

[43]Note that it is possible for the subsidy to increase welfare because firms are bidding above marginal cost. The welfare estimates do account for any other external effects of rooftop solar such as pollution reductions or intermittency.

the baseline case. Firm profits grow because the decrease in prices is more than offset by the increase in total sales. Seller profits are still lower than the "no platform" with the subsidy case.

In the final column, we see that coupling the platform together with the ITC policy leads to substantial gains in solar PV adoption and welfare. In comparison to the baseline scenario, solar system installations expand by over 200%. At the same time, both producer and consumer surplus rise by over 300%. Both sides of the market benefit from the sizable increase in system purchases and consumers benefit from lower prices and larger choice sets. The increase in welfare for market participants is partly offset by the cost subsidizing each purchase, but even after accounting for the subsidy bill, overall welfare more than doubles relative to the baseline case.

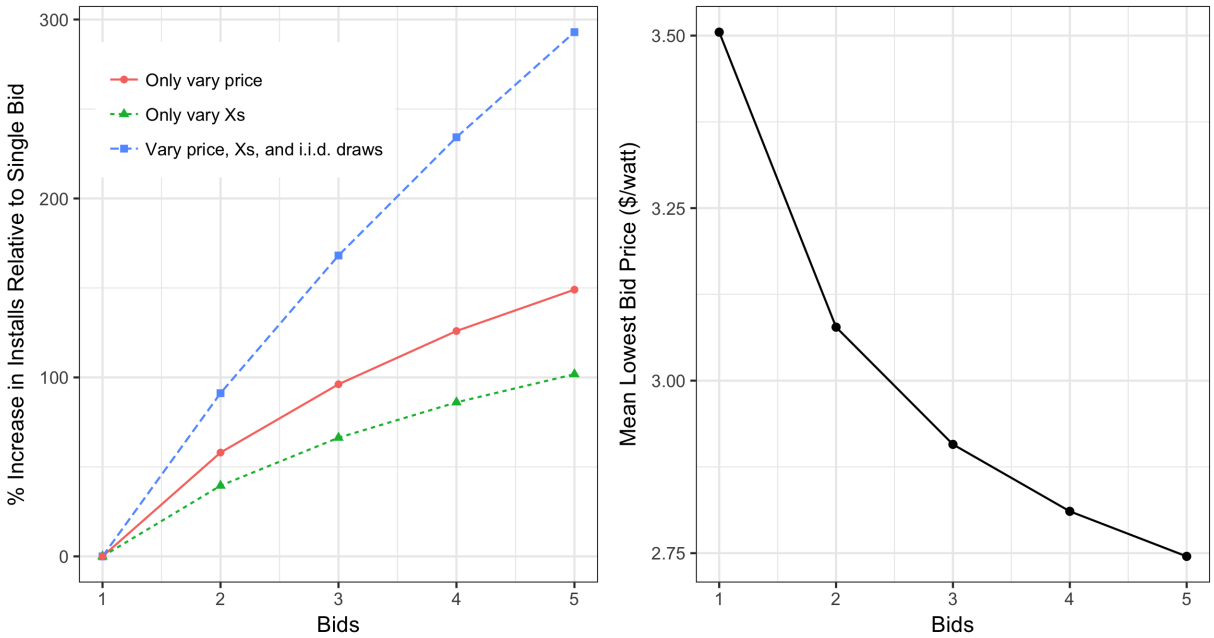
Overall, the counterfactual analysis reveals considerable benefits of increasing the number of bids for potential rooftop solar projects. Perhaps the most surprising result is that expanding the number of bids per project can increase solar installs by as much (or more) than the ITC subsidy policy that only reduces the price. While the effect of the platform on solar purchases is relatively large, the magnitude is consistent with the reduced form regression estimates. Namely, the linear regression estimates from section 2.2.1 that included a larger set of spatial and time controls (such Zip code fixed effects) implied that doubling the number of bids was associated with a 68-76% increase in solar purchases depending on the specification.

A remaining question then is to understand better why the platform is so effective at increasing the number of sales. As discussed in previous sections, there are a few potential channels through which the platform could increase overall adoption. First, bid prices decline when buyers have access to the platform. Second, the platform could increase adoption by connecting buyers with superior sellers or enhanced hardware offerings. Reduced prices do not appear to be the only driving factor behind the significant increase in adoption rates induced by the platform. To see this, notice that adoption is actually higher with the platform and no subsidy (Column 3) than with no platform with the subsidy (Column 2) even though prices are lower in the latter case. This means that non-price factors may be a significant avenue through which the platform increases adoption.

To better understand the channels through which auction participation impacts solar adoption, I use the model to conduct a series of experiments. First, I simulate the number of solar installations that would occur if each household were to obtain precisely one bid. Next, I solve for the number solar adoptions if buyers obtain two, three, four, and five quotes. The blue dashed line in the left panel of Figure 7 shows how installs evolve as the number of bids varies. In the context of the consumer choice model, there are three factors that contribute to more purchases as the number of bids grows: (1) lower prices,[44] (2) improved observable non-price

[44]With each additional bid, buyers will increase the odds of finding lower cost installers and each installer will reduce their markup.

Figure 7: Decomposing Increase in Solar Installations from Additional Bids



Notes: The left panel shows how the total amount of installations changes under different assumptions about the number of bids each consumer receives. The blue-dashed line shows how total solar installations evolve if sellers are allowed to update their prices and buyers gain access to new installers (that are differentiated in quality and in the type of hardware they use). The solid red line shows how total installs change if we hold seller quality and hardware type fixed and also hold fixed the i.i.d. draws (consumers only receive one utility shock that is the same for each inside option). The dotted green line shows how total installs changes if we allow seller quality (non-price Xs) to vary but hold prices constant at the monopoly price level and also fix the utility shocks to be the same for each inside option. The right column shows how the equilibrium low bid price (average) changes as the number of bidders changes.

factors (Xs that enter the buyer's utility function),[45] and (3) improved matches due to unobserved factors (each bid gets a new utility shock ε_{ij}).[46] The additional plots in the left panel isolate the impact of the price mechanism (1) and the quality mechanism (2).

The solid red line represents the pure price effect (1) of increasing the number of bids. To isolate the price channel, I allow prices to vary with the number of bids but holding the non-price Xs and utility shocks constant as the number of bids increases. I implement these simulations by setting each bidder's X vector equal to the mean X vector that households obtained in the monopoly bidder case.[47] Additionally, I set the λ parameter equal to 0.99 to approximate a

[45] An additional bid means the buyer gets a new draw of Xs. This means the consumer can potentially gain access to installers from higher quality groups and get bids for different hardware such as premium panels and microinverters. Section 4.1 showed that these non-price factors are pivotal in consumers' choices.

[46] Finally, added bids imply that consumers collect an additional utility shock ε_{ij}. The importance of the additional utility shocks on the adoption will depend on the correlation parameter λ. If λ was equal to one, then the utility shocks associated with each bid are perfectly correlated, and therefore additional utility shocks do not increase the probability of choosing the inside good. For lower λ values, incremental ε_{ij} draws will contribute to more adoption. One interpretation of a lower λ is that buyers have heterogeneous project needs or preferences over installers and that additional bids help to find better-matched installers.

[47] The mean X's are calculated separately by auction type.

situation where buyers get the same utility shock for each bid.[48] The right panel of Figure 7 shows how the lowest bid price changes as the number of bids grows. We see that there is a substantial decline in price (12.3%) from increasing the number of quotes from one to two. As auction participation is scaled up further, the price decline becomes more modest, with the next three added bids contributing an additional 9% reduction in prices. The solid red line shows that reduced prices drive up the number of solar purchases by 150% in the five bid case compared to the single bid case. However, the price channel alone can only account for 51% of the overall increase in solar installs.

The green dotted line shows the trend in solar adoptions if we allow the non-price Xs to vary with each added bid but hold each bidder's price constant at the average monopolist price and again fix $\lambda = 0.99$.[49] As consumers collect additional bids, they are more likely to access sellers from better quality groups and get bids for premium hardware. Due to these factors alone, getting five bids doubles the number of adoptions compared to receiving one bid. All in, observable non-price factors can explain 35% of the total increase in installs.

The results displayed in Figure 7 are instructive about why the platform can lead to substantial gains in solar panel adoption. The impact of lower prices and subsidies has been established in previous research and remains an important channel here. However, lower prices can only explain half of the adoptions that are caused by the platform. The platform is also effective because it increases access to premium hardware and higher-rated sellers. Moreover, when consumers have heterogeneous preferences, the platform also increases adoption by connecting buyers to better-matched sellers.

5 Discussion and Conclusions

Increasing renewable energy investment is viewed as a pertinent part of a larger goal to curb carbon emissions from the electricity sector worldwide. In particular, many national and local governments have emphasized expanding the adoption of residential solar PV systems. Despite falling hardware costs and increased public interest in solar energy, only a small fraction of households have adopted the technology.

In this paper, I estimate a structural model of consumer choice and seller bidding in the residential solar PV market. The model allows me to quantify the importance of price, hardware quality, and installer quality in consumers' solar adoption decision. I also use the model to measure the extent of market power in the solar industry. In particular, I use high-resolution data on firm bidding behavior to identify marginal costs, markups, and bid preparation costs.

My estimates provide evidence that both hardware and installer quality are critical factors in buyers' solar purchase decision. Namely, the average buyer is willing to pay substantially more for a high-quality firm to install their PV system. I also show that supplier markups compose a

[48] The nested logit choice probabilities are not defined exactly at $\lambda = 1$. Instead, the choice probability approaches the case of perfectly correlated errors as λ approaches one from the left.

[49] The average monopolist (single bidder) prices are calculated separately for each auction type.

substantial portion of overall PV system prices. Market power is of particular concern in this setting because solar PV provides positive environmental benefits.

Most current public policies have used subsidies, tax credits, and rebates to encourage the adoption of residential solar PV systems. These policies have been shown to be an effective method for increasing solar PV purchases. However, these programs also require large public expenditures and risk being cut if government funds are limited. This paper provides evidence that using platforms to increase competition and expand buyers choice sets can also serve as an effective way to increase adoption. Platforms can reduce system prices by making the market more competitive. Additionally, increasing buyer participation on platforms can also lead to more purchases by linking consumers to higher quality installers.

Both state and federal governments could channel more effort to informing the public about existing platforms or by developing their own platforms to link buyers and sellers. Since platforms have already been designed to connect buyers and sellers in this market, it is likely that increases in adoption from expanding platform participation could come at a relatively low cost and lead to significant welfare gains for both consumers and producers.

References

Aggarwal, V. (2015). EnergySage's Solar Installer Survey 2015 Results. Technical report, EnergySage Inc.

Bollinger, B. K. and Gillingham, K. (2014). Learning-by-doing in solar photovoltaic installations.

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):S85–S122.

Brown, J. R. and Goolsbee, A. (2002). Does the Internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, 110(3):481–507.

Brown, Z. Y. (2017). An Empirical Model of Price Transparency and Markups in Health Care. Technical report, Working Paper, Columbia University.

Burr, C. (2014). Subsidies and Investments in the Solar Power Market. Technical report, Working Paper, University of Colorado.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity. *Econometric Theory*, 13(2):185–213.

De Groote, O. and Verboven, F. (2016). Subsidies and myopia in technology adoption: Evidence from solar photovoltaic systems. Technical report, Working Paper, KU Lueven.

Diamond, P. A. (1971). A model of price adjustment. *Journal of Economic Theory*, 3(2):156–168.

Fu, R., Chung, D., Lowder, T., Feldman, D., Ardani, K., and Margolis, R. (2016). *U.S. Solar Photovoltaic System Cost Benchmark: Q1 2016.*

Gehrig, T. (1993). Intermediation in search markets. *Journal of Economics & Management Strategy*, 2(1):97–120.

Gerarden, T. (2017). Demanding innovation: The impact of consumer subsidies on solar panel production costs. Technical report, Working paper, Harvard University.

Gillingham, K., Deng, H., Wiser, R., Darghouth, N., Nemet, G., Barbose, G., Rai, V., Dong, C. G., and others (2016). Deconstructing solar photovoltaic pricing. *The Energy Journal*, 37(3).

Gillingham, K. and Tsvetanov, T. (2014). Hurdles and steps: Estimating demand for solar photovoltaics. Technical report, Working Paper, Yale University, Working Paper, Yale University.

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

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal Nonparametric Estimation of First-Price Auctions. *Econometrica*, 68(3):525–574.

Hall, G. and Rust, J. (2003). Middlemen versus market makers: A theory of competitive exchange. *Journal of Political Economy*, 111(2):353–403.

Hughes, J. E. and Podolefsky, M. (2015). Getting green with solar subsidies: Evidence from the California solar initiative. *Journal of the Association of Environmental and Resource Economists*, 2(2):235–275.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The Quarterly Journal of Economics*, 122(3):879–924.

Krasnokutskaya, E. and Seim, K. (2011). Bid preference programs and participation in highway procurement auctions. *The American Economic Review*, 101(6):2653–2686.

Krasnokutskaya, E., Song, K., and Tang, X. (2019). The Role of Quality in On-Line Service Markets (forthcoming). *Journal of Political Economy.*

Langer, A. and Lemoine, D. (2017). Designing Dynamic Subsidies to Spur Adoption of New Technologies. Technical report, Working Paper, University of Arizona.

Luco, F. (2016). Who Benefits from Information Disclosure? The Case of Retail Gasoline. Technical report, Working Paper, Texas A&M.

Margolis, R., Feldman, D., and Boff, D. (2017). Q4 2016/Q1 2017 Solar Industry Update. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

O'Shaughnessy, E. J. and Margolis, R. M. (2017). The Value of Transparency in Distributed Solar PV Markets. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Pless, J. and van Benthem, A. (2017). The Surprising Pass-Through of Solar Subsidies. Technical report, National Bureau of Economic Research.

Reddix II, K. D. (2015). *Powering Demand: Solar Photovoltaic Subsidies in California.* PhD thesis, The University of North Carolina at Chapel Hill.

Salz, T. (2017). Intermediation and competition in search markets: An empirical case study. Technical report, Working Paper, Columbia University.

Spulber, D. F. (1996). Market making by price-setting firms. *The Review of Economic Studies*, 63(4):559–580.

Stahl, D. O. (1989). Oligopolistic pricing with sequential consumer search. *The American Economic Review*, pages 700–712.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Yoganarasimhan, H. (2015). Estimation of beauty contest auctions. *Marketing Science*, 35(1):27–54.

For Online Publication Only

Appendices

A Algorithm for Solving Counterfactuals

1. For each auction type (i.e., CT-2016Q1-Small), start with a vector of all bids submitted for projects of that type ($\mathbf{B_0}$) and start with an entry probability for each potential seller for that auction type ($\mathbf{E_0}$).

2. Calculate each firm's optimal price given the current distribution of prices and entry probabilities from step one. Store the new vector of bids $\mathbf{B_1}$.

   - Equation 8 is the first order condition for each firm's optimal price. The first order condition does not have a closed form, so simulate S=1000 iterations of each auction type to approximate the integrals numerically.

3. Use the updated prices (and conditional winning probabilities) from Step 2 to calculate each potential entrant's expected marginal profit of entering the auction. Then use the new expected profits to update each firm's entry probability. Store the new entry probabilities $\mathbf{E_1}$.

   - Note: For the "no platform" counterfactuals, the entry probabilities are set to be equal to one half the entry probabilities observed in the data. For these counterfactuals, the entry probabilities are fixed and are not updated endogenously. Therefore, $\mathbf{E_0} = \mathbf{E_1}$

4. Measure the difference between each of the original prices and the updated prices and measure the difference between the original and updated entry probabilities. Stop if $||abs(\mathbf{P_1} - \mathbf{P_0})||_\infty < \delta_p$ and $||abs(\mathbf{E_1} - \mathbf{E_0})||_\infty < \delta_e$. Otherwise replace $\mathbf{P_0}$ with $\mathbf{P_1}$ and $\mathbf{E_0}$ with $\mathbf{E_1}$ and then start over at Step 1.

   - I set $\delta_p = .00001$ and $\delta_e = .005$. It is possible to obtain convergence with smaller values of δ_e if we increase the number iterations, S, that each auction type is simulated to calculate the first order conditions. For a fixed number of simulated auctions (i.e., S=1000), a small change in entry probabilities can cause a few firms to change their bids by a couple of cents. However, increasing the number of auction iterations becomes computationally burdensome, so S=1000 is used in order to do more counterfactuals and robustness checks.

B Assumption for "No Platform" Counterfactual

The baseline counterfactual simulation assumes that households would receive two price quotes on average if they were unable to access the platform. While this is a reasonable starting point, some buyers may collect more (or fewer) than two quotes in the absence of the platform. Table 7 in the appendix shows how the predicted increase in adoption changes under alternative assumptions about the "no platform" counterfactual. In the first column, I assume that (in expectation) each buyer receives 1.5 quotes. I implement this counterfactual by adjusting each firm's entry probability by a factor of 0.25 relative to the entry probabilities observed in the data. Sellers are allowed to update their equilibrium bids accordingly. Under this assumption, the platform would induce a 252% increase in adoption relative to this new baseline. The effect is much larger because as the average number of bids approaches zero, the number of purchases declines non-linearly (as seen in Figure 7) so the platform looks relatively more attractive. In the third column, we assume that each consumer would get 2.5 price quotes (in expectation) without the platform.[50] Under this assumption, the platform would increase overall adoption by 45%.

An important takeaway from this exercise is that the precise estimate of the effect of the platform on solar panel adoptions is relatively sensitive to the assumption about the number of bids buyers would obtain off of the platform. However, regardless of the specific assumption, the main narrative that additional bids lead to substantial increases in adoption holds true. Even in the most conservative case, given in column 3, where we speculate that the platform only increases the average number quotes by one (from 2.49 to 3.52), we still see a greater than 40% increase in the number of systems installed.

Table 7: No Platform Counterfactual - Alternate Asssumptions about Bids Received

Number of Bids (Mean)	1.49	2.08	2.49
Low Bid Price (Mean)	\$3.29	\$3.13	\$3.06
Effect of Platform on Adoption (% Incr.)	252%	84%	45%

Notes: This table shows the simulated effects of providing consumers with access to the platform under alternate assumptions about the "no platform" counterfactual. Each of the counterfactuals is implement by adjusting the observed probabilities that each seller will bid on a given project (entry probabilities). The entry probabilities are scaled by 0.25 in column 1 and by 0.5 and 0.65 in columns 2 and 3 respectively.

[50] This counterfactual is implemented by scaling down each installer's observed entry probabilities by a factor of 0.65.

C Tables and Figures

Figure 8: Google Maps Photo of the Rooftop for a Potential Project

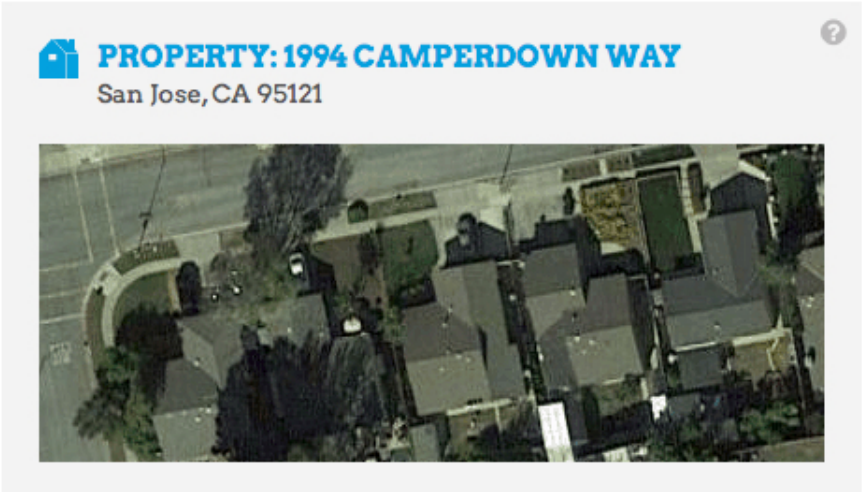


Figure 9: EnergySage Dashboard for Comparing Submitted Quotes

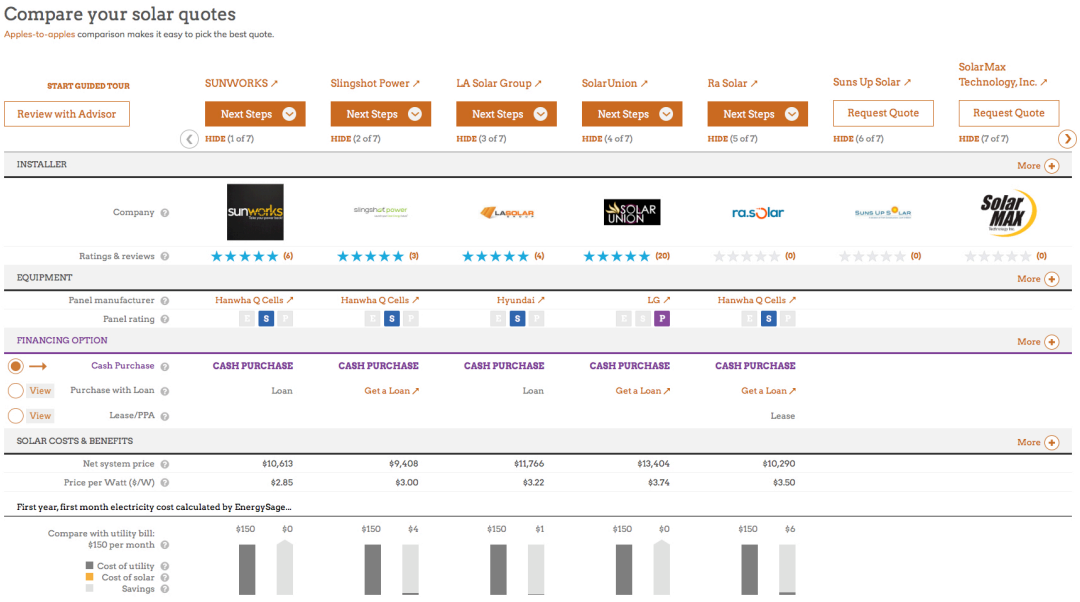


Table 8: Project Summary Statistics by State (Means)

	CA North	CA South	CT	MA	NY	Total
Realized Number of Bids	3.947	4.649	3.509	3.669	3.096	3.953
	(1.811)	(1.950)	(1.808)	(1.776)	(1.660)	(1.903)
Project Size (Watts)	6312.3	6630.4	8288.0	7498.0	8887.3	7155.9
	(2902.9)	(2827.0)	(3101.4)	(2950.4)	(3546.9)	(3138.6)
Distinct Panel Brands	3.051	2.816	2.457	2.662	2.620	2.795
	(1.281)	(1.228)	(1.162)	(1.235)	(1.398)	(1.277)
# of Bids w/ Premium Panel	1.027	2.250	1.353	1.468	0.886	1.486
	(0.994)	(1.628)	(1.293)	(1.336)	(0.876)	(1.398)
# Bids w/ Premium Plus Panel	0.214	0.0727	0.266	0.0945	0.235	0.159
	(0.471)	(0.281)	(0.538)	(0.311)	(0.463)	(0.410)
# of Bids w/ Microinverter	2.564	3.619	2.196	3.364	2.156	2.929
	(1.613)	(1.973)	(1.812)	(1.825)	(1.363)	(1.854)
# of Bids from Permanent Sellers	3.221	3.944	2.735	3.271	2.011	3.238
	(1.593)	(1.751)	(1.403)	(1.811)	(1.215)	(1.733)
Observations	10545					

Standard deviations are listed in parentheses.

Table 9: Mean Bid Characteristics by State

	CA North	CA South	CT	MA	NY	Total
Price ($/watt)	3.590	3.476	3.739	3.773	3.765	3.611
	(0.482)	(0.416)	(0.452)	(0.486)	(0.650)	(0.494)
Premium Panel (0,1)	0.266	0.502	0.396	0.408	0.292	0.386
	(0.442)	(0.500)	(0.489)	(0.492)	(0.455)	(0.487)
Premium Plus Panel (0,1)	0.0554	0.0162	0.0778	0.0263	0.0774	0.0412
	(0.229)	(0.126)	(0.268)	(0.160)	(0.267)	(0.199)
Microinverter (0,1)	0.665	0.807	0.642	0.936	0.710	0.761
	(0.472)	(0.395)	(0.479)	(0.245)	(0.454)	(0.426)
Permanent Seller (0,1)	0.835	0.879	0.800	0.910	0.662	0.841
	(0.371)	(0.326)	(0.400)	(0.286)	(0.473)	(0.365)
Observations	40575					

Standard deviations are listed in parentheses.

Table 10: Seller Summary Statistics

	Mean	SD	10-%tile	50-%tile	90-%tile
Total Bids Submitted	289.05	488.14	9.00	87.50	830.00
Avg. Star Rating (if Rated)	4.80	0.64	4.00	5.00	5.00
5 Star Rating	0.84	0.36	0.00	1.00	1.00
Below 5 Star Rating (if Rated)	0.16	0.36	0.00	0.00	1.00
Number of Ratings	6.98	10.66	1.00	3.50	16.00
Total Residential Installs	1579.18	3487.32	0.00	410.00	3500.00
Years Installing Solar	9.02	6.01	3.00	8.00	15.00
N	168				

Table 11: Effect of Additional Bids on Purchases

	(1)	(2)	(3)	(4)
% Increase in purchases if buyers obtain an additional bid (relative to 2 bids)	34.11***	35.88***	38.18***	36.16***
	(5.452)	(5.793)	(7.231)	(10.84)
State_FE	Yes	No	No	No
Quarter_FE	Yes	No	No	No
State-Time FE	No	Yes	Yes	No
Zipcode FE	No	No	Yes	No
Zipcode-Time FE	No	No	No	Yes
N	10545	10545	10545	10545

This table uses the regression estimates from equation 1 to calculate the percentage increase in solar purchases that results from providing buyers an additional bid, relative to the case where buyers receive two bids. This is calculated as $\frac{\mathbb{E}[1[Purchase_i]|bids=3]}{\mathbb{E}[1[Purchase_i]|bids=2]} * 100$. Delta method standard errors are in parentheses. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 12: Regressions of Bid Price ($/watt) on Order of Bid

	(1) Price ($/watt)	(2) Price ($/watt)
Order of Bid	0.00263**	0.000374
	(0.00133)	(0.00122)
Panel Quality Controls	No	Yes
NBids Control	Yes	Yes
System Size Control	Yes	Yes
State FE	Yes	Yes
Time FE	Yes	Yes
Installer FE	Yes	Yes
N	40565	40565
R^2	0.449	0.537

Panel quality controls include panel quality (rated by EnergySage), and a dummy for if the system includes a microinverter or power optimizer. All standard errors are listed in parenthesis. * indicates $p < 0.10$, ** indicates $p < 0.05$, and *** indicates $p < 0.01$

Table 13: Robustness Checks: Demand Estimates with Additional Demographic Controls

	(1)	(2)	(3)	(4)
λ	0.294^{***}	0.289^{***}	0.286^{***}	0.286^{***}
	(0.058)	(0.059)	(0.059)	(0.059)
Price ($/watt)	-0.794^{***}	-0.788^{***}	-0.786^{***}	-0.792^{***}
	(0.108)	(0.109)	(0.111)	(0.110)
Income and Educ. Controls	No	Yes	Yes	Yes
Political Controls	No	No	Yes	Yes
Age, Race, and HH Size Controls	No	No	No	Yes
Observations	10,545	10,545	10,545	10,545

Notes: The first column contains estimates for the baseline model. Zip code level demographic data was collected from the 2016 American Community Survey. Controls include median household income, percentage of adults that are college graduates, median age, percentage of white residents, and average household size. I also control for political preferences using the county-level democratic vote share in the 2016 presidential election. Each specification includes hardware quality dummies as well as quarter, state, and quality group fixed effects.

Table 14: Demand Specification with Experience Effects

λ	0.278^{***}
	(0.059)
Price ($/watt)	-0.802^{***}
	(0.109)
Cumulative Sales on Platform	-0.002
	(0.002)
Equipment Quality Controls	Yes
State FE	Yes
Time-Period FE	Yes
Quality Group FE	Yes
Size Control, Large Project FE	Yes
# Quality Groups	10
Observations	10,545

Notes: Equipment quality controls include premium panel, premium plus panel, and microinverter dummies. Cumulative Sales is the number of total auctions the installer has won on EnergySage platform at the time the bid was placed.

Figure 10: Histogram - Number of Bids per Project

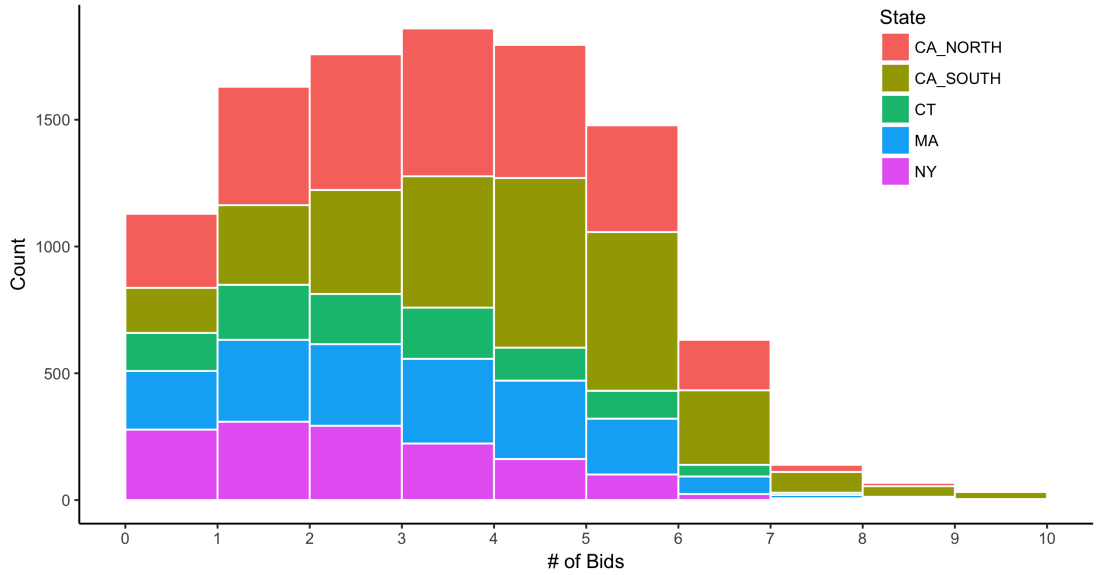


Figure 11: Number of Distinct Installers Submitting Bids on the Platform

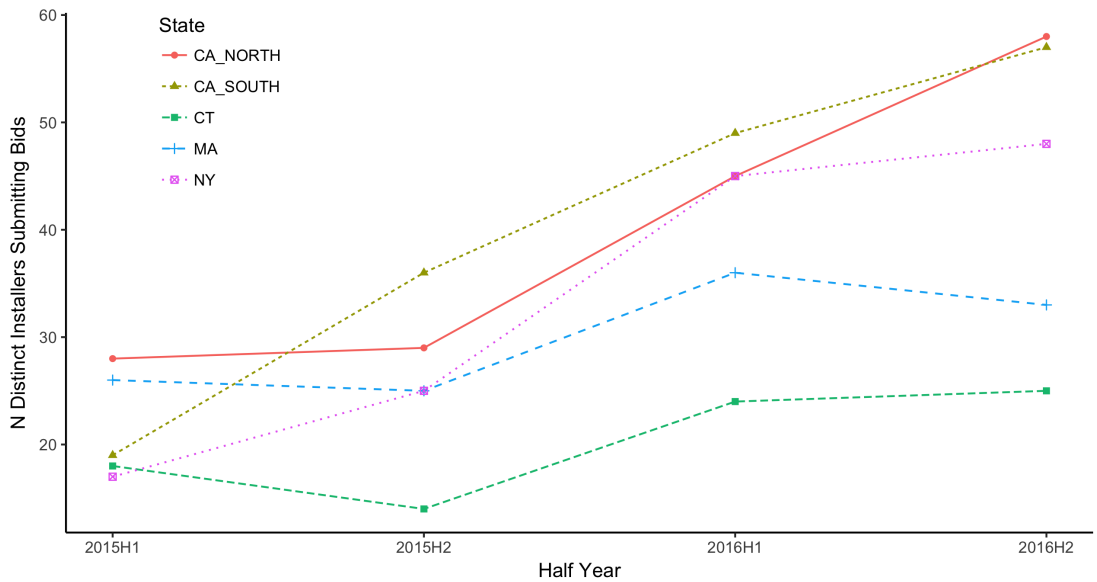


Figure 12: 2016 Q1 Price Distributions by State

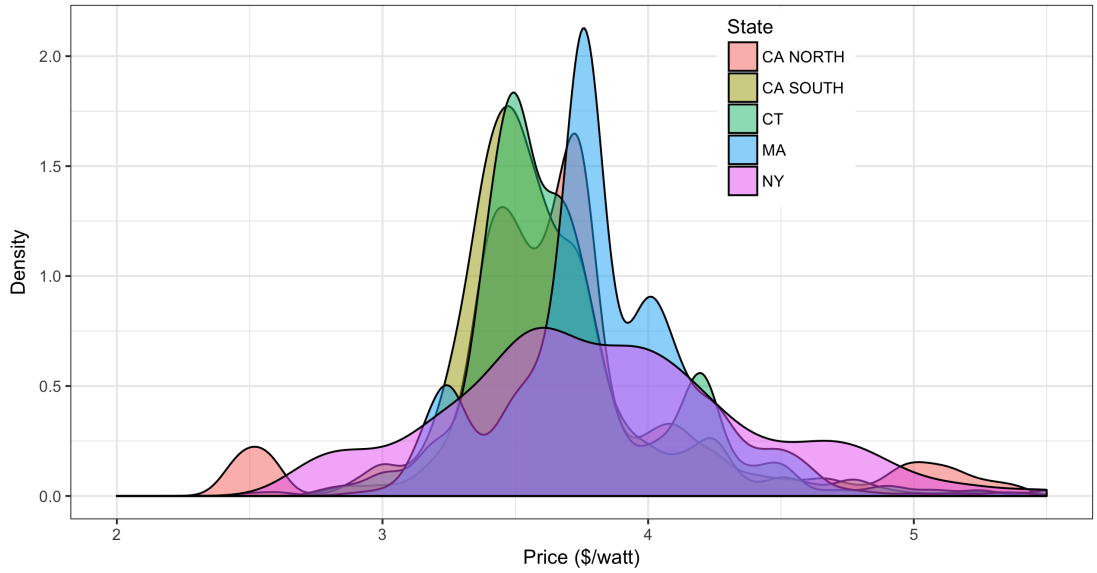


Figure 13: Average Bid Prices over Time

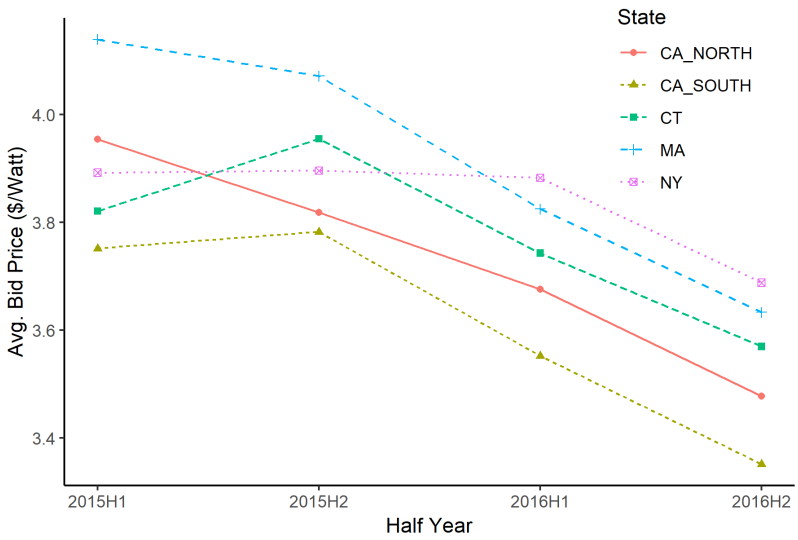
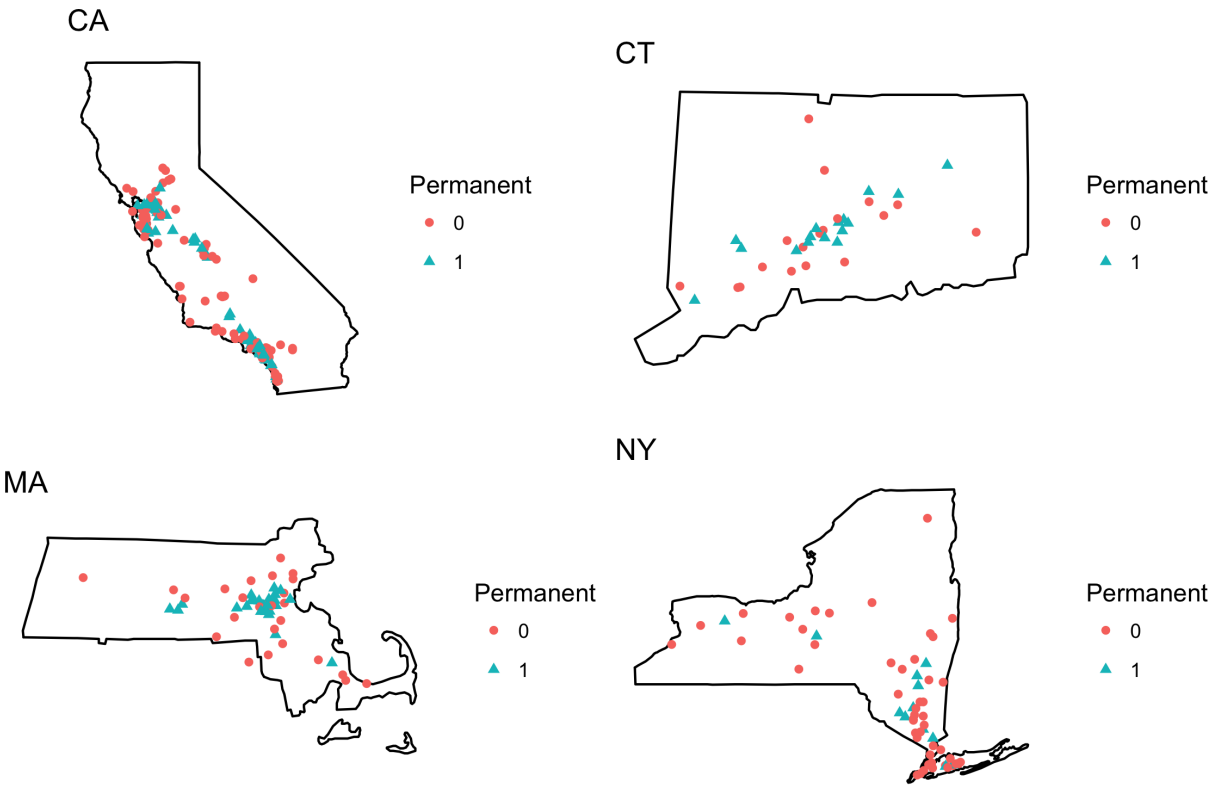
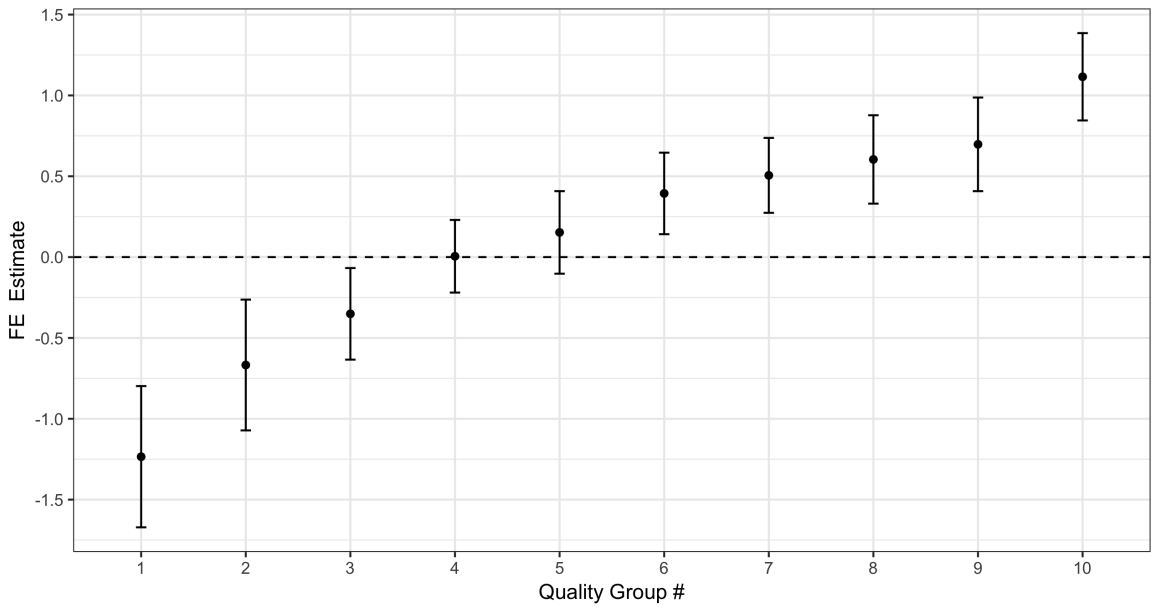


Figure 14: Imputed Installer Locations



Notes: Estimated installer locations are the centroid of all project locations that the installer submitted bids for throughout the sample. The red circles represent "transient" sellers that submitted fewer than 100 total bids, and the blue triangles represent permanent sellers.

Figure 15: Quality Group Fixed Effects Estimates



Notes: This figure plots the point estimates for each of the quality group fixed effects from the preferred demand specification (column 1 of Table 4). The bars indicate 95% confidence intervals.

Table 15: Robustness Checks: Alternative Demand Specifications

	(1)	(2)	(3)	(4)
λ	0.294^{***}	0.285^{***}	0.279^{***}	0.292^{***}
	(0.058)	(0.058)	(0.060)	(0.059)
Price ($/watt)	-0.794^{***}	-0.750^{***}	-0.781^{***}	-0.783^{***}
	(0.108)	(0.110)	(0.109)	(0.110)
State Incentive ($/watt)		0.445^{***}		
		(0.172)		
Loan Offered		-0.194^{***}		
		(0.060)		
$>$ 6 Months on Platform				-0.021
				(0.081)
Equipment Quality Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	No	Yes
Time-Period FE	Yes	Yes	Yes	No
Quality Group FE	Yes	Yes	Yes	Yes
State-Year FE	No	No	Yes	No
Size Control, Large Project FE	Yes	Yes	Yes	Yes
# Quality Groups	10	10	10	10
Observations	10,545	10,545	10,545	10,545

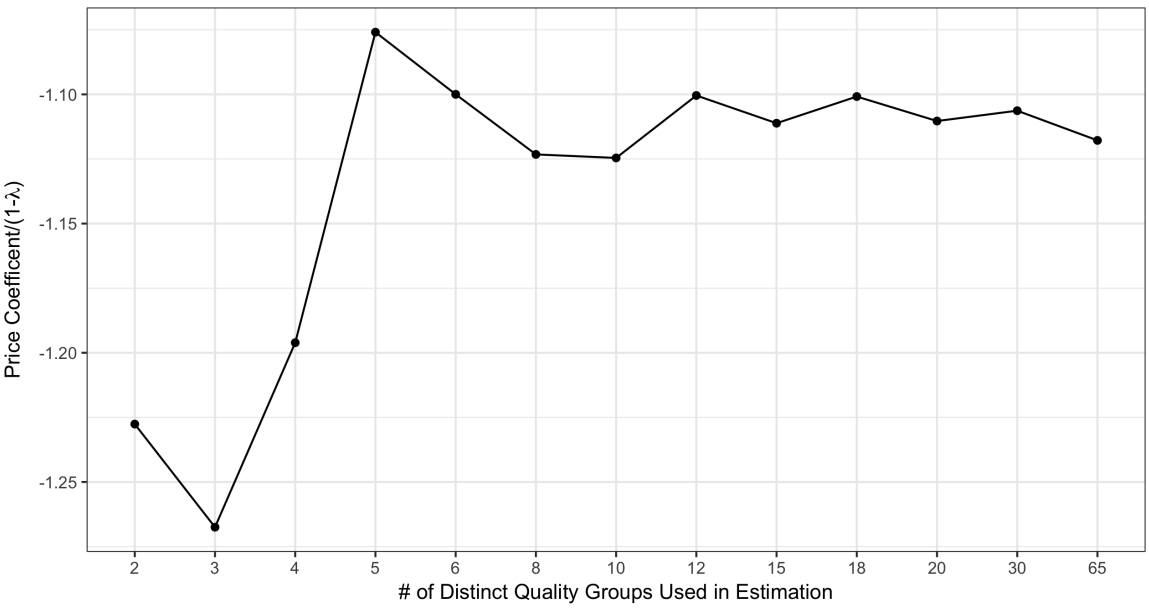
Notes: Equipment quality controls include premium panel, premium plus panel, and microinverter dummies. State incentives are reported by EnergySage and include both rebates and tax credits. NY was the only state that had an incentive policy that varied over the sample period. '>6 months on the Platform' is dummy determined by installer experience on EnergySage.

Table 16: Robustness Checks: Alternative Demand Specifications 2

	Baseline	K-Means	Iterated
	(1)	(2)	(3)
λ	0.294***	0.275***	0.254***
	(0.058)	(0.059)	(0.061)
Price ($/watt)	−0.794***	−0.820***	−0.835***
	(0.108)	(0.110)	(0.112)
Equipment Quality Controls	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Time-Period FE	Yes	Yes	Yes
Size Control, Large Project FE	Yes	Yes	Yes
# Quality Groups	10	10	10
Observations	10,545	10,545	10,545

Notes: Equipment quality controls include premium panel, premium plus panel, and microinverter dummies. The 'K-means' specifiactions uses, a k-means clustering algorithm to sort the sellers into groups based off the estimated firm FEs. The 'Iterated' specification updates the quality grouping after each estimation round based off of each firm's average residuals, and continues to update the groupings until convergence.

Figure 16: Robustness: Alternative Quality Group Specifications



Notes: This figure shows how the ratio of the estimated price coefficient over $1 - \lambda$ changes if different numbers of quality groups are used in estimation. The case with 65 quality groups corresponds to the firm fixed effects specification.

Table 17: Marginal Cost Regressions

	Marginal Cost ($/watt)
Premium Panel	0.115***
	(0.005)
Premium Plus Panel	0.919***
	(0.011)
Microinverter	0.082***
	(0.006)
Size (KW)	−0.029***
	(0.001)
Quality Group (1-3)	0.017**
	(0.007)
Quality Group (4-7)	−0.121***
	(0.007)
Quality Group (8-10)	−0.286***
	(0.008)
State FE	Yes
Time-Period FE	Yes
Observations	40,575
R^2	0.263
Adjusted R^2	0.263
Residual Std. Error	0.427 (df = 40556)
F Statistic	803.892*** (df = 18; 40556)

Notes: 'Transient Sellers' are the omitted quality group. 'Premium' and 'Premium Plus' panel brands are determined by EnergySage based off of panel efficiency, warranty, and other factors.

Table 18: Markup and MC Estimates - Alternative Project Classifications

	(1)	(2)	(3)	(4)	(5)
Mean Markup	1.42	1.43	1.42	1.48	1.43
Mean Marginal Cost	2.19	2.19	2.19	2.14	2.19
Mean Own-Price Elasticity	-1.79	-1.78	-1.79	-1.72	-1.78

Notes: All markups and marginal costs are recorded in dollars per watt. The first column reports summary statistics for the estimates under the baseline specification of project types. Column two allows for three projects size categories instead of two. The third column uses shorter time-periods, specifically, each time period is two months long instead of three. The fourth column uses 6-month time periods. Finally, the fifth column allows for more geographic heterogeneity by splitting NY and CA into rural and urban areas.

Figure 17: Estimated Marginal Costs and Markups by State over Time

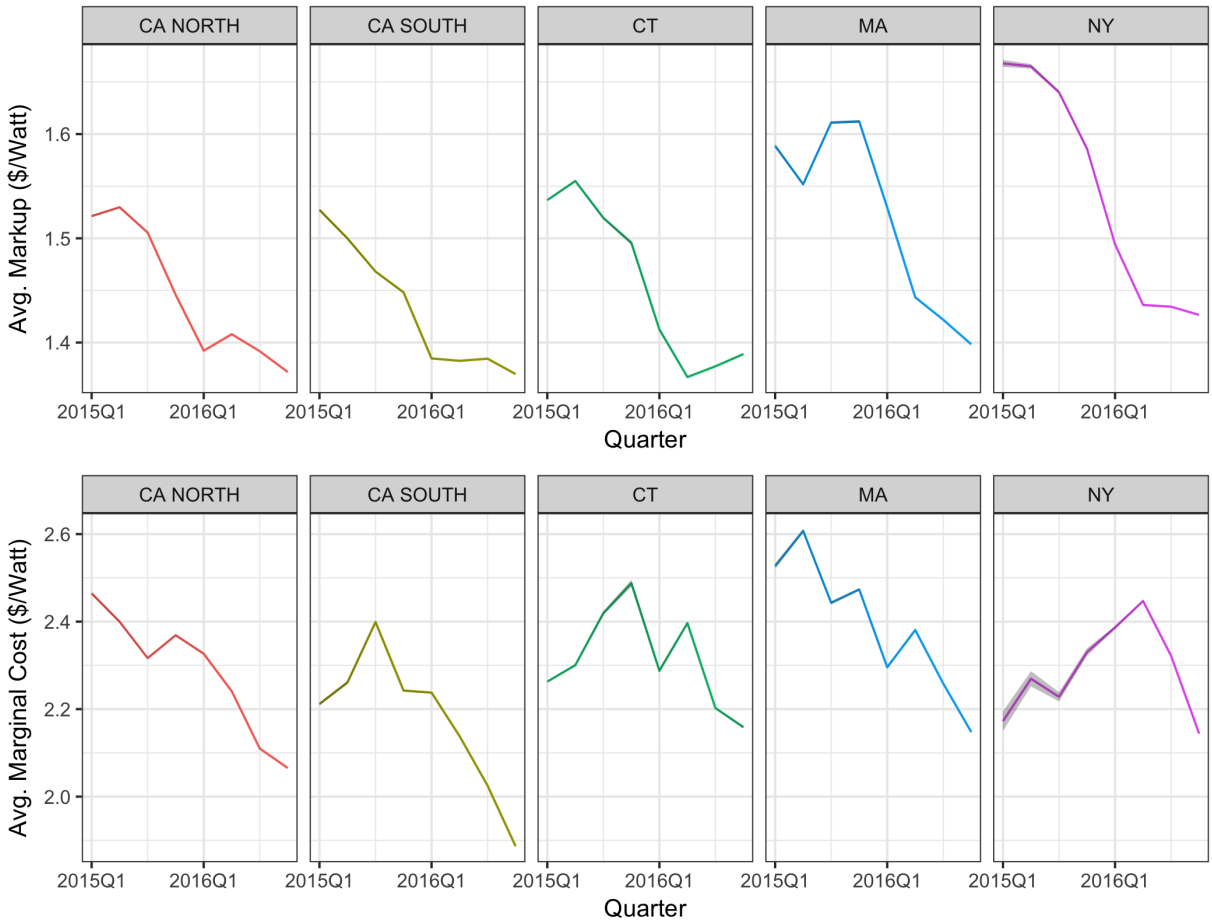
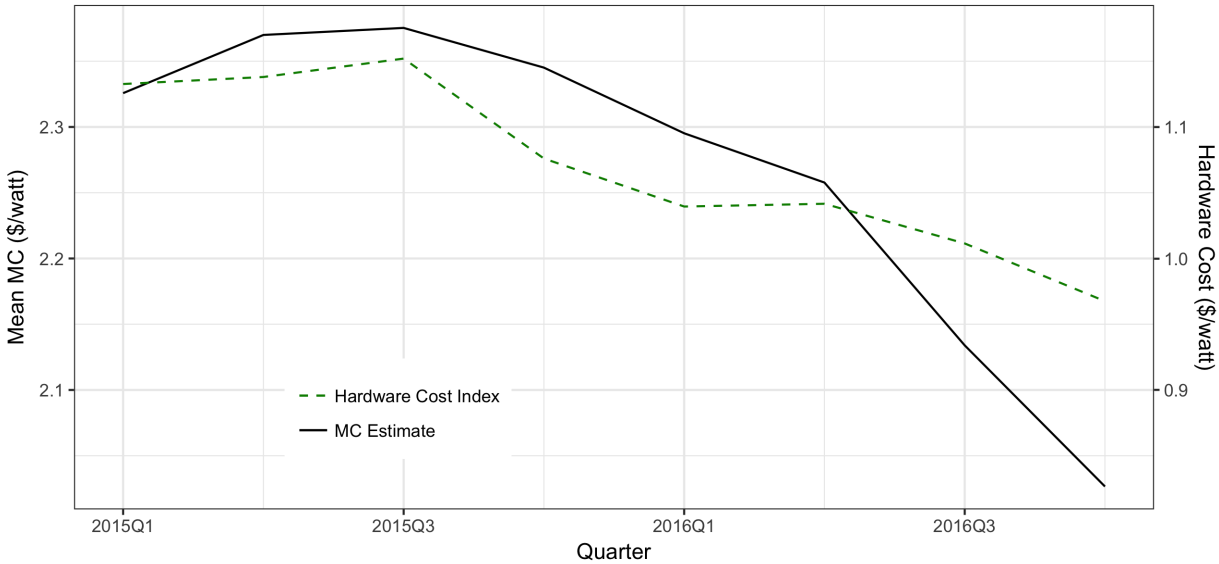
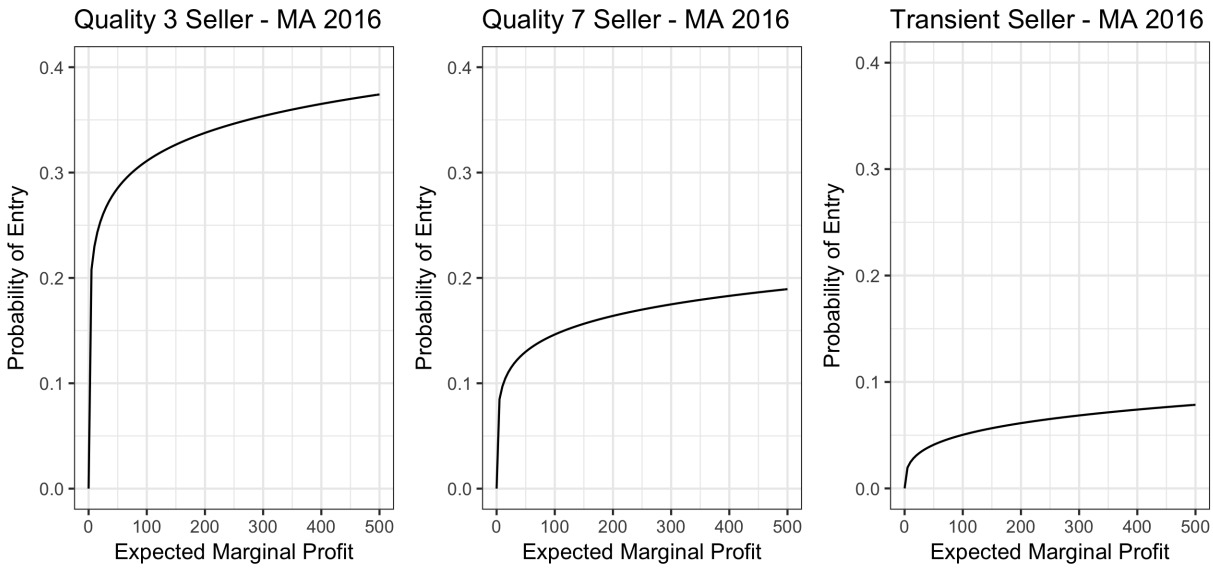


Figure 18: Estimated Average Marginal Cost and Bloomberg Hardware Cost Index



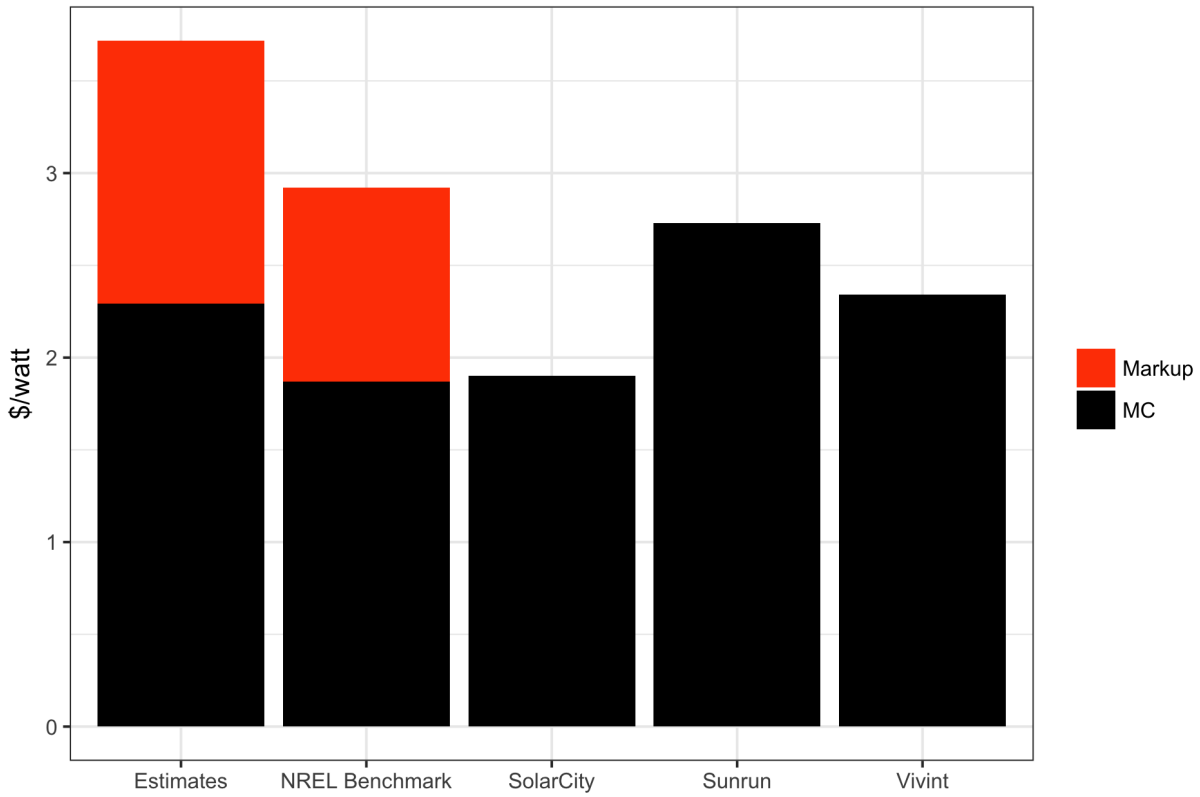
Notes: Hardware costs were obtained from Bloomberg New Energy Finance. The hardware index plotted above is calculated by summing Bloomberg's multicrystaline silicon module overall avg spot price and the PV Insights residential inverter price index (also collected from Bloomberg).

Figure 19: Predicted Probabilities of Entry



Notes: This figures plots the predicted probabilities of entry for three seller types at varying levels of marginal profits ($). The entry probabilities are determined using the estimated entry cost parameters in Table 20.

Figure 20: Comparison of Marginal Costs to Existing Estimates



Notes: The first bar plots the average marginal cost and markup estimates from the structural model for 2016Q1. The second bar plots estimates from the National Renewable Energy Laboratory (NREL) 2016 Q1 benchmark analysis. The last three bars plot marginal costs reported in quarterly reports by large publicly traded installers, Solar City, Vivint, and Sunrun.

Table 19: Summary of Estimated Markups and Marginal Cost by Location and Time-Period

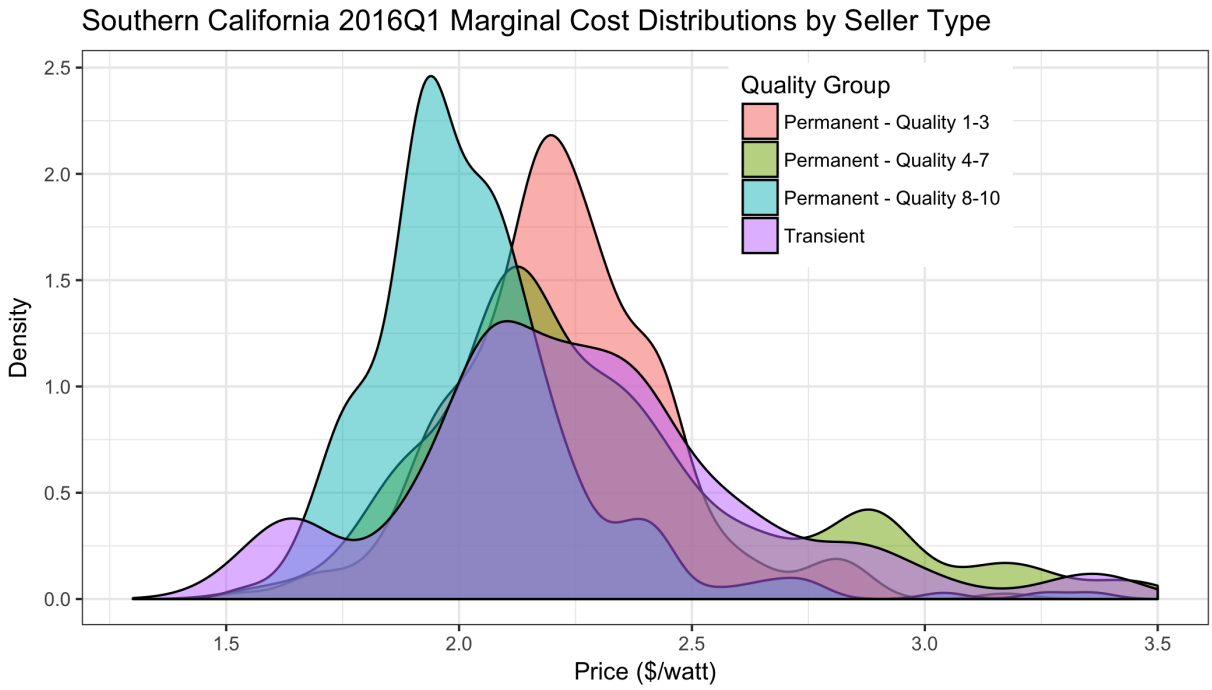
Panel A: Marginal Costs ($/watt)

State	Time-Period	Mean	SD
CA NORTH	2015 Q1	2.46	0.42
CA NORTH	2015 Q2	2.4	0.45
CA NORTH	2015 Q3	2.32	0.51
CA NORTH	2015 Q4	2.37	0.47
CA NORTH	2016 Q1	2.33	0.58
CA NORTH	2016 Q2	2.24	0.45
CA NORTH	2016 Q3	2.11	0.44
CA NORTH	2016 Q4	2.07	0.46
CA SOUTH	2015 Q1	2.21	0.42
CA SOUTH	2015 Q2	2.26	0.6
CA SOUTH	2015 Q3	2.4	0.48
CA SOUTH	2015 Q4	2.24	0.4
CA SOUTH	2016 Q1	2.24	0.39
CA SOUTH	2016 Q2	2.14	0.36
CA SOUTH	2016 Q3	2.02	0.44
CA SOUTH	2016 Q4	1.89	0.35
CT	2015 Q1	2.26	0.37
CT	2015 Q2	2.3	0.39
CT	2015 Q3	2.42	0.57
CT	2015 Q4	2.49	0.61
CT	2016 Q1	2.29	0.43
CT	2016 Q2	2.4	0.49
CT	2016 Q3	2.2	0.41
CT	2016 Q4	2.16	0.46
MA	2015 Q1	2.53	0.76
MA	2015 Q2	2.61	0.66
MA	2015 Q3	2.44	0.51
MA	2015 Q4	2.47	0.45
MA	2016 Q1	2.3	0.44
MA	2016 Q2	2.38	0.46
MA	2016 Q3	2.26	0.48
MA	2016 Q4	2.15	0.43
NY	2015 Q1	2.17	0.71
NY	2015 Q2	2.27	0.67
NY	2015 Q3	2.23	0.7
NY	2015 Q4	2.33	0.73
NY	2016 Q1	2.39	0.66
NY	2016 Q2	2.45	0.68
NY	2016 Q3	2.32	0.65
NY	2016 Q4	2.14	0.62

Panel B: Markups ($/watt)

State	Time-Period	Mean	SD
CA NORTH	2015 Q1	1.52	0.1
CA NORTH	2015 Q2	1.53	0.09
CA NORTH	2015 Q3	1.51	0.08
CA NORTH	2015 Q4	1.45	0.07
CA NORTH	2016 Q1	1.39	0.06
CA NORTH	2016 Q2	1.41	0.07
CA NORTH	2016 Q3	1.39	0.06
CA NORTH	2016 Q4	1.37	0.05
CA SOUTH	2015 Q1	1.53	0.14
CA SOUTH	2015 Q2	1.5	0.14
CA SOUTH	2015 Q3	1.47	0.11
CA SOUTH	2015 Q4	1.45	0.12
CA SOUTH	2016 Q1	1.38	0.08
CA SOUTH	2016 Q2	1.38	0.09
CA SOUTH	2016 Q3	1.38	0.08
CA SOUTH	2016 Q4	1.37	0.07
CT	2015 Q1	1.54	0.13
CT	2015 Q2	1.55	0.13
CT	2015 Q3	1.52	0.13
CT	2015 Q4	1.5	0.13
CT	2016 Q1	1.41	0.09
CT	2016 Q2	1.37	0.08
CT	2016 Q3	1.38	0.08
CT	2016 Q4	1.39	0.08
MA	2015 Q1	1.59	0.21
MA	2015 Q2	1.55	0.16
MA	2015 Q3	1.61	0.19
MA	2015 Q4	1.61	0.23
MA	2016 Q1	1.53	0.16
MA	2016 Q2	1.44	0.1
MA	2016 Q3	1.42	0.1
MA	2016 Q4	1.4	0.09
NY	2015 Q1	1.67	0.1
NY	2015 Q2	1.66	0.09
NY	2015 Q3	1.64	0.1
NY	2015 Q4	1.59	0.12
NY	2016 Q1	1.49	0.1
NY	2016 Q2	1.44	0.09
NY	2016 Q3	1.43	0.09
NY	2016 Q4	1.43	0.08

Figure 21: Non-Parametric Marginal Cost Distributions



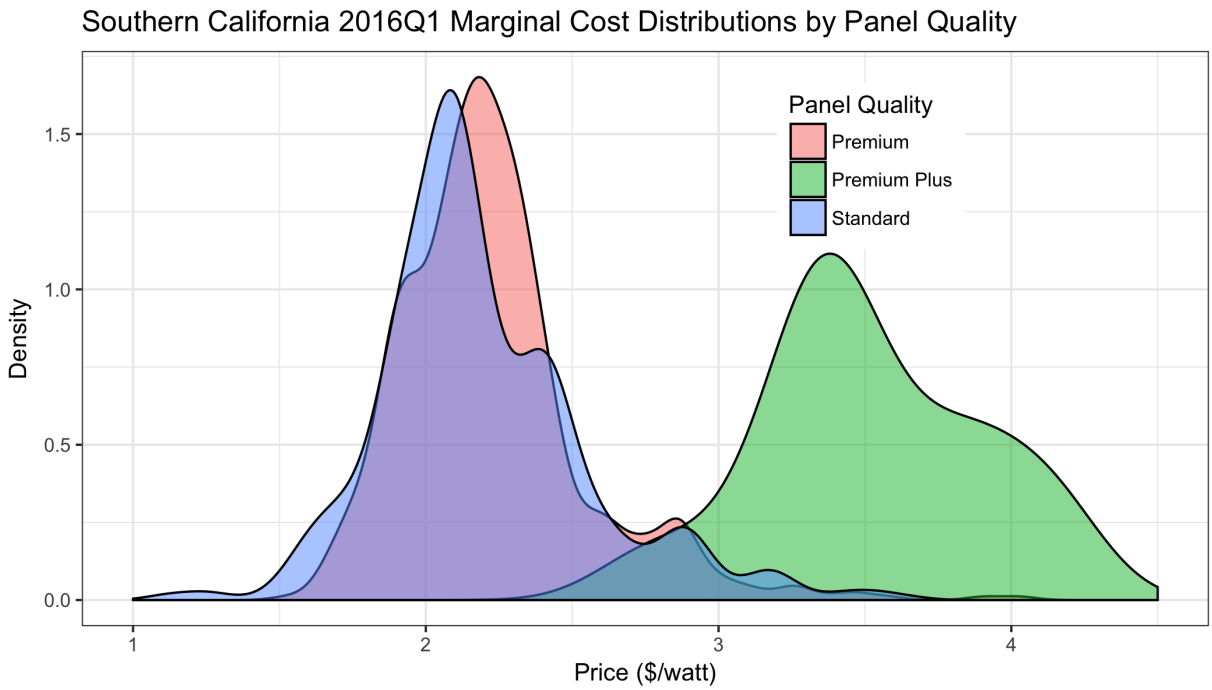
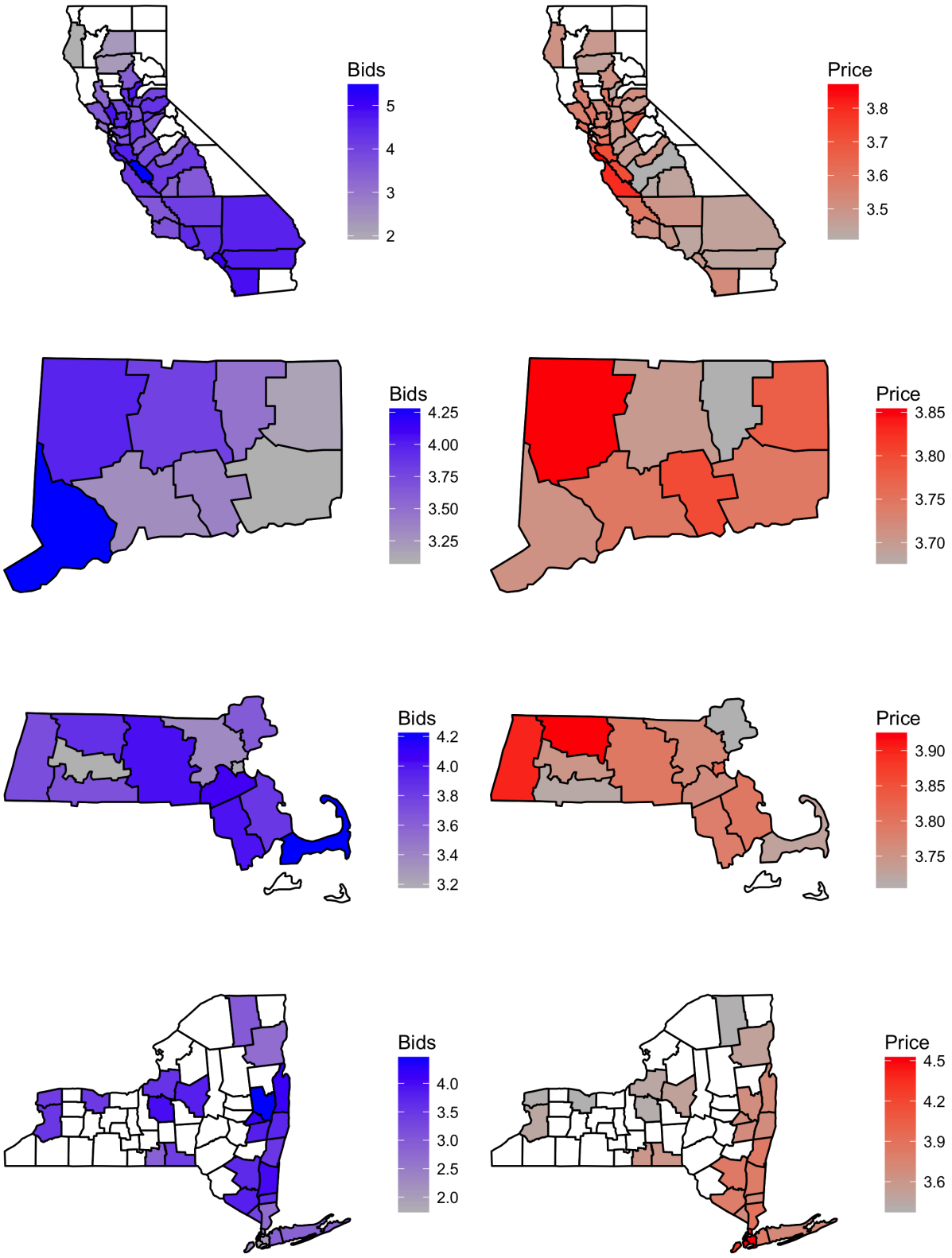


Table 20: Entry Cost Estimates

	Estimates
μ - Transient	13.196***
	(0.591)
μ - Quality Group 1	7.729***
	(0.329)
μ - Quality Group 2	10.574***
	(0.475)
μ - Quality Group 3	6.149***
	(0.183)
μ - Quality Group 4	9.047***
	(0.294)
μ - Quality Group 5	10.522***
	(0.366)
μ - Quality Group 6	8.812***
	(0.261)
μ - Quality Group 7	11.379***
	(0.400)
μ - Quality Group 8	12.886***
	(0.497)
μ - Quality Group 9	12.815***
	(0.477)
μ - Quality Group 10	10.169***
	(0.256)
σ - Transient	7.102***
	(0.482)
σ - Permanent	9.349***
	(0.640)
State FE	Yes
Year FE	Yes
Observations	280,670

Notes: 'Permanent Sellers' include all sellers in quality groups 1-10.

Figure 22: Average Number of Bids Received and Bid Prices ($/watt) by County



Notes: Maps only include counties that had at least 10 potential projects during the sample.