Unequal Uptake: Assessing Distributional Disparities in the Residential Solar Market

Jackson Dorsey[*]

Derek Wolfson[†]

July 31, 2024

Abstract

We examine technology adoption and consumer welfare disparities across demographic groups using data from an online solar photovoltaic (PV) marketplace. Low-income households are 25% less likely to purchase solar through the platform and obtain 53% lower expected consumer surplus than high-income households. Moreover, Black and Hispanic households are relatively less likely to purchase solar through the platform and obtain lower consumer surplus than White and Asian households. We develop a method to decompose the drivers of consumer welfare disparities between demographic groups. Differences in demand fully account for the consumer surplus disparities between high- and low-income households and between White and Hispanic households. However, supply-side factors explain 37% of the consumer surplus gap between White and Black households. Black households get relatively fewer bids and face higher prices, and installers have higher implied costs to serve them. Lastly, we assess counterfactuals that offer targeted price discounts to certain demographic groups.

JEL Codes: D22, D44, D63, H23, L11, Q40, Q41, Q48
Keywords: solar PV, renewable energy, technology adoption, new energy technologies, inequality, equity, distributional analysis

[*]Department of Economics, University of Texas at Austin (e-mail: jackson.dorsey@austin.utexas.edu)

[†]Ironcloud Metalworks (e-mail: derekwolfson@gmail.com)

We thank Spencer Fields for providing instrumental help with data access. We also thank Catie Hausman, Todd Gerarden, Arik Levinsohn, and seminar participants at the AERE Summer Conference, the NBER Conference on Distributional Consequences of New Energy Technologies, UC Berkeley, and Indiana University for helpful comments and suggestions. Finally, we thank Moyan Li and Aishwarya Agarwal for excellent research assistance.

1 Introduction

Over the past decade, billions of dollars in subsidies for electric vehicles, solar PV panels, and energy efficiency retrofits have accelerated clean energy investment. However, many subsidy programs, such as U.S. clean energy tax credits, disproportionately benefit higher-income households (Borenstein and Davis, 2016). This is because Black, Hispanic, and low-income households are less likely to adopt clean energy technologies, such as rooftop solar PV systems (Sunter et al., 2019; O'Shaughnessy et al., 2021; Reames, 2020). To address these inequities, Congress and the Environmental Protection Agency (EPA) have introduced programs aimed to increase adoption among low-income households (Environmental Protection Agency, 2023; Internal Revenue Service, 2023).[1]

In markets for new energy technologies, such as home energy efficiency retrofits, residential battery storage, and rooftop solar PV, contractors often customize and price projects individually. Consequently, sellers may use household or neighborhood characteristics, such as income and race, to determine which customers to serve or to adjust bid prices. Thus, adoption disparities may stem from both supply-side and demand-side factors. On the demand side, different consumer groups may purchase a new technology at varying rates because of differences in willingness to pay or other underlying preferences. On the supply side, firms can contribute to disparities in equilibrium adoption by changing their service offerings or bidding behavior across consumer demographic groups. Our study aims to disentangle the relative importance of these supply and demand factors in explaining technology adoption and welfare disparities.

Figure 1 illustrates how supply-side and demand-side factors contribute to disparities in clean technology adoption and consumer welfare. Figure 1a and Figure 1b depict the demand and supply for rooftop solar PV systems among high-income and low-income households, respectively. We see that low-income households have more elastic demand and also face higher prices. In practice, low-income households may face higher prices because they tend to purchase smaller solar arrays (with higher marginal costs) and live, on average, further away from solar installers.[2]

Unsurprisingly, Figure 1a and Figure 1b demonstrate that, in this example, high-income households will purchase more solar PV systems and receive greater consumer surplus compared to their lower-income neighbors. Figure 1c illustrates the difference in consumer surplus achieved by high-income households relative to low-income households, which we call the *consumer surplus gap*.

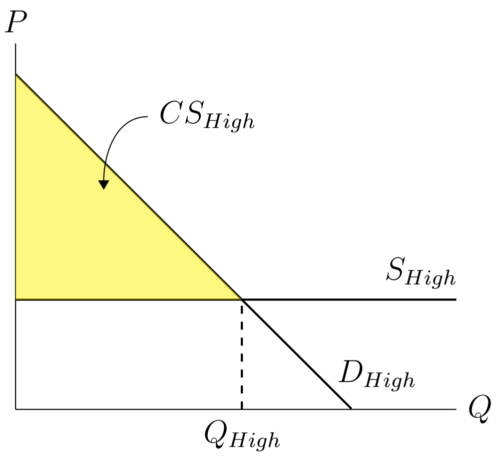
[1]U.S. Inflation Reduction Act (IRA) formulated a provision for low- and middle-income individuals and the EPA offered grants through the Solar for All program.

[2]Solar PV installations often feature economies-of-scale where the per-unit cost of an installation tends to decline with the size of the system (Dorsey, 2024).

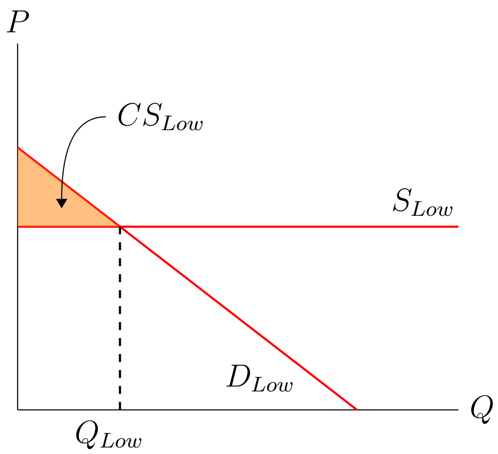
To effectively address distributional inequities, evaluating both the supply and demand side of the market is essential. For example, consider a naïve subsidy offered to low-income households equal to $S_{Low} - S_{High}$ in Figure 1. Such a subsidy would equalize effective solar prices across income groups. However, Figure 1 indicates that price equalization alone is insufficient to achieve parity in adoption and consumer welfare. When low-income households' demand is relatively elastic, equilibrium adoption and consumer surplus for low-income households will lag behind high-income households, even with equal prices. In such cases, more ambitious targeted subsidies would be needed if policymakers seek to eliminate gaps in consumer surplus and adoption.

Figure 1d further decomposes the consumer surplus gap into three components: a *demand*

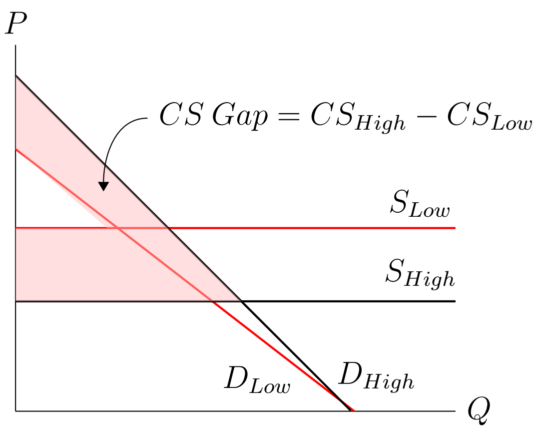
Figure 1: Consumer Surplus for High- and Low-Income Households in the Solar PV Market



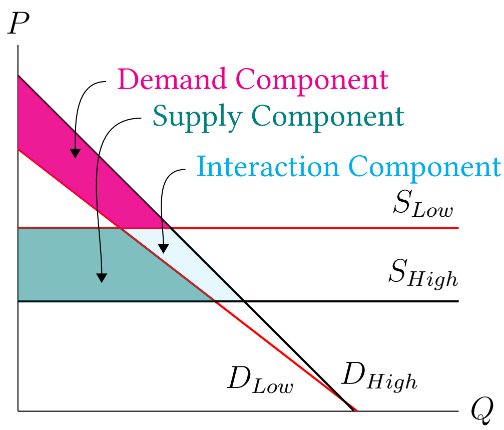
(a) High-Income Consumer Surplus



(b) Low-Income Consumer Surplus



(c) Consumer Surplus Gap



(d) Decomposition of CS Gap

Notes: Panel (a) shows the demand and supply curves of a high-income household. The shaded area represents the consumer surplus. The shaded area in Panel (b) shows the consumer surplus obtained by low-income households. Panel (c) shows the difference in consumer surplus between the two groups. Finally, Panel (d) decomposes this difference into three components - a demand component, a supply component, and an interaction component.

component, a *supply component*, and an *interaction component*. The demand component represents the portion of the consumer surplus gap that would persist if both income groups faced the same supply curve (the low-income supply curve). The supply component signifies the part of the consumer surplus gap that would remain if both groups had the same demand (the low-income demand curve) but faced different supply curves. Lastly, the interaction component depicts the residual piece of the consumer surplus gap arising from simultaneous supply and demand curve shifts.

Understanding the underlying components of consumer surplus disparities can help inform the choice of policy instruments used to address inequities. If the supply component primarily explains disparities, policies that incentivize firm investment or remove barriers to entry in underserved markets may be fruitful. Examples of supply-side programs include grants, loans, tax abatement, reducing regulatory or permitting costs, workforce development, and other initiatives designed to reduce firms' investment costs and effectively increase supply. Demand-side policies may be relatively appealing if the demand component primarily drives the consumer surplus gap. Such policies include Pigouvian taxes, product-market subsidies, or behavioral nudges.

Our study uses detailed data from a major online marketplace to investigate socioeconomic and demographic disparities in the residential solar PV market. While there is growing evidence of income and racial disparities in adopting clean energy technologies, relatively little research documents the fundamental mechanisms contributing to observed adoption inequities. Our paper aims to fill this gap in the literature by addressing four main objectives: (1) use new data to quantify the gap in solar PV adoption across household income and racial/ethnic demographics, (2) develop a structural model to estimate disparities in consumer welfare in the solar PV market, (3) use our model to decompose the mechanisms that explain the measured disparities in consumer welfare, and (4) evaluate the welfare impacts of offering targeted price discounts to disadvantaged groups.

We begin with a descriptive analysis, showing that solar PV purchase rates are indeed lower for low-income and minority households in our data. Specifically, low-income households are 25% less likely to purchase a solar PV system through the online platform than high-income households.[3] Moreover, purchase rates for Black households are over 31% lower compared to White or API households. Similarly, the gap for Hispanic households is nearly 18%.[4]

Given that purchase choices are determined by the equilibrium behavior of both sellers and buyers, it is not clear whether disparities in the solar uptake across households are driven

[3]We define low-income households as those having a census-block group median income within the bottom quintile of our sample and high income as households in the top quintile.

[4]The platform close rate does not account for the possibility that adoption occurs offline. We also show a similar adoption gap arises if we compare responses to buyer exit surveys that ask whether the households decided to install solar (including offline purchases).

primarily by differences in available choice sets (*i.e.* supply) or heterogeneity in buyers' preferences (*i.e.* demand). Importantly, we find that low-income and Black households obtain relatively fewer bids from installers. For example, low-income households receive 19% fewer bids compared to high-income households. Additionally, we show that Black households receive 38% fewer bids relative to White households and that this disparity is primarily explained by buyers' neighborhoods and not buyers' names. On the other hand, Hispanic and API households tend to receive as many bids as White households with similar incomes. These findings underscore the considerable heterogeneity in choice sets available across socioeconomic and demographic groups.

We next develop a structural model to address our last three research objectives. In the model, each prospective buyer arrives at the platform, and sellers learn about the project's characteristics and their private installation cost. Given this information, sellers submit profit-maximizing bids. Then, the household chooses one of the bidding sellers or the outside option—choosing an off-platform installer or not installing a solar system. The auction allocation mechanism is a multi-attribute (beauty contest) auction in which buyers can consider factors other than prices, such as the installers' ratings, experience, and hardware quality. We use a similar estimation approach to Yoganarasimhan (2015), Krasnokutskaya et al. (2019), and Dorsey (2024) to estimate the model.

The model estimates reveal substantial disparities in consumer welfare across socioeconomic and demographic groups. For instance, high-income households receive over 50% higher consumer surplus than low-income households. Similarly, White and API households obtain over double the expected consumer surplus of Black and Hispanic households.[5]

Having estimated sizable disparities in consumer welfare, we build on and adapt methods from the labor economics literature (Oaxaca, 1973; Blinder, 1973) to empirically decompose the consumer surplus gap between two groups (*e.g.*, the Black-White consumer surplus gap). Our decomposition approach separates the consumer surplus gap into the three components illustrated in Figure 1d—a demand component, a supply component, and an interaction component. Intuitively, the demand component is measured by evaluating the counterfactual level of consumer surplus for each demographic group, holding the choice set available to each group of buyers fixed. The supply component measures the portion of the consumer surplus gap due to differences in the choice sets available to the two groups of buyers—including differences in the number of bids received or the distribution of bid prices. More concretely, we derive the supply component by evaluating a set of counterfactuals that allows the choice

[5] Each buyer's expected consumer surplus is equal to their willingness to pay for the full set of installation bids that they receive through the platform. Consequently, consumer surplus will generally increase with the number of bids received, decrease with bid prices, increase with the quality of sellers making bids, and decrease with the buyer's price sensitivity.

sets to vary across the two groups while holding the buyers' preferences fixed.[6] Finally, the interaction component is the remaining portion of the consumer surplus gap not unilaterally accounted for by the demand and supply components. To our knowledge, we are the first paper to decompose consumer welfare disparities using this approach.[7]

We first decompose the consumer surplus disparity between low-income and high-income households and find that the demand component accounts for the entire consumer surplus gap between these groups. This implies that a similar disparity in welfare would persist even if these groups faced similar choice sets. We next decompose the consumer surplus gap between race and ethnic groups. Our decomposition yields a similar result when we compare White and Hispanic households—Hispanic households would obtain roughly the same consumer surplus if they faced identical choice sets as White households.

In contrast, the supply component explains 37% of the gap in consumer surplus between Black and White households. In particular, Black households receive higher average bid prices and fewer bids, contributing substantially to the consumer surplus disparity across these groups. Our supply-side estimates further indicate that installers face higher implied marginal costs of serving Black households, which helps explain why these households face these higher prices, have limited choice sets, and consequently get lower consumer surplus.[8]

Targeted subsidies or tax credits have recently emerged as a popular policy tool for addressing distributional disparities in solar PV and related markets. Our last set of results investigates the effect of targeted price discounts in mitigating consumer surplus disparities. Our analysis reveals that households in the lowest income quintile would need to be offered relatively large price discounts to achieve the same expected consumer surplus as households in the highest income quintile. In particular, we find that observed bid prices (after existing tax incentives) submitted to low-income households would have to fall by 27%—$0.57 per watt—to achieve parity in consumer surplus with their high-income counterparts.

Overall, our results highlight large distributional differences in adoption and welfare among our sample of solar PV buyers. Moreover, we find that these disparities can be attributed to a combination of heterogeneous preferences and fundamental differences in supply across households. In the short run, low-income households would require substantial price discounts to reach the level of consumer surplus obtained by their higher-income counterparts. While we do not explicitly investigate supply-side policies, our decomposition analysis sug-

[6]The supply-side component of the decomposition incorporates everything that goes into defining the equilibrium choice sets, including sellers' strategic pricing with respect to a buyer's expected preferences.

[7]While the intuition for our approach is similar to the Oaxaca-Blinder decomposition used to measure discrimination in the labor market, the implementation differs because the counterfactual outcomes in our setting are determined using a non-linear random utility model instead of an OLS regression. In addition, other papers in the industrial organization literature have carried out decompositions that are similar in spirit to our decomposition (e.g., Olley and Pakes, 1996).

[8]The marginal costs implied by our model are the sum of explicit costs (*e.g.* labor and materials) and implicit costs (*e.g.* distaste for serving minority buyers, the time cost of traveling across town, etc.).

gests that supply-side policies that reduce firms' costs of entry in majority Black communities may be complementary to consumer subsidies in reducing disparities in adoption and consumer welfare.[9]

Our paper relates to a broader theoretical literature on discrimination and inequality beginning with Becker (1971) and empirical work that documents inequalities in the various markets. For instance, in the labor market, Bertrand and Mullainathan (2004) provide evidence that firms discriminate against individuals based on their name in remote interactions. In particular, the authors experimentally demonstrate resumes submitted with Black-sounding names receive fewer job interview call-backs than identical resumes with White-sounding names (*e.g.,* Lakisha Washington versus Emily Walsh). More recent research corroborates these findings in the labor market (*e.g.,* Kline and Walters (2021)). We contribute to this literature by documenting whether firms' bidding behavior varies across race and socioeconomic groups in an online solar PV marketplace. Moreover, we uncover new mechanisms for price and adoption disparities across demographic groups using detailed bidding data. Unlike much of the related literature, we find that racial bidding disparities are primarily explained by neighborhood location rather than buyers' names.

Particularly relevant to this paper is a suite of research documenting discrimination against minorities in terms of price and access in online markets such as Airbnb (Edelman et al., 2017), eBay (Ayres et al., 2015), and ridesharing (Ge et al., 2020). Outside the online domain, researchers document that low-income individuals and minorities, especially Black individuals, pay more for goods and services in many sectors including electricity (Lyubich, 2020; Bednar and Reames, 2020), groceries (Butters et al., 2022), vehicles (Ayres and Siegelman, 1995), and housing (Avenancio-León and Howard, 2022; Christensen and Timmins, 2022, 2023).

Lastly, our work contributes to a growing literature documenting the distributional consequences of energy policy, environmental policy, and the clean energy transition. Existing work has investigated the distributional impacts of carbon pricing (Metcalf, 2009), fossil fuel extraction dynamics (Blonz et al., 2023), renewable energy policy (Reguant, 2019), carbon capture (Waxman et al., 2023), electric vehicles (Jacqz and Johnston, 2023), and residential energy subsidies (Hahn and Metcalfe, 2021). In the residential solar market Nemet et al. (2017), Barbose and Darghouth (2023), and O'Shaughnessy et al. (2021) document considerable pricing variation in solar installations and Dauwalter and Harris (2023) estimates the distribution of environmental benefits from solar adoption. We build on this literature by estimating a model to evaluate the distribution of welfare in this market and to understand the underlying mechanisms that drive distributional disparities.

This paper proceeds as follows. In Sections 2 and 3, we introduce our data and provide sum-

[9]In the longer run, supply and demand-side policies may be complementary. For example, Gerarden (2023) shows that demand-side subsidies encouraged upstream investment on the supply side of the solar industry.

mary statistics and a descriptive analysis. We introduce our model and estimation strategy in Sections 4 and 5, and discuss the results of our model in Section 6. Section 7 concludes.

2 Data

The primary data for our analysis comes from the EnergySage online marketplace. We augment the EnergySage data with household characteristics and rich demographic data from the American Community Survey to investigate ethnic, racial and income disparities in bidding behavior and rooftop solar adoption.

2.1 Solar Auction Data

The EnergySage data contains a set of bid prices and consumer purchase choices for solar auctions on the online platform. Our main data set includes the bids to all households within the platform's 15 largest markets from 2017-2020—which includes 243,120 individual bids submitted to 56,011 potential buyers through the platform.[10]

EnergySage Inc. runs a quote aggregation platform that facilitates connections between potential solar customers and a network of solar PV installers. More specifically, the EnergySage platform enables households to conduct multi-attribute auctions to select installers for their projects. Multi-attribute auctions (also called beauty contest auctions) refer to a procurement mechanism in which each bidder submits a multi-dimensional bid that includes a price and a vector of other characteristics, such as solar panel brand or inverter type. The buyer then selects the winning bidder based on their preference over these multi-dimensional bids.[11]

Each auction includes several stages. First, consumers create an account on the platform's website and provide information, including the physical household address for the potential installation, a monthly electricity bill, and an indication of whether they obtained other solar installation bids off-platform. Second, registered installers receive a project notification, including details such as a Google Maps photo of the buyer's roof, the buyer's monthly electricity usage, and whether the buyer has other off-platform quotes. Installers then choose to submit a project quote to the buyer. A bid contains information about pricing, the system size, and specific hardware characteristics (*e.g.,* panel brand, panel ratings, inverter type, and brand, etc.). The platform also gives buyers details about the seller, such as their rating—stars on a scale from 1 to 5—and a description of their solar installation experience. Finally, after installers submit their bids, the consumer is free to select one of the quotes and complete the

[10]See Appendix Table A.1 for a listing of the markets included in this study.

[11]Multi-attribute auctions are related to scoring auctions but differ in that the auctioneer does not explicitly announce the choice rule ex-ante (*i.e.* weights on each characteristic), as they would in a scoring auction.

transaction or opt out and not purchase any of the offers.[12]

We access several key variables on buyers and sellers in the EnergySage data. First, we observe characteristics of each potential buyer, including the census block where the home is located, the household's average monthly electricity bill, and roof age. We further observe survey-based data from EnergySage regarding each household's preferences over equipment and financing and retrospective survey data on whether households adopted solar outside the platform. Second, we observe detailed information on the sellers' bids submitted to each buyer. The bid data includes the price, hardware specifics (*e.g.* panel brand, panel quality, etc.), the capacity of the solar array, and attributes of the seller (*e.g.* quality "star" ratings). Importantly, we observe a unique installer ID associated with each bid, so we can investigate how a particular installer's bidding behavior changes across projects. Importantly, we can also infer installers' approximate locations (See Appendix A.2.1 for more details). Finally, we observe which bid, if any, is selected by each buyer.

2.2 Household and Neighborhood Demographic Data

EnergySage did not collect consumer demographic information during the sample period of this study. However, they report each buyer's location at the census block level. We use this locational information to collect demographic characteristics of each household from data available in the 2017 American Community Survey (ACS). The census block is the smallest geographic unit in the US Census. Thus, we can merge precise information about each buyer's neighborhood demographics from the ACS.

The main variables we extract from the ACS are median household income and the racial and ethnic composition of their block group. We collect the racial and ethnic composition variables for homeowners only because owning a home is generally required to purchase a solar PV system.[13] We assign the median household income to each household for the entire census block group (*i.e.,* for both renters and owners) since income variables by home-ownership status are redacted at the block group level in the publicly available US Census data.

We use the ACS data to create binary measures of race and ethnicity following the approach used in Diamond et al. (2019). This approach leverages information about both the buyer's name and the buyer's home location, we discuss the details of this categorization in Appendix A.2.2.

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

[13]This choice helps us avoid potential mismeasurement introduced by differences in the racial or ethnic composition of neighborhoods for renters and buyers.

2.3 Summary Statistics

Households in our sample generally belong to higher-income block groups, with a median income of $103,000.[14] Despite this, there is considerable variation, with median block group incomes ranging from $11,625 to $250,000. In Appendix Table A.2, we report summary statistics separately by household income quintile. High-income households have larger monthly electricity bills and a stronger preference for cash purchases over loans, with non-response rates to optional onboarding questions decreasing with income.

Appendix Table A.3 shows that API and Pacific Islander (API) households account for 13.1% of the sample, Black households for 1.2%, Hispanic households for 7.2%, White households for 64.6%, and Unclassified households for 13.9%. White and API households tend to live in higher-income neighborhoods, whereas Black and Hispanic households live in lower-income areas. However, Appendix Figure A.1a illustrates that the income distributions overlap considerably. White households have the largest average monthly electricity expenditures, API households are the most likely to prefer making a cash purchases, and Hispanic households are the most likely to prefer a loan or lease. All the groups report having roughly similar roof ages and equipment preferences. However, Black and Hispanic households were less likely to complete the optional survey questions about financing, equipment, and roof age.

Finally, Appendix Table A.4 shows how proximity to installers varies by demographic characteristics. Higher-income households are closer to installers—21% more installers are located within 10 miles of the highest-income households relative to the lowest-income households. In addition, Black households generally have fewer nearby installers compared to the other racial groups, potentially limiting their access to solar PV installation services. The top row of Appendix Table A.4 Panel B shows that Black households live in markets with fewer total installers, and also have fewer installers located within 5 or 10 miles of their home.

3 Descriptive Analysis

In this section, we examine how market outcomes vary by race and income, controlling for time and geographic market. We evaluate differences in the number of bids, bid prices, and adoption rates across these groups, using two adoption measures: the *close rate* (adoption via the EnergySage platform) and the *adoption rate* (including off-platform adoption reported in exit surveys). We estimate these relationships with the following regression equation:

[14] Our analysis focuses on the 15 core-based statistical area (CBSAs) with the most EnergySage activity, which represents major metropolitan areas with relatively high earnings compared to the rest of the country (see Appendix Table A.1 for a listing of the CBSAs in this study). Additionally, this is a sample of homeowners who have selected to shop for an expensive investment in durable capital.

$$Y_i = \alpha + \sum_{q=1}^{q=4} \beta_d \mathbb{1}[i \in \text{Income Quintile = q}] + \sum_{r \in R} \theta_r \mathbb{1}[i \in \text{Race = r}] + \text{Size}_i + \gamma_y + \varepsilon_i. \quad (1)$$

where Y_{icy} represents our dependent variable for buyer i. The β coefficients represent the change in the dependent variable by income group, the θ coefficients represent the change in the dependent variable associated with changes in the households' race/ethnicity, Size_i is a control for household i's system size as measured by the mean system capacity bid submitted to the household.[15] Lastly γ_y coefficient represents a fixed effect for the year of the solar auction. We also estimate a separate specification of the model with CBSA-by-year (core-based statistical area) fixed effects. These two specifications allow us to investigate disparities both across and within geographic markets. In these regressions, the omitted category is high-income (*i.e.* fifth income quintile) White households.

Our measure of price is the the median price of all bids a household receives. We also take logs of the dependent variables for price and number of bids to interpret the β and θ coefficients as an approximate percentage change in the dependent variable relative to the omitted group. For our adoption measures—the close rate and adoption rate—we estimate the linear probability models and divide the estimated β and θ coefficients by the constant parameter, α. The regression constant represents the mean for the omitted group conditional on the CBSA-year fixed effects—so our reported estimates can again be interpreted as a percentage change relative to the omitted group.[16]

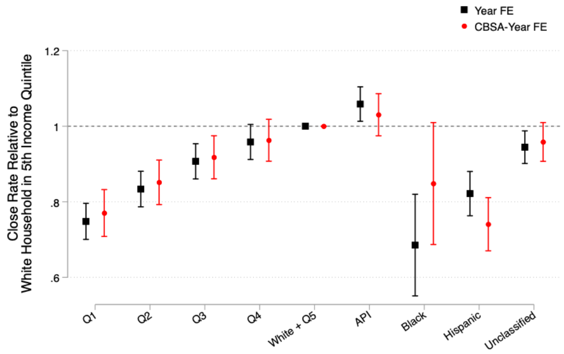
We begin by documenting differences in on-platform close rates and survey-reported adoption rates (including offline adoption) across demographic groups, shown in Figures 2a and 2b. Higher income correlates with higher adoption rates: the close rate for the lowest income quintile is 25% lower than the highest income quintile, and their overall adoption rate is 9% lower.

Disparities also exist across races. In the regressions with year fixed effects, Black households' on-platform close rate is roughly 31% lower, and overall adoption is roughly 10% lower compared to White households. These estimated disparities are attenuated when we include CBSA-Year fixed effects, and while the overall pattern remains similar, the estimated disparities are no longer statistically significant. Across markets, API households exhibit about 6%

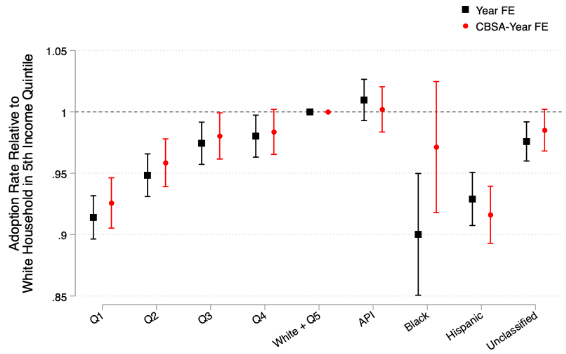
[15] We control for system size to account for installation economies-of-scale that may vary across demographics.

[16] We are restricted to reporting these as relative percentage changes because EnergySage protects the close rate levels as a trade secret. Because our model controls for the size of the system, these normalized estimates should be interpreted as the proportional difference relative to the omitted group if the system size equals 0 kW. We find a small negative coefficient on system size in all our regressions, so all of the estimated disparities are conservative estimates compared to if we considered larger system sizes. In addition, Appendix Figure A.3 shows that we obtain very similar estimates for the bid price regressions if we omit the system size control.

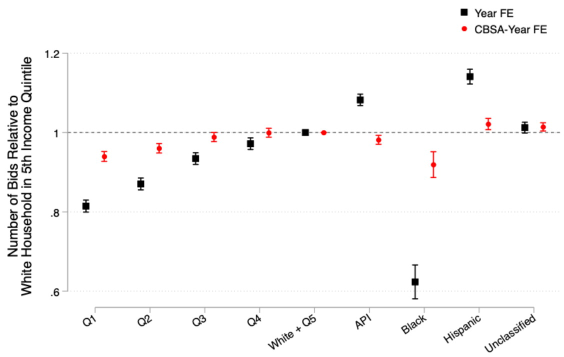
Figure 2: Descriptive Regression Estimates



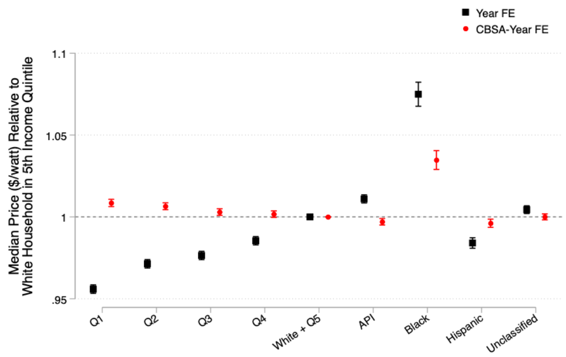
(a) Relative Close Rate by Income & Race



(b) Relative Overall Adoption by Income & Race



(c) Number of Bids per Project by Income & Race



(d) Median Bid Price by Income & Race

Notes: Each panel presents regression coefficients estimated from a regression of the selected outcome variables on income quintile dummies and race dummies. Panel (a) and (b) are linear probability models for the households' decision to purchase solar through the platform or reported overall adoption (on- or off-platform), respectively. The dependent variables in Panel (c) and (d) are the log number of bids and the log of median bid price, respectively. All regressions include year fixed effects and controls for system size (mean capacity bid submitted to the household). Brackets around the point estimates represent 95% confidence intervals. For example, in panel (c), the dependent variable is the logarithm of the number of bids received by a household, and the explanatory variables are income quintile and race. All coefficients are normalized relative to White households in the 5th-income quintile.

higher on-platform close rates and about 1% higher overall adoption rates than White households, although the latter of these results is not significant. Hispanic households also have a nearly 18% lower close rate and a 7% lower overall adoption rate compared to White households.[17]

Next, we consider supply-side decisions by examining the number of bids and bid prices offered across households. In Figure 2c, we plot the coefficients and 95% confidence intervals from our regressions relating the number of bids a household receives to race, ethnicity, and income. In the specification with year fixed effects, we find that API households receive

[17] The estimates in Figure 2b may be biased if there is non-random selection in reporting off-platform adoption across demographics. We, therefore, re-estimate the regression on only the sample of households that completed the exit survey. Appendix Figure A.4 shows that the estimates remain similar among this alternative sample.

roughly 8% more bids relative to White households with similar incomes. However, the sign reverses in the model with CBSA-year fixed effects. This pattern suggests that API households are more likely to live in CBSAs with more competitive solar auctions but that API households get fewer bids than White households that live in the same market. Black households receive the fewest bids of any of the racial groups, obtaining roughly 38% fewer bids than White households with similar incomes. This bid disparity drops to 8% when we control for CBSA-year, which indicates that Black households obtain fewer bids than White households in the same market but also that Black households live in markets with less competitive solar auctions. We also find that Hispanic and unclassified households receive more bids than White households with similar incomes, but this gap is small in magnitude when controlling for CBSA-year.

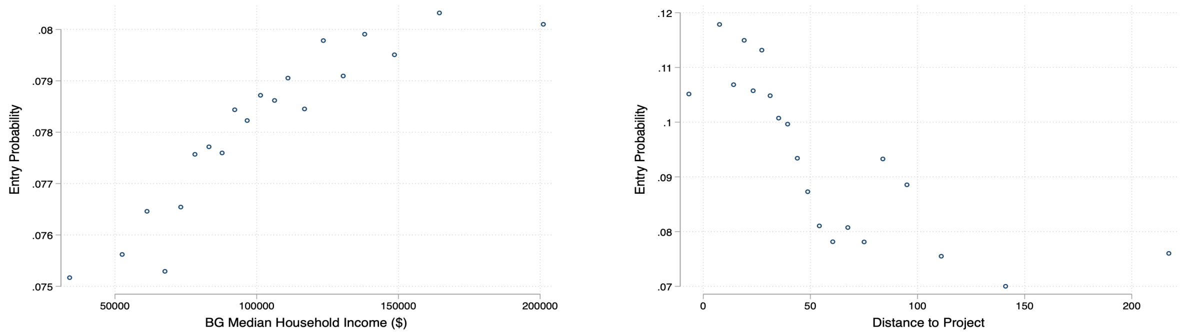
In both models, the number of bids is monotonically increasing across income quintiles. Across markets, the lowest income quintile receives about 19% fewer bids, the 2nd quintile receives about 13% fewer bids, the 3rd quintile receives about 7% fewer bids and the 4th percentile receives 3% fewer bids than the highest income group.[18] This disparity in bid quantity remains but is much smaller in magnitude when we include CBSA-year fixed effects in the regression. This pattern suggests that higher-income CBSAs are more likely to have more solar installers, but also that installers bid more often to high-income households within a particular CBSA market.

To explore why low-income households obtain fewer bids, Figure 3a and 3b illustrate the relationship between the empirical probability that each *potential installer* bids on a project as a function of the household's block group income and installers' distance to the project. The binned scatter plot in Figure 3b demonstrates that installers are less likely to bid on projects that are further away. In particular, installers located within 10 miles of a project submit a bid 10-12% of the time, whereas the probability of bidding falls to below 7.5% when a project is over 100 miles away. Recall, that low-income households live farther away from potential installers (see Appendix Table A.4). Therefore, these patterns suggest that installers' locations are a potentially important factor explaining the lower number of bids received by low-income households.

For bid prices (Figure 2d), we see that when comparing households across CBSAs (year FE model), lower-income households tend to receive bids with a lower average price-per-watt. Specifically, the households in the lowest quintile receive bids over 4% lower than those in the highest income quintile. However, this relationship reverses direction when we compare households in the same market using the model with CBSA-year fixed effects. We see lower-income households obtain higher prices than their higher-income neighbors in the

[18]Given the distribution of race and income shown in A.1a, some of these impacts are likely compounding for Black households. For example, Black households in the 1st quintile of income receive over 50% fewer bids than high-income White households, on average.

Figure 3: Household Characteristics and Installer Entry Probability



(a) Household Income

(b) Installer's Distance to Project

Notes: Panels (a) and (b) depict the probability of installer entry by the household's block group median income and installers' distance to the project. Each point is calculated as the mean probability of entry among each quantile bin after controlling for CBSA fixed effects. The sample includes all installer-project pairs in which the installer is located less than 250 miles from the project.

same CBSA. Quantitatively, these coefficients in this second regression are precisely estimated but small in magnitude–the wedge between the lowest income and highest income household is about 1%.

Overall, the price regressions indicate that CBSAs with lower incomes are likely to have lower per-unit prices, possibly because of lower labor costs or lower markups in these markets. On the other hand, prices within a market are slightly decreasing with household incomes. This within-market pattern is consistent with installation economies-of-scale—with larger systems quoted to richer households in the same market. Appendix Figure A.5 shows that system size is relatively constant across incomes when comparing households across all markets. However, system size increases with income when we compare households within a market. All else equal, larger systems tend to have lower per-unit costs because some components, such as inverters and permitting costs, are fixed. High-income households tend to have larger electricity bills and larger roofs. Therefore, they are more likely to install larger PV systems, contributing to the price disparities observed within CBSA markets.[19]

Regarding race and ethnicity, prices obtained by API and Hispanic households are very similar to prices obtained by White households, especially when we compare households within the same market. For Black households, however, we estimate substantial price gaps relative to White households. Across CBSAs, prices for Black households are 7.5% higher compared to White households. This price disparity persists at 3.5% when comparing Black and White

[19]We show further evidence in Appendix Figures A.2a and A.2b which contain binned scatter plots of the relationship between bid prices and household income, and bid prices and the solar PV system's size in kilowatts (controlling for CBSA). The scatter plots confirm that bid prices (per watt) are declining with income, but importantly that price-per-watt is also declining with the size of the PV system.

households in the same CBSA.[20]

Given the disparities between White and Black households, we explore mechanisms that explain the disparities in bid quantity and price across households' races and ethnicities. Our measure of household race/ethnicity outlined in Appendix A.2.2 contains two sources of information: (1) the racial composition of the household's census block group and (2) the racial information contained in the household's surname (*i.e.,* the probability that a name belongs to race/ethnic group). Bertrand and Mullainathan (2004) show that firms may discriminate based on information contained in individuals' names in remote interactions. Therefore, we estimate regressions to test whether the disparities in bidding across racial groups are primarily explained by racial information contained in buyers' name versus the buyer's neighborhood (*i.e.,* the race/ethnic makeup of the buyer's census block group.) In particular, we estimate regressions analogous to Equation 1 except we omit our preferred binary race/ethnicity variables and instead include as regressors: (1) the proportion of each race/ethnicity group within the households' block group and (2) the probability that each buyer's name belongs to a race/ethnicity group.

The results in Appendix Table A.5 show that neighborhoods and not names overwhelmingly explain the racial disparities in bid prices and the number of bids between Black and White households. Specifically, the bid price premium for Black households is more than fully explained by the information in the census block racial composition. In contrast, the coefficients on the racial proportion of the buyer's name are all relatively small in magnitude and indicate that buyers with Black names are associated with slightly *lower* prices after conditioning on the block group racial composition. Similarly, we find that the gap in the bids (both prices and number of bids) between API and White households is primarily explained by Census block group racial composition. However, we do find that API names are associated with higher prices and fewer bids, but the magnitudes of the API name coefficients are relatively small in magnitude.

Altogether, we take these results as evidence that installers are primarily adjusting bidding behavior based on the locations of buyers. Whereas, we do not find strong evidence of installers screening on buyer's names as in Bertrand and Mullainathan (2004). One caveat to these findings is that we do not directly observe each buyer's race. Moreover, our information about buyers' names is imprecise because we only use information from the buyer's *last name* and not the first name. Therefore, it's possible that our estimates may be attenuated due to measurement error in the the race categorization variables.

[20]Appendix Table A.6 tests for heterogeneity in racial bid price differences across installer ratings. To do so, we regress logged bid price on household race/ethnicity indicators and interact the household race indicators with an indicator for whether the installer has a five-star rating. The estimates reveal that low-rated installers quote higher prices to Black households relative to White Households, but the Black-White price gap is relatively larger among five-star installers.

4 Model

Motivated by the data patterns in the previous section, we develop a structural model that incorporates heterogeneous buyer preferences and strategic bidding by sellers, following Dorsey (2024). Buyers in our model make a discrete choice between the installation bids submitted for their project and the outside option. When estimating the buyers' choice rule, we allow for heterogeneity in price sensitivity across the household's income and race. On the supply side, installers place bids to maximize their expected profits, given their expectations about demand and competing supply bids.

The model allows us to investigate further the distributional disparities we documented in the previous section. We use the demand model to evaluate demand elasticities and consumer surplus separately across income and race. In addition, we use the supply model to separate bid prices into a markup and a cost element. We interpret the cost element implied by our model as the sum of explicit costs (*e.g.* labor and materials) and implicit costs (*e.g.* distaste for serving minority buyers, the time cost of traveling across town, etc.). As such, the supply model allows us to understand better why installation prices vary across household demographics.

In the following subsections, we describe the details of the demand and supply model in detail.

4.1 Demand

Let $\mathcal{K}_i \subset \mathcal{N}(\mathbf{z}_i)$ 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. Buyer i's utility from selecting option j is given by:

$$\begin{aligned} u_{ij} = B_{ij}\alpha_i + \mathbf{x}_{ij}'\beta + \mathbf{w}_j'\gamma + \delta_i + \zeta_{ig} + (1-\lambda)\varepsilon_{ij} \\ \alpha_i = \mathbf{z}_i^{(d)\prime}\tilde{\alpha}, \quad \delta_i = \mathbf{z}_i^{(s)\prime}\tilde{\delta} \end{aligned} \tag{2}$$

Here B_{ij} is the bid price for option j, and α_i is the price sensitivity of buyer i. Buyer price sensitivity, α_i, is a function of an m-dimensional vector of household demographic characteristics denoted $\mathbf{z}^{(d)} = \mathbf{z}_{1:m}$, including the households' race and income. Utility is affected by $\mathbf{x}_{ij}$, the non-price characteristics of the bid, such as the panel brand quality. Buyers' utility also depends on fixed attributes of each seller across bids, $\mathbf{w}_j$, such as installer fixed effects. The δ_i term is a demand shifter for buyer i allowing the utility for all of the installation bids to vary depending on a p-dimensional vector of the household's survey responses, $\mathbf{z}^{(s)} = \mathbf{z}_{m+1:m+p}$, such as the household's geographic market, the year the bids were solicited, the household's monthly electricity expenditure, roof age, equipment preferences, and financing preferences. The variables determining the project type, $\mathbf{z}$, include both sets of household-level variables in $\mathbf{z}^{(d)}$ and $\mathbf{z}^{(s)}$.

Choices are also influenced by ε_{ij}, a random term we assume is independent and identically distributed from 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 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. We specify one group to be the outside option, and the other group to contain all of the project bids. 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" becomes perfectly correlated. Finally, the overall level of utility is not identified, so we normalize the utility of the outside option to equal zero plus an error term. Notably, the outside option in our context subsumes choosing an off-platform installer or a decision not to install solar PV.

In modeling the buyer's choice, we assume each buyer chooses the installation option that delivers the highest utility per unit of capacity. Therefore, B_{ij} in utility is the bid price in dollars-per-watt. Current public policies largely dictate each buyer's optimal system capacity. In particular, net-metering rules allow residential solar customers to sell electricity generated by their rooftop system to their utility at the retail electricity rate as long as the household's annual generation does not exceed their annual consumption. Any solar generation that exceeds the household's annual consumption is compensated far below the retail rate. As a result, the system capacity that will deliver the largest net present benefit to the buyer is the capacity that equates the expected annual solar generation with the expected annual electricity use.[21]

4.2 Supply

We model the installers' bidding decisions in a multi-attribute auction (Dorsey, 2024; Krasnokutskaya et al., 2019; Yoganarasimhan, 2015). In the multi-attribute auction setting, sellers choose to enter an auction knowing their own marginal cost and the mean utility parameters of the buyer but face uncertainty about both the number of competing bidders, the identity of the competing bidders, and the buyer's preference shocks. We assume that sellers know the distribution of buyers' preference shocks and have rational expectations over the entry probabilities and price bids of potential competing sellers. More specifically, they know the characteristics of other active sellers, the price distribution of those sellers' bids, and the probability that those sellers will bid in the auction.

We index sellers by j and differentiate them based on a vector $\mathbf{w}_j$. A seller's type could be

[21] In practice, each installer can propose a different system capacity when bidding through the platform. In a later section, Dorsey (2024) shows that the demand estimates are robust to controlling for the proposed system capacity as a non-price bid attribute and that the demand elasticities are relatively similar using a discrete-continuous choice utility formulation.

distinguished by a relatively parsimonious measure such as a star-rating category, a relatively higher dimensional variable such as a unique installer ID (*i.e.,* seller fixed effects), or a combination of variables. Each seller's type is observable to both the buyer and the other potential sellers. If a seller chooses to participate in the auction for project i they then also select a price bid B_{ij}. Each seller is only permitted a single bid for each project. Sellers' bids are characterized by their price in addition to a vector of non-price characteristics $\mathbf{x}_{ij}$, such as panel quality and inverter type. In contrast to the seller's type $\mathbf{w}_j$, $\mathbf{x}_{ij}$ is allowed to vary across projects for a given seller.

We focus on modeling firms' bid pricing problems conditional on choosing to participate in an auction. After deciding to enter an auction, firms learn their exact marginal cost, the non-price characteristics (such as the panel quality of the system), and the capacity of the system. As in Dorsey (2024), we assume that sellers do not strategically choose the non-price price characteristics and system capacity. These assumptions are necessary for the tractability of the model and are supported in the data (Dorsey, 2024).[22] Conditional on this revealed information set, firms choose a bid price to maximize expected profits. More explicitly, firm j solves the following problem when setting a bid price for project i:

$$\max_{B_{ij}} \quad q_{ij}[B_{ij} - c_{ij}] \cdot \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) \tag{3}$$

Where q_{ij} is the system capacity, B_{ij} is firm j's per-unit price bid, and c_{ij} is firm j's marginal cost.[23] $\mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)$ is the equilibrium expected probability of winning the auction conditional on placing a bid price of B_{ij}. The equilibrium expected probability of being selected is also a function of the project type $\mathbf{z}_i$, the seller's type $\mathbf{w}_j$, and the non-price characteristics of the bid $\mathbf{x}_{ij}$. The project type, $\mathbf{z}_i$, is characterized by the geographic market where the project is located, the time period, and the household's characteristics. We categorize sellers into types $\mathbf{w}_j$ using a relatively parsimonious measure using either the seller's ratings and reviews or seller-specific indicators (*i.e.,* seller fixed effects).[24]

[22]Specifically, Dorsey (2024) shows that installers largely bid the same hardware across subsequent products regardless of buyer characteristics and that system capacity is relatively fixed across bids and not affected by the expected number of bidders for a given project.

[23]Here, c_{ij}, is installer j's average cost-per-watt from installing household i's system, relative to not performing the installation.

[24]Our model is static and therefore rules out dynamic incentives for both buyers and sellers. Dorsey (2024) provides a more detailed discussion on why a static demand model is likely to provide a reasonable approximation in this setting because solar installation prices have become relatively stable during the sample period relative to the early 2010s. However, a limitation of this framework is that it rules out the possibility of dynamic pricing incentives that may arise due to peer effects (Bollinger et al., 2022) or learning-by-doing Bollinger and Gillingham (2019).

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

$$(B_{ij} - c_{ij})\frac{\partial \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)}{\partial B_{ij}} + \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) = 0 \tag{4}$$

Given a vector of non-price characteristics, Equation 4 implicitly defines a seller j's optimal bid for project i. Therefore, we can use observed bid prices along with Equation 4 to infer a seller's marginal cost associated with each project bid. We discuss the supply model in more detail in Appendix A.3.

5 Estimation

We estimate the structural parameters in two steps. First, we solve for the demand parameters via maximum likelihood. Second, we use the estimated demand parameters to simulate firms' first-order conditions for each bid in our data and recover bid-specific markups. We discuss the details of each step in the following subsections.

Demand Estimation

From Equation 2 we observe four sets of variables that affect a buyer's utility: (1) B, the bid price; (2) $\mathbf{z}^{(d)}$, household demographics that determine price-sensitivity; (3) $\mathbf{z}^{(s)}$, households' survey responses that shift the buyer's preference for all of the installation bids (*i.e.,* shifts the likelihood of picking the outside option); and (4) $\mathbf{x}$, variables that characterize the non-price attributes of each participating installer's bid.

We measure the bid price that enters a buyers' utility for option j in dollars-per-watt and scale this value to 70% of each installer's gross bid price to account for the 30% Investment Tax Credit (ITC). We refer to this after-incentive price as the "net price".[25] We allow price sensitivity to vary across household income, race, and ethnicity. In particular, we separate the sample into quintiles based on median income and race/ethnicity. The income quintiles are determined based on the median household income of the household's census block group. We also identify the household head's race/ethnicity as being either (1) Asian or Pacific Islander, (2) Black, (3) Hispanic, (4) White, or Unclassified. We refer to the fourth group simply as "White" households. Therefore, we have nine total variables (including a constant) that shift the price coefficient $\mathbf{z}^{(d)}$, with the constant term representing the price coefficient for White households belonging to the fifth income quintile.

The vector $\mathbf{z}^{(s)}$, which shifts utility for all of the inside options, includes fixed effects for each

[25] This specification implicitly assumes that buyers value a one-dollar reduction in the gross price the same as a one-dollar increase in the tax credit. This assumption is consistent with the existing literature (*e.g.,* Langer and Lemoine, 2022) and is reasonable in this context because, in most cases, the tax credit is paid out to the buyer within one year of the system purchase upon filing an annual tax return.

CBSA and each year of the sample. We also include a set of variables that indicate households' responses to a set of survey questions when they create an account through the platform. All households are required to report their electricity expenditure and an indication of whether they already have quotes from off-platform installers. Therefore, we include the natural log of the household's reported monthly electricity expenditure and a dummy variable for whether the household obtained off-platform quotes in $\mathbf{z}^{(s)}$. Households can optionally report information about the age of their roof, their solar panel equipment preferences, and their financing preferences. Accordingly, we include three sets of indicator variables in $\mathbf{z}^{(s)}$. First, we include an indicator for households with a roof less than 20 years old (or plans to replace the current roof), an indicator for a roof over 20 years old, with an omitted category representing a missing survey response about roof age. Second, we add an indicator for preference over high tech/high production/attractive panels, and an indicator for high-value panels (*i.e.* economical), with the omitted category counting those with a missing response about technology preferences. Third, we have an indicator for a preference for a cash purchase and an indicator for a preference for a loan or lease. Again, the omitted group represents those households who did not report a response about financing preferences.

Finally, the utility for each option is also a function of several non-price characteristics. We include one set of fixed effects for the quality of the solar panels as rated by EnergySage, which include: "Excellent", "Very Good", "Good", "Fair/Poor", and "Missing Rating". We also include fixed effects for the installer's star rating category and a set of fixed effects that measure the installer's installation experience. Moreover, we allow for additional heterogeneity in seller quality by including "permanent" installer fixed effects for each installer that placed over 1000 total bids through the platform during the sample period. These permanent sellers account for over 80% of the bids in our data.

Inferring Markups and Marginal Costs

Next, we use these demand estimates to recover a markup for each bid. We solve for this markup by inputting our final demand estimates into each firm's first-order condition for an optimal bid defined by Equation 4. The first order condition does not have a closed form since it contains two expectations $\frac{\partial \mathcal{P}_{ij}(B_{ij},\mathbf{x}_{ij},\mathbf{w}_j|\mathbf{z}_i)}{\partial B_{ij}}$ and $\mathcal{P}_{ij}(B_{ij}, \mathbf{x}_{ij}, \mathbf{w}_j \mid \mathbf{z}_i)$. Therefore, we solve Equation 4 by integrating 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. We recover marginal costs and markup for each bid using the procedure detailed in Appendix A.4.

6 Results

We show a selected subset of the parameter estimates pertaining to buyers' utility in Table 1. The left-hand panel shows estimates of the price sensitivity coefficients that we allow to vary by the income and race/ethnicity of the household head. The constant term within the set of price coefficients represents the price sensitivity for White households in the highest income quintile. The highest income quintile corresponds to households living in census block groups with a median income above \$137,000. Thus, the interaction terms are interpreted as shifts in price sensitivity relative to high-income White households. We find the price coefficients associated with the third and fourth income quintiles—corresponding to incomes between \$86,000 to \$137,000 are small and not statistically significant. The coefficient for the second income quintile is larger in magnitude but is also not statistically significant. On the other hand, the interaction for the first income quintile—including incomes below \$64,000—is both negative and statistically significant. The demand estimates also indicate notable heterogeneity in price sensitivity across race and ethnicity. While White households and API households exhibit similar price sensitivity, our results suggest Black and Hispanic households are relatively more price sensitive, although the interaction associated with Black households is not quite statistically significant (t-stat of 1.1).

Table 1: Demand Estimates

Nesting Parameter		**Household Survey Responses** (× Inside Option)	
λ	0.391 (0.025)	Log(Electricity Bill)	-0.553 (0.036)
Price Coefficients		Has Off-Platform Quotes {0,1}	0.267 (0.036)
Constant	-0.740 (0.065)	Roof Age: $<=$ 20 years $\{0, 1\}$	0.201 (0.128)
× Income - Quintile 1	-0.113 (0.028)	Roof Age: $>$ 20 years $\{0, 1\}$	0.048 (0.133)
× Income - Quintile 2	-0.039 (0.025)	Purchase Preference: Loan/Lease $\{0, 1\}$	0.771 (0.132)
× Income - Quintile 3	-0.0003 (0.023)	Purchase Preference: Cash Purchase $\{0, 1\}$	0.983 (0.130)
× Income - Quintile 4	0.005 (0.022)	Equipment Preference: Premium Technology $\{0, 1\}$	0.153 (0.049)
× Black Owner	-0.121 (0.113)	Equipment Preference: Value $\{0, 1\}$	0.193 (0.048)
× Hispanic Owner	-0.213 (0.038)		
× Asian/PI Owner	-0.005 (0.021)		

Fixed Effects		**Log Likelihood**
CBSA FEs	Yes	-21155.34
Year FEs	Yes	
Panel Rating FE	Yes	
Permanent Installer FE	Yes	
Transient Installer Star Rating FE	Yes	
Transient Installer # of Reviews FE	Yes	

Notes: The utility specifications include CBSA, year, panel rating, and permanent seller fixed effects. Permanent sellers are those that submitted over 1000 total bids. For transient sellers, we include a set of fixed effects for the installer's star rating and the installer's number of reviews. The right side of the table shows the coefficients associated with the household survey responses, which we allow to shift the utility of all of the installation bids (*e.g.* the inside options). For the survey responses that include dummy variables, the omitted group represents buyers who did not answer the survey question. For example, the "Roof Age: $<=$ 20 years" variable is relative to buyers who did not report the age of their roof. Standard errors are in parenthesis.

Table 2: Mean Price Elasticities Across Demographics

Income Quintile	**Mean Own-Price Elasticity**	**Avg Number of Bids**
1	-2.63	3.99
2	-2.44	4.19
3	-2.31	4.41
4	-2.3	4.52
5	-2.35	4.59

Race/Ethnicity	**Mean Own-Price Elasticity**	**Avg Number of Bids**
Asian, Pacific Islander	-2.4	4.69
Black	-2.79	2.88
Hispanic	-3.03	4.62
White	-2.33	4.28

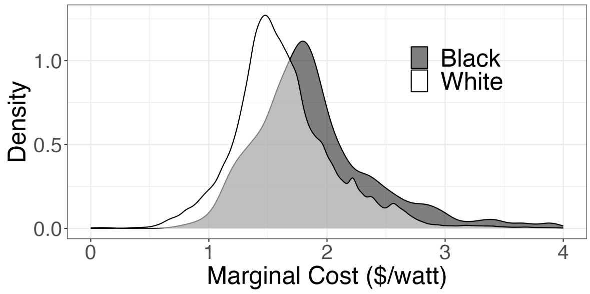
Notes: The mean own-price elasticity is calculated based on the realized choice sets and does not account for ex-ante uncertainty in seller participation.

The right-hand side of Table 1 shows how the household-specific survey responses shift the utility that buyers obtain from each bid relative to the outside option. Notably, buyers with higher electricity bills obtain lower utility. Recall that utility is measured per unit of capacity and households with higher electricity expenditures will tend to be quoted for larger systems. If we transform utility to account for differences in system capacity, households with higher electricity usage obtain higher *total* utility despite obtaining lower utility per unit of capacity. We also find that buyers who report their roof age as being less than 20 years old receive higher utility compared to buyers who did not respond to that survey question (the omitted group). We do not find a statistically significant difference between buyers who report having a roof older than 20 years old relative to non-respondents. Perhaps surprisingly, we discover similar coefficients for buyers who prefer to purchase a solar installation with cash as buyers who prefer to finance their solar installation with a loan or lease. We also find similar willingness-to-pay among buyers who report a preference for value-based technology and buyers with a preference for high-performance or aesthetically pleasing solar panels. However, we find that answering the survey question about financing preferences and the survey question about equipment preferences correspond to a higher willingness to pay relative to the group of non-respondents.

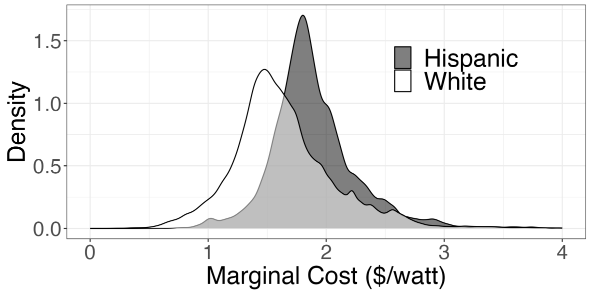
We also report the mean price elasticities separately for each income and racial group in Table 2. We show the lowest income households are more price elastic than the highest income households—with a mean own-price elasticity of 2.63 compared to 2.35 for the highest income group. Black households and Hispanic households are approximately 20% and 30% more elastic than White households, respectively. These results suggest that the gaps in close rates and reported solar adoption across demographic groups are partly explained by differences in price sensitivity. However, Table 2 reiterates that these demographic groups receive substantially different bids from installers. For example, Black households obtain almost 40% fewer bids than White households. Therefore, households' choice sets may also partly explain the disparity in purchase rates.

Before moving on, it is important to note that our empirical specification is restrictive in it only allows for heterogeneity in the price coefficients across demographics. Households may also vary in their preferences over observed attributes like solar panel quality. In Appendix Table A.10, we estimate a models that allow for more flexible heterogeneity in preferences between high- and low-income households, allowing for interactions between income and panel brand quality and the household's survey responses. We find that the most flexible model (Column 2) lacked statistical power—only 2 out of the 12 interactions in the model were statistically significant. Given the lack of statistical power, we proceed using the more restrictive specification with only price heterogeneity. As such, a limitation of our main specification is that we are unable to capture non-price heterogeneity that could vary across demographics and affect our consumer welfare estimates.

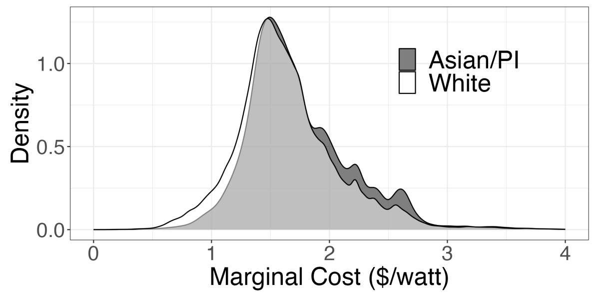
Figure 4: Marginal Cost Distributions by Race and Income



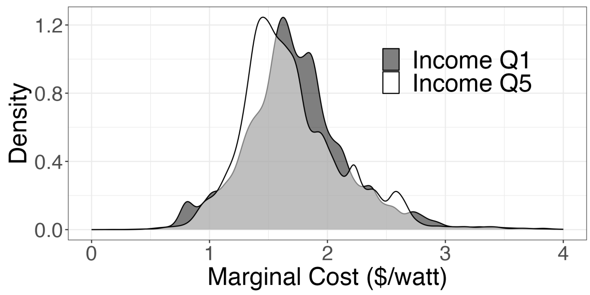
(a) Black MC and White MC



(b) Hispanic MC and White MC



(c) Asian/PI MC and White MC



(d) High Income MC and Low Income MC

Notes: Each panel compares the distribution of marginal costs in dollars per watt across all bids submitted to two designated groups of households. For example, the dark blue distribution in Panel 4a shows the distribution of estimated marginal costs across all bids submitted to Black households, and the red distribution in Panel 4a shows the distribution of marginal costs associated with all bids offered to White households.

6.1 Marginal Costs and Markups

We now turn to the supply side to investigate differences in markups and marginal costs across households. The supply-side model allows us to disentangle potential mechanisms that could explain price differences in solar installation bids across households. The results from the previous section indicate that Black and Hispanic households tend to be more price-elastic than White households. Therefore, these differences in preferences across households should translate into differences in the sellers' optimal price markup across households. On the other hand, heterogeneity in underlying installation costs may also explain the price variation across households.

Figure 4 compares the distribution of installation marginal costs across demographic groups. Figure (4a)-(4c) contrasts the estimated marginal cost distributions for bids made to White households, with the distribution for Black, Hispanic, and API households, respectively. These figures show the marginal cost distribution for Black households is shifted to the right of the distribution for White households indicating that the higher bid prices obtained by Black households cannot fully be explained by markup incentives. A similar pattern emerges in Figure 4b, which illustrates that marginal costs for Hispanic households also tend to be higher relative to White households. In panel 4c, we see that the estimated marginal cost distribution for API households looks very similar to the distribution for White households. Finally, Figure

4d shows that marginal costs for low-income households are slightly higher than marginal costs for high-income households.

While these distributions suggest substantial differences in the underlying cost structure across demographic groups, Figure 4 shows only the unconditional distribution of marginal costs without controlling for any other variables that may help to explain the heterogeneity. In Appendix Table A.7, we run a series of regressions to better understand how installation markups and marginal costs vary across household demographics. In particular, we show in Column (3) that log markups vary across income and race/ethnicity after controlling for market fixed effects, installer fixed effects, hardware quality, and each household's survey responses, including their monthly electricity expenditure. These results imply that mean optimal markups for bids made to Black households is approximately 13% lower relative to observably similar White households, the mean optimal markup for Hispanic households is 25% lower than for White households, and the optimal markup for API households is nearly the same as for White households. With respect to income, the mean optimal markup for bids made to households in the lowest income quintile is roughly 14% below the optimal markup for an observably similar household in the highest income quintile. In the sixth column of Table A.7, we show results for an analogous regression using log marginal cost as the dependent variable. These results show implied marginal costs are 15% higher for Black than White households, 16% higher for Hispanic than White households, and roughly the same for API and White households. Similarly, we find mean marginal cost for the lowest income group is 12% above the marginal cost for the highest income group.

These results document considerable heterogeneity in marginal costs across households. In particular, we find that solar installation costs tend to be higher for low-income, Black, and Hispanic households. This indicates that underlying supply conditions may contribute to the adoption disparities we observed in Section 3. There are several reasons why marginal costs could vary systematically across demographic groups, such as heterogeneity in labor, materials, and transportation costs. For example, Table A.4 showed that low-income and minority households tend to live further away from installers which could plausibly explain some of the estimated disparities in installation costs. In addition, households with lower income often buy smaller PV arrays which results in higher per-watt prices due to fixed costs like permitting and inspection fees. Another possibility is that installers have a taste-based preference against certain types of households (or certain neighborhoods) and adjust their bid prices to reflect these tastes. For example, some sellers may prefer to serve certain households or neighborhoods- either because of intolerant views, crime rates, or other variables correlated with household demographic characteristics. If these preferences are factors in firms' bids, they would be incorporated implicitly into our marginal cost estimates.

6.2 Consumer Surplus

We next use our estimates to evaluate disparities in expected consumer surplus across household demographics. A household's consumer surplus depends on both their preferences and their choice set. Specifically, the expected consumer surplus for a household in demographic group A is given as follows:

$$\mathbb{E}[CS|i \in A] = \mathbb{E}\Big[\frac{\overline{q_i}}{\alpha_i} \log\big(\sum_{j \in \mathcal{K}_i} \exp(B_{ij}\alpha_i + \mathbf{x}'_{ij}\beta + \mathbf{w}'_j\gamma + \delta_i)\big)\Big]. \tag{5}$$

where A indicates the set of households in demographic group A. This formulation makes explicit that the utility parameters, such as the price coefficient, α_i, and the responses associated with the household survey responses, δ_i, can vary across individuals within a demographic group. For instance, households within a particular racial group can have heterogeneous price coefficients depending on their income. Here, $\overline{q_i}$ is the mean system size bid submitted to household i—this term scales utility to adjust for differences in system size across households. We estimate the expectation using the empirical mean consumer surplus of all households within a demographic group.

The top panel of Table 3 reveals stark differences in expected consumer surplus across household income. The second column shows the expected consumer surplus per unit of solar capacity, based on the mean system size quoted to each household. The fourth column shows the total expected consumer surplus for each group, accounting for differences in system capacity across income groups. We see that the expected total consumer surplus for the lowest income quintile ($577) is less than half of the surplus obtained by the highest income quintile ($1,229). Similarly, the bottom panel indicates that total consumer surplus is highest for API and White households ($1,097 and $988). In contrast, Black households obtain 72% less surplus ($281), and Hispanic households obtain 56% less surplus ($431) than White households.

Appendix Table A.8 indicates that the disparities in consumer surplus across income and race largely persist after controlling for the household's geographic market, electricity expenditure, and other observables. In this table, we report results from regressions with the log of consumer surplus as the dependent variable and include controls for all variables that enter the demand model that varies across households. Even after adding this set of controls, the coefficient estimates indicate consumer surplus for the bottom income quintile is approximately 59% lower than the top income quintile. Moreover, the regression shows that Black households' consumer surplus is 81% lower than comparable White households, and Hispanic households' surplus is 73% lower, all else equal. On the other hand, API households get roughly equal surplus as White households.

Table 3: Mean Expected Consumer Surplus Across Income and Race

Panel A: Income Quintile

Income Quintile	Bids	CS/watt ($/watt)	Mean System Size (kW)	Total CS ($)
1	3.99	0.07	9.03	576.93
	(1.93)	(0.07)	(4.22)	(526.29)
2	4.19	0.1	8.75	804.67
	(1.94)	(0.09)	(4.16)	(687.03)
3	4.41	0.13	8.62	1022.32
	(1.91)	(0.11)	(4.07)	(821.53)
4	4.52	0.15	8.61	1135.78
	(1.86)	(0.12)	(4.11)	(883.12)
5	4.59	0.16	8.86	1228.94
	(1.77)	(0.11)	(4.26)	(904.79)

Panel B: Race/Ethnicity

Race/Ethnicity	Bids	CS/watt ($/watt)	Mean System Size (kW)	Total CS ($)
Asian, Pacific Islander	4.69	0.17	7.1	1097.15
	(1.78)	(0.12)	(3.5)	(836.74)
Black	2.88	0.03	9.18	281.38
	(1.64)	(0.04)	(4.07)	(322.31)
Hispanic	4.62	0.06	8.38	430.54
	(1.86)	(0.05)	(3.89)	(376.1)
White	4.28	0.12	9.08	988.03
	(1.9)	(0.1)	(4.23)	(820.19)

Notes: The second column reports the mean number of bids obtained across all households within the group. The third column reports the mean expected consumer surplus ($/watt) across all households within the group. The fourth column shows the mean system size quoted to all households in the group; first, we calculate the mean capacity bid for each household, and then we average that figure across households. The total consumer surplus for each household is calculated at the expected consumer surplus per unit multiplied by the mean capacity bid for that household. The standard deviations of each variable are listed in parentheses.

6.3 Decomposing Disparities in Consumer Surplus

The results in the previous subsection highlight the large disparities in consumer surplus across household demographics. In this section, we quantify the fraction of these disparities that are explained by differences in household preferences (demand) versus differences in choice sets (supply). The evidence in Section 3 and 6.1 shows that low-income and some minority households obtain fewer bids suggesting that supply-side factors may partly explain the income and racial disparities in consumer surplus. However, the demand estimates in Table 1 showed that low-income and minority households also tend to be relatively more price sensitive, which implies that the disparities in consumer surplus might be largely explained by differences in willingness-to-pay for solar PV.

We develop a decomposition method similar in spirit to Oaxaca (1973) and Blinder (1973) to better understand how supply and demand factors separately affect disparities in consumer surplus. First, consider the difference in consumer surplus between two groups—demographic group A and demographic group B:

$$\Delta CS = \mathbb{E}[CS|i \in A] - \mathbb{E}[CS|i \in B]. \tag{6}$$

Analogous to Figure 1, we seek to empirically decompose ΔCS into three components: (1) the *demand component*—the share of the CS gap explained by differences in preferences across the two groups, (2) the *supply component*—the portion of the CS gap explained by differences in choice sets across the two groups, and (3) the *interaction component*.

To derive the decomposition, we first introduce the notation $CS(\theta_i, \mathcal{K}_{i'})$, which concisely represents the consumer surplus that a household with preferences $\theta_i \equiv \{\alpha_i, \delta_i\}$ would obtain if that household were faced with the choice set offered to household i'. Formally, we define $CS(\theta_i, \mathcal{K}_{i'})$ as:

$$CS(\theta_i, \mathcal{K}_{i'}) = \frac{\overline{q_i}}{\alpha_i} \log\big(\sum_{j \in \mathcal{K}_{i'}} \exp(B_{i'j}\alpha_i + \mathbf{x}'_{i'j}\beta + \mathbf{w}'_j\gamma + \delta_i)\big). \tag{7}$$

The expression makes clear that $CS(\theta_i, \mathcal{K}_{i'})$ is evaluated by using the utility parameters of household i ($\theta_i \equiv \{\alpha_i, \delta_i\}$), the system size of household i ($\overline{q_i}$); and the bid prices ($B_{i'j}$), installer attributes ($\mathbf{w}'_j$), and system attributes ($\mathbf{x}'_{i'j}$) offered to household i'. Given this notation, we can express the *counterfactual* expected consumer surplus that a demographic group A would obtain if their choice sets were drawn from the distribution of choice sets offered to households in a demographic group B as follows:

$$\mathbb{E}[CS|i \in A, i' \in B] = \int\int CS(\theta_i, \mathcal{K}_{i'}) dF(\theta_i|i \in A) dG(\mathcal{K}_{i'}|i' \in B). \tag{8}$$

Where F represents the joint distribution of preferences, θ_i, conditional on the household belonging to group A, and G is the distribution of choices sets, $\mathcal{K}_{i'}$, offered to households belonging to group B. We leverage the above expression to decompose the expected gap in consumer surplus between group A and group B using the following formula:

$$
\begin{aligned}
\Delta CS = & \underbrace{\mathbb{E}[CS|i \in A, i' \in B] - \mathbb{E}[CS|i \in B]}_{\text{Demand Component}} \\
& + \underbrace{\mathbb{E}[CS|i \in B, i' \in A] - \mathbb{E}[CS|i \in B]}_{\text{Supply Component}} \\
& + \underbrace{\mathbb{E}[CS|i \in A] - \mathbb{E}[CS|i \in B] - \text{Demand Component} - \text{Supply Component}}_{\text{Interaction Component}}. \qquad (9)
\end{aligned}
$$

The first line of Equation 9 represents the demand component—the portion of the consumer surplus gap explained by differences in preferences between demographic group A and demographic group B. More specifically, the demand component captures the difference between the counterfactual expected consumer surplus if group A households faced group B's distribution of choice sets relative to Group B's observed consumer surplus. Analogously, the supply component of the consumer surplus gap (second line of Equation 9) represents group B's expected gain in consumer surplus from receiving group A's choice sets, compared to their own choice sets.

The final line introduces the interaction component, which accounts for the residual gap not explained by differences in the groups' preferences or choices set on their own. This term includes any additional gains that arise from changing both preferences and choice sets of the two groups simultaneously. Intuitively, an improved choice set should benefit all consumer groups but may benefit consumer groups with higher and more inelastic demand by relatively more (see Figure 1d). Thus, the interaction component measures the additional benefit that Group A obtains from an improved choice set relative to Group B.[26]

Intuitively, if we swap the base group for the decomposition from B to A, the new demand component should be equal in absolute value to the sum of the old demand component and the old interaction component. As a validity check, we show the alternative results of the decomposition when we swap the base group from the low-income group to the high-income group in Appendix Table A.15.[27]

We use Equation 9 to decompose the disparity in consumer surplus between high- and low-income households and then to decompose the gap between White and minority households.

[26] Consider a case where Group B obtains extremely low value from adopting solar (e.g., $u_{ij} = -1000 - 2B_{ij} + \varepsilon_{ij}$) whereas Group A obtains higher value (e.g., $u_{ij} = 10 - B_{ij} + \varepsilon_{ij}$). In this case, the supply component will equal approximately zero because Group B will obtain approximately zero consumer surplus gain from any modest improvement in the choice set. The interaction component, in this case, would, therefore, capture the benefit that Group A obtains from the improved choice set.

[27] We see that the demand component in Panel B after swapping the base group is (-\$618.95) is approximately equal in absolute value to the sum of the demand component in Panel A and interaction component in Panel A (\$607.73=\$663.38-\$55.65). The terms are not exactly equal due to simulation error in approximating the integrals in Equation 9.

We evaluate the counterfactual expected consumer surplus terms in the decomposition (e.g., $\mathbb{E}[CS|i \in A, i' \in B]$) by simulating from the relevant empirical distributions $F(\theta_i|i \in A)$ and $G(\mathcal{K}_{i'}|i' \in B)$.

The decomposition results are mixed. The top panel of Table 4 decomposes ΔCS between households in the highest income quintile and households in the lowest income quintile. The results show the $652 gap in consumer surplus between these income groups is fully explained by the demand component. That is, high-income households would achieve an approximately equal level of consumer surplus if they faced similar choice sets to low-income households. Similarly, Panel C shows the $557 consumer surplus gap between White and Hispanic households, is more than fully explained by the demand component. For this decomposition, the supply component is negative which indicates that providing Hispanic households with similar choice sets to White households would reduce their expected consumer surplus.

On the other hand, Panel B, which decomposes the $706 consumer surplus gap between White and Black households, paints a different picture. Here, we see that only 34% of the consumer surplus gap between these groups is explained by differences in preferences. In contrast, 37% is explained by the supply component, which means that Black households expected consumer surplus would increase by $263 if they had access to similar choice sets as White households. The estimates for this decomposition are less precise due to the smaller sample of Black households, but the supply component is statistically distinguishable from zero at the 1% level. Recall that Black households obtain 1.4 fewer bids on average than White households–a 33% reduction–which indicates that Black households receiving more restricted choice sets than White households contributes to a substantial portion of the CS gap. The interaction component explains 29% of ΔCS, which means that White households would obtain a relatively larger gain in consumer surplus from a comparable improvement in their choice sets relative to Black households. Lastly, Panel D includes the decomposition of the consumer surplus gap between White and API households for completeness. However, the overall consumer surplus gap between these groups is not statistically significant and nor are any of the components in the decomposition.

Table 4: Decomposition of Consumer Surplus Gap by Income and Race

Panel A: High-Low Income Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	652.01	(117.28)
Demand Component ($)	663.38	(111.72)
Share Explained by Demand Component	1.02	(0.05)
Supply Component ($)	44.29	(19.83)
Share Explained by Supply Component	0.07	(0.03)
Interaction Component ($)	-55.65	(24.35)
Share Explained by Interaction Component	-0.09	(0.04)

Panel B: White-Black Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	706.65	(153.3)
Demand Component ($)	237.77	(115.17)
Share Explained by Demand Component	0.34	(0.15)
Supply Component ($)	263.79	(103.32)
Share Explained by Supply Component	0.37	(0.3)
Interaction Component ($)	205.09	(105.98)
Share Explained by Interaction Component	0.29	(0.15)

Panel C: White-Hispanic Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	557.49	(97.18)
Demand Component ($)	713.32	(112.4)
Share Explained by Demand Component	1.28	(0.05)
Supply Component ($)	-23.64	(6.26)
Share Explained by Supply Component	-0.04	(0.01)
Interaction Component ($)	-132.19	(23.32)
Share Explained by Interaction Component	-0.24	(0.04)

Panel D: White-Asian/PI Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	-109.12	(68.17)
Demand Component ($)	-87.39	(70.76)
Share Explained by Demand Component	0.8	(0.62)
Supply Component ($)	74.55	(18.65)
Share Explained by Supply Component	-0.68	(2.08)
Interaction Component ($)	-96.28	(22.58)
Share Explained by Interaction Component	0.88	(2.55)

Notes: The second row of each panel shows the mean consumer surplus gap between the two groups of households (*e.g.* White mean consumer surplus minus Black mean consumer surplus). The middle section of each panel decomposes the gap in consumer surplus into a portion explained by price sensitivity and a portion that is unexplained by price sensitivity. The top row describes the reference group used to measure the decomposition. The bottom section of each panel shows the gap in the mean number of bids received across the two groups. Bootstrapped standard errors are in parentheses.

We provide additional intuition for our decomposition results in Appendix Tables A.11-A.14. In these tables, we further decompose the gap in consumer surplus into smaller sub-components to understand better the key factors driving our main supply-demand decomposition results. Recall that when we calculate the demand component, we evaluate the counterfactual consumer surplus with group B's choice sets but with group A's preferences and system sizes. Therefore, we can further break down this decomposition by unilaterally varying specific elements. For example, we can determine the portion of ΔCS explained by differences in system size across the two groups by evaluating the counterfactual consumer surplus with group B's choice sets and preferences but with group A's system sizes, $\bar{q}_i$. Similarly, we can estimate the share of ΔCS explained by differences in price sensitivity by calculating the counterfactual consumer surplus with group B's choice sets, group B's system sizes ($\bar{q}_i$), and group B's utility intercepts (δ_i), but with group A's price coefficients (α_i). In this former case, we calculate the counterfactual consumer surplus by sampling from group A's marginal distribution of $\bar{q}_i$. In the latter case, we sample from group A's marginal distribution of α_i.

We find that differences in system size across the groups explain only a minor share of the disparities in consumer surplus. Specifically, system size accounts for 10% of ΔCS between high- and low-income households, none of ΔCS between White and Black households, and 16% of ΔCS between White and Hispanic households. In comparison, differences in price coefficients alone can explain 46% of the gap between high- and low-income groups, 23% of the gap between White and Black households, and 81% of the White-Hispanic gap.

The CBSA and year fixed effects, which shift the utility intercept (δ_i) for each bid relative to the outside option, explain a significant portion of the disparities, 36% between high- and low-income households, 9% between White and Black households, and 9% between White and Hispanic households. These estimates underscore how differences in solar preferences across markets with varying demographic characteristics influence consumer surplus disparities. For instance, households in higher-income areas such as Washington D.C. generally experience higher levels of utility from solar bids (higher mean δ_i), while those in lower-income CBSAs like Houston experience lower utility from these bids (lower mean δ_i).

A modest share, 7-19% of the consumer surplus disparities, are explained by within-CBSA differences in household survey responses across the groups, which similarly shift the utility intercept (δ_i). The survey response covariates include indicators for the household's financing preference, roof age, equipment preferences, off-platform search, and monthly electricity usage. For example, high-income households are more likely to prefer cash purchases and tend to have newer roofs, factors estimated to be associated with higher utility from each bid.[28]

[28]Note that each element of these sub-decompositions are performed independently and therefore the shares need not sum to one.

6.4 Eliminating Consumer Surplus Disparities through Prices

The previous sections document substantial gaps in consumer surplus across socioeconomic and demographic groups in the residential solar PV market. Consequently, policymakers have expressed interest in crafting policies to reduce distributional inequities in the solar PV market and in related markets. For example, the EPA launched the "Solar for All" initiative in 2023 that will award up to 60 grants to states, territories, Tribal governments, municipalities, and nonprofits to expand the number of low-income and disadvantaged communities to invest in residential solar energy (Environmental Protection Agency, 2023). In addition, the Inflation Reduction Act (IRA) proposed a low- and middle-income (LMI) adder that offers supplemental tax credits of 10 to 20 percent, in addition to the standard Investment Tax Credit (ITC), for small solar and wind projects (Internal Revenue Service, 2023). These projects must either meet environmental justice standards at a community level or satisfy income requirements at a residential scale.[29] Finally, California's Low-Income Weatherization Program offers eligible households subsidies that cover the entire cost of solar installations, solar hot water heaters, heat pump technology, and other energy efficiency retrofits (California Department of Community Services and Development, 2016).[30] A common theme across all of these policies is to offer subsidies or grants that reduce the upfront cost of installing solar for low-income and disadvantaged households.

Motivated by these policies, we evaluate the price adjustment needed to eliminate disparities in consumer surplus across different demographic groups. In particular, we calculate how much each bid price would have to change to equalize expected consumer surplus across two groups if all else was held equal (*i.e.,* each buyer's choice set).

[29]According to the IRA, those who own projects situated in impoverished communities or on Indian Land are eligible for a further 10 percent increase in tax credits. Moreover, project owners who cater to affordable housing residents or who dedicate a portion of their projects to serving low-income customers are entitled to an additional tax credit boost of 20 percent.

[30]The stated policy objective of California's Low-Income Weatherization Program is to reduce greenhouse gases while also lowering low-income households' energy costs (California Department of Community Services and Development, 2016). Notably, there are many other alternate policies that could help achieve the goals. For example, California implemented a cap and trade program as a market-based approach to mitigate greenhouse gases. In addition, Borenstein et al. (2021) discuss electricity rate reforms that could help alleviate energy cost burdens for low-income households.

Table 5: Price Adjustments Required to Close Consumer Surplus Gaps

Group	Uniform Price Adjustment ($/watt)	Mean Price After Adjustment ($/watt)	Mean Consumer Surplus After Adjustment ($)
Base: High Income (Q5)	0	2.13	1228.94
Comparison: Low Income (Q1)	-0.57	1.56	1228.94
Base: White	0	2.07	988.03
Comparison: Black	-0.89	1.18	988.03
Base: White	0	2.07	988.03
Comparison: Hispanic	-0.54	1.53	988.03
Base: White	0	2.07	988.03
Comparison: Asian/PI	0.09	2.16	988.03

Notes: The second column calculates uniform change to all bid prices (net price after the ITC) that would imply that the base group (*e.g.,* high-income households) and the comparison group (*e.g.,* low-income households) obtain equal expected consumer surplus. The third column reports the mean bid for each group after the price adjustment. The last column reports the expected consumer surplus for each group after implementing the uniform bid price adjustment.

The top portion of Table 5 considers the top income quintile of households as the baseline group, and the lowest income quintile as the comparison group. We solve for the (uniform) increment to low-income households' bid prices that would result in the same level of expected consumer surplus as the high-income households.[31] We see that if each bid submitted to a low-income household decreased by $0.57 per watt–approximately 27%—then low-income households would obtain the same expected consumer surplus as the high-income households. The lower portions of the table calculate the changes to prices that would eliminate the racial gaps in expected consumer surplus. We find that a substantial price reduction of $0.89 per watt (43%) would be necessary to eliminate the gap in consumer between White and Black households, and a $0.54 per watt (26%) decrease would be required to eliminate the gap between White and Hispanic households. Finally, prices submitted to API households would need to rise by $0.09 per watt (4%) to equate the expected consumer surplus with White households, since API households obtain slightly higher consumer surplus than White households.

The above results underscore that in the short run, substantial subsidies or grants may be necessary to eliminate disparities in consumer surplus and solar PV adoption that we currently observe. For instance, our results suggest that the IRA LMI adder of 10-20% may not be sufficient to fully eliminate the gap in consumer surplus between the highest and lowest-income households in our sample. Whereas, programs like the California Low-Income Weatherization program—which offers free solar installations to eligible households—may be more than sufficient to close the gap between the highest and lowest-income households in our sample.

An alternate means of eliminating consumer surplus disparities is to increase the prices for demographic groups that currently obtain higher consumer surplus. For instance, policymak-

[31] All prices are reported as the net price after incorporating the 30% ITC.

ers concerned about distributional outcomes may prefer eliminating subsidies paid to high-income households, particularly if program budgets are limited. As an example, the Inflation Reduction Act set eligibility limits for electric vehicle tax credits based on income. In Appendix Table A.9, we swap the comparison groups in Table 5 and calculate the uniform price *increase* that would reduce the mean consumer surplus of high-income households to the level of low-income households.[32] We find that if prices for high-income households increased uniformly from $2.01/watt to $2.67 per watt, this would close the consumer surplus gap between the high- and low-income groups. Interestingly, a price change of this magnitude could be achieved by unilaterally eliminating the 30% ITC for high-income households conditional on a subsidy pass-through rate greater than 80%.

An important caveat of this exercise is that it does not provide insights about longer-run changes in seller behavior that may be caused by targeted subsidies or grants. In the longer run, targeted subsidies and grants may encourage more installer entry and bids submitted to low-income households. These longer-run changes in entry and participation could further help to reduce welfare disparities in the solar PV market and other related markets. In addition, our calculations only consider changing prices of bids made to buyers through the online platform and therefore do not incorporate changes in consumer surplus that buyers may accrue from offline solar installers through targeted subsidy programs.[33] Finally, while our analysis is informative about the magnitude of price adjustment that would equalize consumer surplus across groups, implementing these price changes in practice may also bear considerable social costs. For example, targeted subsidy programs might increase administrative costs or crowd out other public spending.

7 Conclusion

In this paper, we document significant distributional disparities across buyers on a leading online marketplace for residential solar installations. Our findings indicate that low-income, Black, and Hispanic buyers are less likely to install solar conditional on visiting the platform. Consequently, these households derive substantially lower levels of consumer surplus from the market. Our research adds to a growing body of research documenting disparities in the adoption of new energy technologies. Consequently, these households are capturing only a small share of government tax credits and other subsidies commonly offered for these emerging technologies.

We contribute to the literature by leveraging rich data on sellers' bids to further investigate the mechanisms contributing to the adoption gap and corresponding welfare disparities. We

[32]We also carry out the analogous calculations for the consumer surplus disparities across race.

[33]Importantly, in all of the calculations in Table 5 we maintain the normalization that buyers obtain zero utility from choosing the outside option.

show that disparities in consumer welfare across high- and low-income households are fully explained by differences in household preferences. In contrast, differences in choice sets account for a significant share of the welfare gap between Black and White households, as Black households receive fewer bids and higher average bid prices. Notably, installers face higher implied costs of serving Black households, which contributes to the disparities in bids. Furthermore, our analysis reveals that the disparity in bidding between Black and White households is primarily due to neighborhood locations rather than household names.

Recent environmental policy discussions highlight an increased focus on ensuring that social benefits are distributed equitably. In the context of climate policy, this means that the harms from global warming shouldn't be borne disproportionately and that the rewards from green capital investments should be shared more broadly among the population. Our findings suggest that offering targeted subsidies to low-income households—such as the LMI Adder in the Inflation Reduction Act—can help mitigate welfare disparities within our sample of potential solar buyers, although these subsidies come with high fiscal costs. Furthermore, our decomposition results are suggestive that supply-side government policies aimed at reducing the barriers to entry for firms operating in majority Black neighborhoods may complement consumer subsidies, leading to a reduction in the gaps in adoption and consumer well-being. Similar supply-side instruments may be available on online platforms themselves, which may be able to implement policies that adjust the fees or commissions charged to installers for projects located in underserved neighborhoods.

There are a number of important caveats to consider when interpreting our results. First, while our data provides novel insights about buyers' and sellers' behavior in the residential solar PV market, our analysis is inherently limited to households that have selected to use the EnergySage platform to search for quotes. The individuals in our sample, have higher incomes and are more likely to be White relative to the general U.S. population. One reason for these patterns is that home ownership is typically a prerequisite for purchasing a solar PV system, and low-income, Black, and Hispanic households are more likely to rent their homes. By studying individuals who have already selected to shop for solar, our results are likely to understate the extent of welfare disparities that exist across the broader population. Thus, understanding demographic differences in awareness and consideration of clean technology adoption is an important topic for future work. Second, our results demonstrate that low-income and Black households obtain relatively fewer bids. However, we are not able to fully determine the extent to which these patterns reflect *taste-based* discrimination versus differences in true underlying costs of providing installation services to these households. Our research reflects a first step in understanding the mechanisms behind these supply-side disparities, including providing descriptive evidence that sellers are likely to locate closer to high-income and White households. However, apart from location, we know relatively little about the sellers, so research leveraging detailed seller information to more accurately explain the reasons behind any observed deviation in bidding across racial and income groups presents an important opportunity for further research.

Finally, while our findings are informative for policymaking, there are an array of other crucial issues that our work omits. For example, our paper focuses on the residential solar market and does not consider community or utility-scale solar investments as potentially cost-efficient ways to reduce electricity emissions and mitigate energy cost disparities. In addition, our analysis ignores energy cost spillovers from solar adoption to non-solar adopters that may arise due to overlapping policies like net energy metering (Feger et al., 2022; Darghouth et al., 2011; Seybert et al., 2013; Borenstein, 2017). Moreover, our model does not directly incorporate additional market frictions that disadvantaged households may face, such as credit constraints and information failures.

References

Avenancio-León, Carlos F and Troup Howard, "The assessment gap: Racial inequalities in property taxation," *The Quarterly Journal of Economics*, 2022, *137* (3), 1383–1434.

Ayres, Ian and Peter Siegelman, "Race and Gender Discrimination in Bargaining for a New Car," *The American Economic Review*, 1995, *85* (3), 304–321.

_ **, Mahzarin Banaji, and Christine Jolls**, "Race Effects on EBay," *The RAND Journal of Economics*, 2015, *46* (4), 891–917.

Barbose, Galen L and Naïm R Darghouth, "Tracking the sun: Pricing and design trends for distributed photovoltaic systems in the United States-2022 edition," Technical Report, Lawrence Berkeley National Laboratory 2023.

Becker, Gary S., *The Economics of Discrimination* Economic Research Studies, University of Chicago Press, 1971.

Bednar, Dominic J. and Tony G. Reames, "Recognition of and response to energy poverty in the United States," *Nature Energy*, 2020, *5* (6), 432–439. Number: 6 Publisher: Nature Publishing Group.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94* (4), 991–1013.

Blinder, Alan S., "Wage Discrimination: Reduced Form and Structural Estimates," *The Journal of Human Resources*, 1973, *8* (4), 436–455.

Blonz, Josh, Brigitte Roth Tran, and Erin E Troland, "The Canary in the Coal Decline: Appalachian Household Finance and the Transition from Fossil Fuels," Technical Report, National Bureau of Economic Research 2023.

Bollinger, Bryan and Kenneth Gillingham, "Learning-by-doing in solar photovoltaic installations," *Available at SSRN 2342406*, 2019.

_ , _ , A Justin Kirkpatrick, and Steven Sexton, "Visibility and peer influence in durable good adoption," *Marketing Science*, 2022, *41* (3), 453–476.

Borenstein, Severin, "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*, 2017, *4* (S1), S85–S122.

_ and Lucas W. Davis, "The Distributional Effects of US Clean Energy Tax Credits," *Tax Policy and the Economy*, 2016, *30* (1), 191–234.

_ , Meredith Fowlie, and James Sallee, "Designing electricity rates for an equitable energy transition," *Energy Institute Working Paper*, 2021, *314.*

Butters, R. Andrew, Daniel W. Sacks, and Boyoung Seo, "Racial Difference in Retail Prices Paid," *Working Paper*, 2022.

California Department of Community Services and Development, "Low-Income Weatherization Program," 2016. https://www.csd.ca.gov/Pages/Low-Income-Weatherization-Program.aspx.

Christensen, Peter and Christopher Timmins, "Sorting or steering: The effects of housing discrimination on neighborhood choice," *Journal of Political Economy*, 2022, *130* (8), 2110–2163.

_ and _ , "The damages and distortions from discrimination in the rental housing market," *The Quarterly Journal of Economics*, 2023, *138* (4), 2505–2557.

Darghouth, Naïm R, Galen Barbose, and Ryan Wiser, "The impact of rate design and net metering on the bill savings from distributed PV for residential customers in California," *Energy Policy*, 2011, *39* (9), 5243–5253.

Dauwalter, Travis E and Robert I Harris, "Distributional Benefits of Rooftop Solar Capacity," *Journal of the Association of Environmental and Resource Economists*, 2023, *10* (2), 487–523.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–3394.

Dorsey, Jackson, "Solar Market Frictions: The Role of Platfroms and Policies," *Working Paper*, 2024.

Edelman, Benjamin, Michael Luca, and Dan Svirsky, "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment," *American Economic Journal: Applied Economics*, 2017, *9* (2), 1–22.

Environmental Protection Agency, "Solar for All," 2023. https://www.epa.gov/greenhouse-gas-reduction-fund/solar-all.

Feger, Fabian, Nicola Pavanini, and Doina Radulescu, "Welfare and redistribution in residential electricity markets with solar power," *The Review of Economic Studies*, 2022, *89* (6), 3267–3302.

Ge, Yanbo, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf, "Racial discrimination in transportation network companies," *Journal of Public Economics*, 2020, *190*, 104205.

Gerarden, Todd D, "Demanding innovation: The impact of consumer subsidies on solar panel production costs," *Management Science*, 2023.

Hahn, Robert W and Robert D Metcalfe, "Efficiency and equity impacts of energy subsidies," *American Economic Review*, 2021, *111* (5), 1658–1688.

Internal Revenue Service, "Additional Guidance on Low-Income Communities Bonus Credit Program," 2023. https://www.federalregister.gov/documents/2023/06/01/2023-11718/additional-guidance-on-low-income-communities-bonus-credit-program.

Jacqz, Irene and Sarah Johnston, "Staggered Electric Vehicle Adoption, Air Pollution Disparities, and Subsidy Policy," *Working Paper*, 2023.

Kline, Patrick and Christopher Walters, "Reasonable Doubt: Experimental Detection of Job-Level Employment Discrimination," *Econometrica*, 2021, *89* (2), 765–792.

Krasnokutskaya, Elena, Kyungchul Song, and Xun Tang, "The Role of Quality in Internet Service Markets," *Journal of Political Economy*, 2019, *128* (1), 75–117.

Langer, Ashley and Derek Lemoine, "Designing Dynamic Subsidies to Spur Adoption of New Technologies," *Journal of the Association of Environmental and Resource Economists*, 2022.

Lyubich, Eva, "The Race Gap in Residential Energy Expenditures," *Working Paper*, 2020.

Metcalf, Gilbert E, "Designing a carbon tax to reduce US greenhouse gas emissions," *Review of Environmental Economics and Policy*, 2009.

Nemet, Gregory F, Eric O'Shaughnessy, Ryan Wiser, Naïm R Darghouth, Galen Barbose, Ken Gillingham, and Varun Rai, "What factors affect the prices of low-priced US solar PV systems?," *Renewable energy*, 2017, *114*, 1333–1339.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 1973, *14* (3), 693–709.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

O'Shaughnessy, Eric, Galen Barbose, Ryan Wiser, and Sydney Forrester, "Income-targeted Marketing as a Supply-Side Barrier to Low-income Solar Adoption," *iScience*, 2021, *24* (10), 103137.

Reames, Tony G., "Distributional Disparities in Residential Rooftop Solar Potential and Penetration in Four Cities in the United States," *Energy Research & Social Science*, 2020, *69*, 101612.

Reguant, Mar, "The efficiency and sectoral distributional impacts of large-scale renewable energy policies," *Journal of the Association of Environmental and Resource Economists*, 2019, *6* (S1), S129–S168.

Seybert, Ehren, Gabe Petlin, Katie Wu, and Melicia Charles, "California Net Energy Metering Ratepayer Impacts Evaluation," *Energy and Environmental Economics Inc. Report prepared for the California Public Utilities Commission*, 2013.

Sunter, Deborah A., Sergio Castellanos, and Daniel M. Kammen, "Disparities in Rooftop Photovoltaics Deployment in the United States by Race and Ethnicity," *Nature Sustainability*, 2019, *2* (1), 71–76.

Waxman, Andrew, HR Huber-Rodriguez, and Sheila M Olmstead, "What are the likely air pollution impacts of carbon capture and storage?," *Available at SSRN*, 2023.

Yoganarasimhan, Hema, "Estimation of Beauty Contest Auctions," *Marketing Science*, 2015, *35* (1), 27–54.

A Online Appendix

A.1 Additional Tables and Figures

Table A.1: Count of Households and Bids by CBSA Location

CBSA	Bid Count	Bid Rank	Household Count	HH Rank
Los Angeles-Long Beach-Glendale, CA	32933	1	6264	1
San Diego-Carlsbad, CA	28779	2	5232	3
Oakland-Hayward-Berkeley, CA	27836	3	5650	2
Riverside-San Bernardino-Ontario, CA	21011	4	4016	5
Anaheim-Santa Ana-Irvine, CA	18684	5	3376	8
San Jose-Sunnyvale-Santa Clara, CA	17369	6	3337	9
Cambridge-Newton-Framingham, MA	14850	7	2944	11
Sacramento, CA	12654	8	2670	12
New York, NY	12571	9	3802	6
Washington-Arlington-Alexandria, DC-VA-MD-WV	12029	10	4538	4
Orlando-Kissimmee-Sanford, FL	11923	11	3047	10
Houston-The Woodlands-Sugar Land, TX	11132	12	2571	13
Tampa-St. Petersburg-Clearwater, FL	9157	13	2511	14
Phoenix-Mesa-Scottsdale, AZ	7008	14	3575	7
Chicago-Naperville-Arlington Heights, IL	5184	15	2478	15

Notes: Table lists the 15 largest core-based statistical areas on the Energy Sage platform which compose that sample.

Table A.2: Summary Statistics by Income Quintile

	Income Quintile					Total
	1	2	3	4	5	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Electricity Bill (USD/Month)	198.18	200.04	205.98	207.44	215.53	205.43
	(98.75)	(102.61)	(107.48)	(109.88)	(115.73)	(107.22)
Block Group Median Income (1K USD)	49.81	75.39	96.99	121.12	172.75	103.21
	(10.43)	(6.28)	(6.12)	(8.40)	(29.73)	(44.63)
Contract Preference Indicatos						
Loan/Lease	0.37	0.36	0.34	0.31	0.23	0.32
	(0.48)	(0.48)	(0.47)	(0.46)	(0.42)	(0.47)
Purchase/Any	0.33	0.36	0.40	0.45	0.55	0.42
	(0.47)	(0.48)	(0.49)	(0.50)	(0.50)	(0.49)
Missing	0.30	0.28	0.26	0.24	0.22	0.26
	(0.46)	(0.45)	(0.44)	(0.43)	(0.41)	(0.44)
Roof Age Indicators						
Less than 20 Years/Plan to Replace	0.62	0.63	0.63	0.64	0.64	0.63
	(0.49)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)
More Than 20 Years	0.09	0.10	0.11	0.12	0.14	0.11
	(0.28)	(0.30)	(0.32)	(0.33)	(0.35)	(0.32)
Missing	0.29	0.27	0.26	0.24	0.21	0.25
	(0.46)	(0.44)	(0.44)	(0.42)	(0.41)	(0.44)
Eqiupment Preference Indicators						
Technology/Attrative/Production	0.27	0.27	0.28	0.30	0.30	0.28
	(0.44)	(0.45)	(0.45)	(0.46)	(0.46)	(0.45)
Value	0.33	0.33	0.33	0.34	0.35	0.33
	(0.47)	(0.47)	(0.47)	(0.47)	(0.48)	(0.47)
Missing/None	0.41	0.40	0.39	0.37	0.35	0.38
	(0.49)	(0.49)	(0.49)	(0.48)	(0.48)	(0.49)
Income Quintile Lower Bound	11.625	64.242	86.146	107.298	137.321	
Income Quintile Upper Bound	64.239	86.141	107.287	137.270	250.001	
Number of HHs	11205	11201	11203	11205	11197	56011
Proportion of HH	0.200	0.200	0.200	0.200	0.200	1.000
Number of Bids	44708	46983	49392	50644	51393	243120
Proportion of Bids	0.184	0.193	0.203	0.208	0.211	1.000

Notes: We break our sample into income quintiles based on the classification of a household's median block-group income. We show the bounds for each quintile in the first two rows of the bottom panel. The average incomes within each quintile bin are \$49,810, \$75,390, \$96,990, \$121,120, and \$172,750, respectively. Sample means are reported for each group with standard deviation in parentheses below.

Table A.3: Summary Statistics by Race & Ethnicity

	API Mean (SD)	Black Mean (SD)	Hispanic Mean (SD)	Unclassified Mean (SD)	White Mean (SD)	Total Mean (SD)
Electricity Bill (USD/Month)	173.98	189.92	196.87	195.74	215.13	205.43
	(89.78)	(101.00)	(98.51)	(102.19)	(110.98)	(107.22)
Block Group Median Income (1K USD)	125.77	70.33	71.95	100.41	103.33	103.21
	(47.23)	(36.74)	(26.98)	(43.04)	(43.50)	(44.63)
Contract Preference Indicatos						
Loan/Lease	0.26	0.33	0.42	0.34	0.32	0.32
	(0.44)	(0.47)	(0.49)	(0.47)	(0.47)	(0.47)
Purchase/Any	0.51	0.31	0.23	0.38	0.43	0.42
	(0.50)	(0.46)	(0.42)	(0.49)	(0.50)	(0.49)
Missing	0.23	0.36	0.35	0.28	0.25	0.26
	(0.42)	(0.48)	(0.48)	(0.45)	(0.43)	(0.44)
Roof Age Indicators						
Less than 20 Years/Plan to Replace	0.62	0.58	0.56	0.61	0.65	0.63
	(0.48)	(0.49)	(0.50)	(0.49)	(0.48)	(0.48)
More Than 20 Years	0.15	0.08	0.10	0.11	0.11	0.11
	(0.36)	(0.26)	(0.30)	(0.32)	(0.31)	(0.32)
Missing	0.23	0.35	0.34	0.27	0.24	0.25
	(0.42)	(0.48)	(0.48)	(0.45)	(0.43)	(0.44)
Eqiupment Preference Indicators						
Technology/Attrative/Production	0.27	0.28	0.27	0.29	0.28	0.28
	(0.44)	(0.45)	(0.45)	(0.45)	(0.45)	(0.45)
Value	0.40	0.27	0.28	0.33	0.33	0.33
	(0.49)	(0.44)	(0.45)	(0.47)	(0.47)	(0.47)
Missing/None	0.33	0.45	0.44	0.38	0.39	0.38
	(0.47)	(0.50)	(0.50)	(0.49)	(0.49)	(0.49)
Number of HHs	7336	678	4026	7781	36190	56011
Proportion of HH	0.131	0.012	0.072	0.139	0.646	1.000
Number of Bids	34370	1950	18618	33481	154701	243120
Proportion of Bids	0.141	0.008	0.077	0.138	0.636	1.000

Notes: We break our sample into race/ethnicity categories as described in Appendix A.2.2. Sample means are reported for each group with standard deviation in parentheses below.

Table A.4: Number of Potential Installers By Distance

Panel A: Income Quintile

	Income Quintile				
	1	2	3	4	5
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
All Installers	36.48	37.36	37.72	37.43	36.67
	(15.22)	(15.54)	(15.09)	(14.63)	(12.92)
Number of Installers Within:					
5 Miles	0.87	0.96	0.97	1.07	1.02
	(1.58)	(1.65)	(1.58)	(1.51)	(1.41)
10 Miles	2.67	2.89	3.03	3.10	3.23
	(3.20)	(3.40)	(3.31)	(3.20)	(2.95)
25 Miles	9.95	10.64	11.23	11.49	11.70
	(7.18)	(6.89)	(6.73)	(6.28)	(5.93)
50 Miles	18.96	20.63	21.84	21.85	22.21
	(10.70)	(10.77)	(10.48)	(9.54)	(8.53)

Panel B: Race/Ethnicity

	API	Black	Hispanic	Unclassified	White
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
All Installers	39.48	31.74	42.57	37.30	36.11
	(12.79)	(18.93)	(14.61)	(15.22)	(14.69)
Number of Installers Within:					
5 Miles	1.01	0.52	0.81	0.91	1.01
	(1.42)	(0.86)	(1.59)	(1.44)	(1.60)
10 Miles	3.34	1.97	2.72	2.87	2.98
	(2.71)	(1.63)	(3.45)	(3.09)	(3.33)
25 Miles	12.34	9.73	11.35	11.03	10.71
	(5.61)	(6.72)	(7.35)	(6.70)	(6.71)
50 Miles	25.18	18.53	23.04	21.25	20.07
	(9.40)	(11.13)	(11.76)	(10.53)	(9.69)

Notes: Means reported for each group with standard deviation in parentheses below. The variable *All Installers* is constructed by counting the number of registered installers on the platform within 250 miles of the project.

Table A.5: Bidding Disparities: Names versus Neighborhoods

	(1) Log(Median Price)	(2) Log(Median Price)	(3) Log(Bids)	(4) Log(Bids)
Race/Ethnicity - Binary Measure				
Black	0.0413***		-0.0955***	
	(0.0050)		(0.0197)	
Asian/Pac. Islander	0.0072***		-0.0410***	
	(0.0011)		(0.0059)	
Hispanic	0.0008		0.0111	
	(0.0013)		(0.0073)	
Unclassified	0.0051***		0.0029	
	(0.0010)		(0.0053)	
Block Group Race/Ethnicity Prop.				
Black		0.0620***		-0.0722***
		(0.0039)		(0.0164)
Asian/Pac. Islander		0.0070***		-0.0558***
		(0.0025)		(0.0135)
Hispanic		0.0047*		0.0108
		(0.0027)		(0.0147)
Name Race/Ethnicity Prop.				
Black		-0.0112***		0.0242
		(0.0028)		(0.0151)
Asian/Pac. Islander		0.0048***		-0.0253***
		(0.0012)		(0.0062)
Hispanic		-0.0029**		0.0134*
		(0.0013)		(0.0071)
Income Quintiles				
1st Quintile	0.0218***	0.0182***	-0.0898***	-0.0926***
	(0.0012)	(0.0013)	(0.0064)	(0.0068)
2nd Quintile	0.0172***	0.0146***	-0.0634***	-0.0655***
	(0.0011)	(0.0012)	(0.0060)	(0.0063)
3rd Quintile	0.0109***	0.0089***	-0.0291***	-0.0303***
	(0.0011)	(0.0011)	(0.0057)	(0.0058)
4th Quintile	0.0069***	0.0056***	-0.0122**	-0.0128**
	(0.0010)	(0.0010)	(0.0054)	(0.0055)
Observations	56011	56011	56011	56011
R-Sq	0.483	0.486	0.452	0.452
FE	Year-by-CBSA	Year-by-CBSA	Year-by-CBSA	Year-by-CBSA

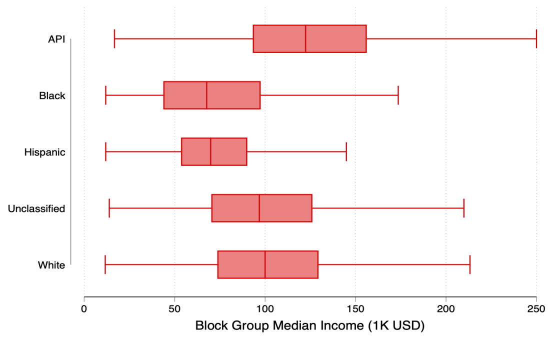
Notes: Columns 1 and 3 report the baseline regression estimates from Equation 1 using our preferred binary measure of the household race/ethnicity. Columns 2 and 4 report analogous regression results but omit the binary race/ethnicity variables as regressors and instead includes as regressors: (1) the proportion of each race/ethnicity group within the households' block group and (2) the probability that each buyer's name belongs to a race/ethnicity group. The dependent variable for columns 1-2 is the logarithm of the median bid price ($/watt) offered to a household. The dependent variable for columns 3-4 is the logarithm of the number of bids the household receives.

Table A.6: Bidding Heterogeneity By Race and Installer Rating

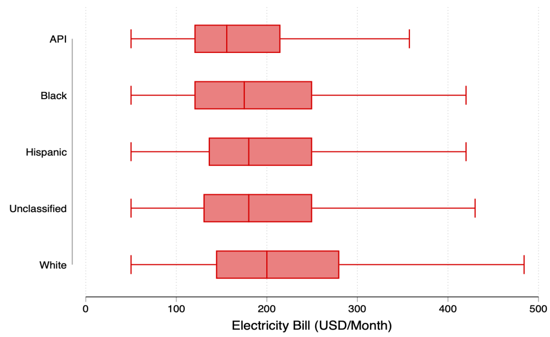
	(1) Log(Price)	(2) Log(Price)	(3) Log(Price)	(4) Log(Price)
Race/Ethnicity - Binary Measure				
Black	0.0231***	0.0169***	0.0173***	0.0120***
	(0.0032)	(0.0031)	(0.0029)	(0.0030)
Asian/Pac. Islander	0.0029***	0.0041***	-0.0036***	-0.0025***
	(0.0006)	(0.0008)	(0.0006)	(0.0007)
Hispanic	0.0017**	0.0007	-0.0033***	-0.0043***
	(0.0008)	(0.0009)	(0.0007)	(0.0009)
Interactions				
Black Owner × Five Star Installer		0.0158***		0.0138**
		(0.0061)		(0.0058)
API Owner × Five Star Installer		-0.0022**		-0.0019**
		(0.0010)		(0.0010)
Hispanic Owner × Five Star Installer		0.0019		0.0021*
		(0.0012)		(0.0012)
System Size Control				
System Size (kW)			-0.0067***	-0.0067***
			(0.0001)	(0.0001)
Observations	243103	243103	243103	243103
R-Sq	0.583	0.583	0.618	0.618
CBSA-Year FE	Yes	Yes	Yes	Yes
Installer FE	Yes	Yes	Yes	Yes

Notes: The dependent variable is the natural logarithm of the bid price submitted by a particular installer to a household. Each regression controls for CBSA-by-year fixed effects and installer fixed effects. Columns 1 and 3 regress the logged bid price on household race/ethnicity indicators (without heterogeneity), and Columns 2 and 4 interact each household race indicator with an indicator for whether the installer has a five-star rating on the platform. Columns 2 and 4 include controls for the system size. Standard errors in parentheses are clustered by household.

Figure A.1: Box Plots of Household Characteristics



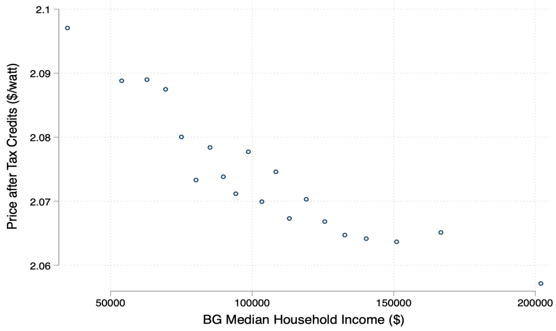
(a) Income



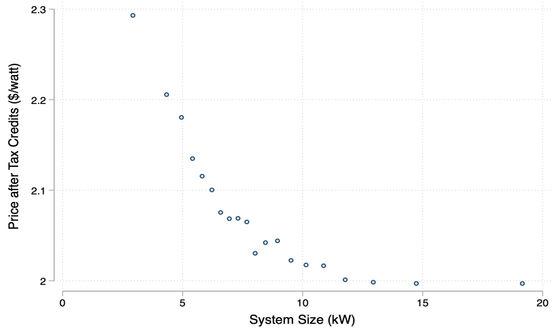
(b) Electricity Bill

Notes: Panel (a) compares the distribution of median income by race. Panel (b) compares the distribution of monthly electricity bill by race.

Figure A.2: Household Characteristics and Prices



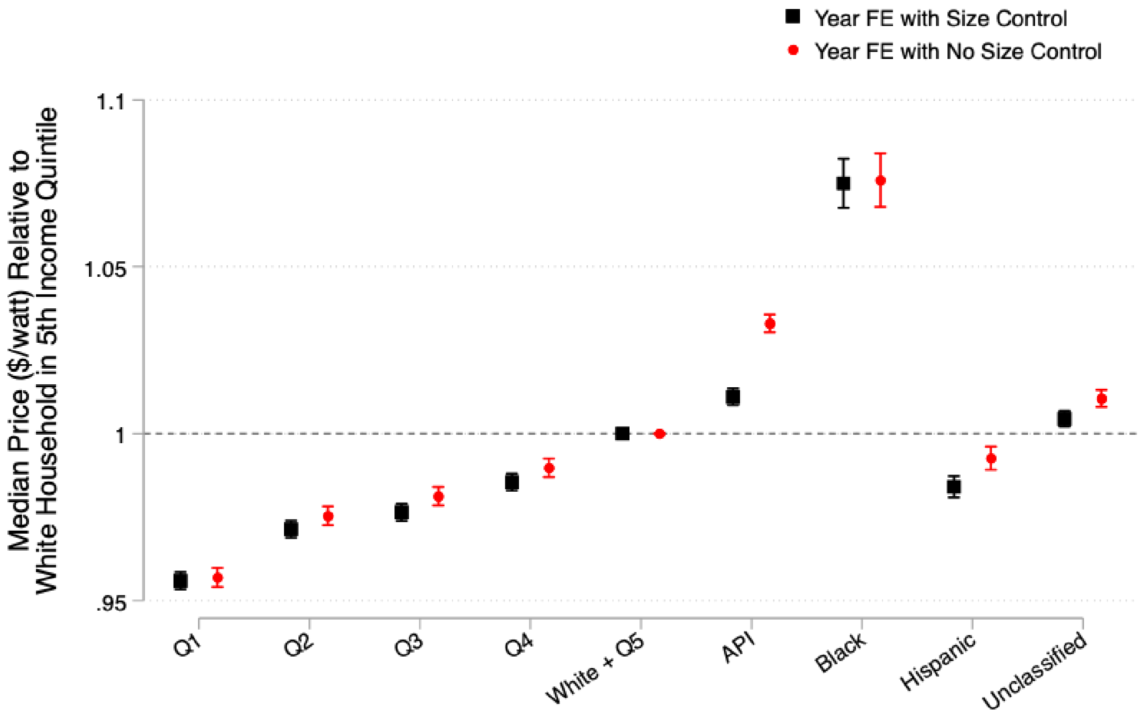
(a) Household Income



(b) System Size (kW)

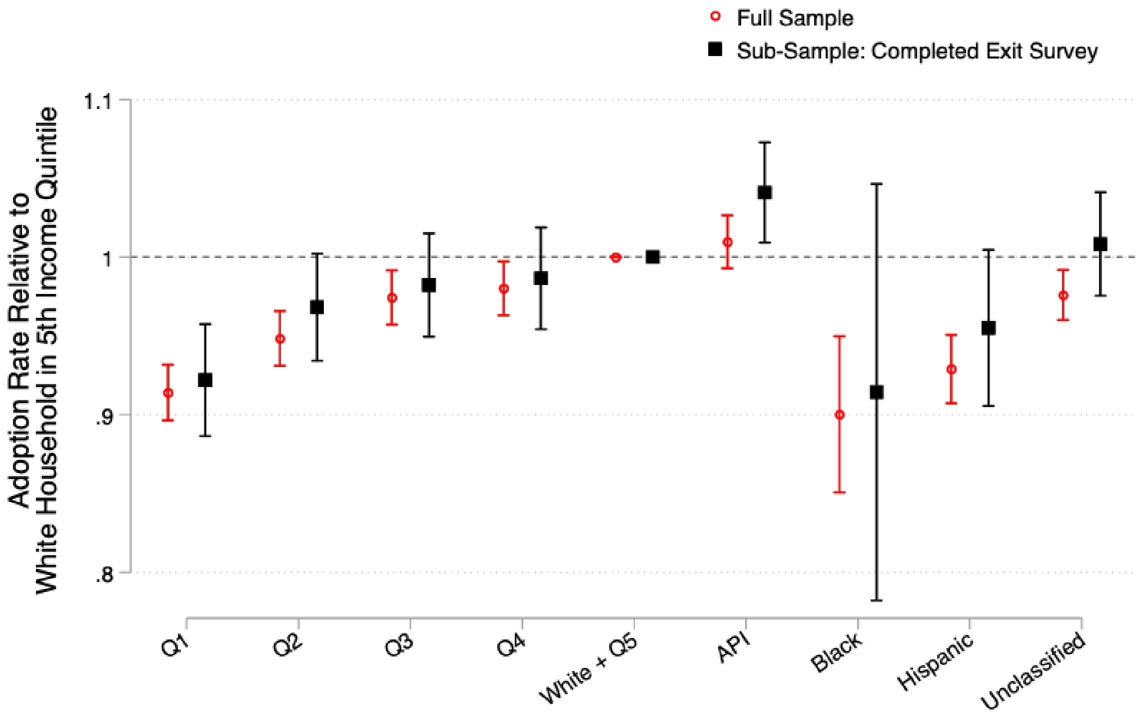
Notes: Panels (a) and (b) depict the average bid prices net of the ITC subsidy by household's block group median income and system size. Each point is calculated as the mean price among each quantile bin after controlling for CBSA fixed effects. The sample includes all bids submitted in our sample.

Figure A.3: Median Bid Price by Income & Race - Robustness to Controlling for System Size



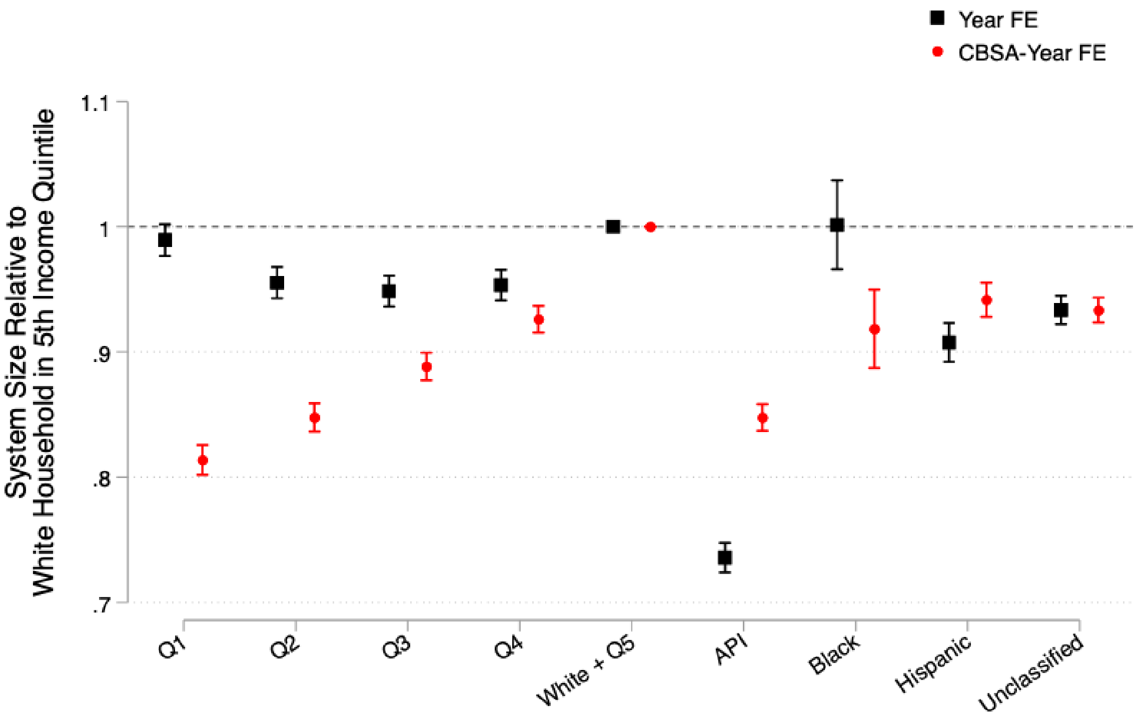
Notes: This figure compares the regression estimates from Figure 2d to an alternate specification that does not control for system size (the mean capacity bid in KW submitted to the household).

Figure A.4: Relative Overall Adoption by Income & Race - Robustness to Selection into Exit Survey



Notes: This figure compares the regression estimates from Figure 2b to an alternate specification that only includes the sample of households that completed the exit survey.

Figure A.5: Descriptive Regressions: System Size by Income & Race



Notes: This figure presents estimates of the regression in Equation 1 with the log of system size as the outcome variable (mean capacity bid submitted to the household) and omitting size as a control. We estimate the regression both with year fixed effects and CBSA-year fixed effects and normalize the coefficient estimates relative to the White households in the 5th income quintile.

Table A.7: Implied Markups and Marginal Costs by Income and Race

	Dependent variable:					
	Log(Optimal Markup)			Log(Implied Marginal Cost)		
	(1)	(2)	(3)	(4)	(5)	(6)
Income - Quintile 1	−0.170		−0.135	0.143		0.118
	(0.001)		(0.0002)	(0.001)		(0.001)
Income - Quintile 2	−0.072		−0.050	0.066		0.050
	(0.0005)		(0.0002)	(0.001)		(0.001)
Income - Quintile 3	−0.010		−0.0004	0.011		0.003
	(0.0005)		(0.0002)	(0.001)		(0.001)
Income - Quintile 4	0.004		0.008	−0.002		−0.005
	(0.0005)		(0.0002)	(0.001)		(0.001)
Black Owner		−0.178	−0.134		0.185	0.146
		(0.001)	(0.001)		(0.004)	(0.004)
Hispanic Owner		−0.275	−0.246		0.185	0.160
		(0.0005)	(0.0003)		(0.001)	(0.001)
Asian/PI Owner		−0.00002	−0.007		−0.010	−0.004
		(0.0004)	(0.0002)		(0.001)	(0.001)
Observations	243,120	243,120	243,120	243,116	243,116	243,116
R^2	0.713	0.813	0.942	0.555	0.556	0.579

Notes: The dependent variable in the first three columns is the natural log of the estimated optimal markup ($/watt), and the dependent variable in the last three columns is the natural log of the estimated marginal cost. Four bids had negative marginal cost estimates were dropped. Each regression includes all other non-price variables that enter the buyer's utility function, including panel brand dummies, installer controls, CBSA fixed effects, year fixed effects, and controls for the household's survey response.

Table A.8: Consumer Surplus by Income and Race

	Dependent variable:		
	Log(Consumer Surplus)		
Income - Quintile 1	−0.592		−0.489
	(0.010)		(0.010)
Income - Quintile 2	−0.284		−0.222
	(0.010)		(0.009)
Income - Quintile 3	−0.080		−0.051
	(0.009)		(0.009)
Income - Quintile 4	−0.019		−0.007
	(0.009)		(0.009)
Black Owner		−0.812	−0.645
		(0.026)	(0.025)
Hispanic Owner		−0.736	−0.631
		(0.011)	(0.011)
Asian/PI Owner		0.027	−0.003
		(0.009)	(0.009)
Observations	56,011	56,011	56,011
R^2	0.641	0.642	0.664

Notes: The dependent variable is the natural log of households' expected consumer surplus ($/watt). Each regression includes all other non-price variables that enter the buyer's utility function, including panel brand dummies, installer controls, CBSA fixed effects, year fixed effects, and controls for the household's survey response.

Table A.9: Price Adjustments Required to Close Consumer Surplus Gaps - Alternate Comparison Groups

Group	Uniform Price Adjustment ($/watt)	Mean Price After Adjustment ($/watt)	Mean Consumer Surplus After Adjustment ($)
Base: Low Income (Q1)	0	2.01	576.93
Comparison: High Income (Q5)	0.66	2.67	576.93
Base: Black	0	2.15	281.38
Comparison: White	1.05	3.2	281.38
Base: Hispanic	0	2.01	430.54
Comparison: White	0.7	2.71	430.54
Base: Asian/PI	0	2.14	1097.15
Comparison: White	-0.09	2.05	1097.15

Notes: The second column calculates uniform change to all bid prices (net price after the ITC) that would imply that the base group (*e.g.*, low-income households) and the comparison group (*e.g.*, high-income households) obtain equal expected consumer surplus. The third column reports the mean bid for each group after the price adjustment. The last column reports the expected consumer surplus for each group after implementing the uniform bid price adjustment.

Table A.10: Demand - Heterogeneity in Preferences for Non-Price Attributes Across Income

	(1)	(2)	(3)	(4)	(5)
λ	0.4 (0.04)	0.39 (0.04)	0.39 (0.04)	0.4 (0.04)	0.39 (0.04)
Price	-1.01 (0.11)	-1.16 (0.14)	-1.1 (0.13)	-0.99 (0.11)	-0.97 (0.11)
Price $\times$ Income Quintile 1	-0.12 (0.03)	0.27 (0.19)	0.12 (0.15)	-0.16 (0.07)	-0.22 (0.08)
Log(Electric Bill)	-0.56 (0.06)	-0.58 (0.07)	-0.53 (0.06)	-0.56 (0.06)	-0.56 (0.06)
Income Quintile 1 $\times$ Log(Electric Bill)		0.04 (0.12)	-0.1 (0.06)		
Has Off-platform Quotes	0.24 (0.06)	0.22 (0.07)	0.21 (0.07)	0.24 (0.06)	0.24 (0.06)
Income Quintile 1 $\times$ Has Off-platform Quotes		0.06 (0.12)	0.1 (0.12)		
Roof Age $\leq$ 20 years	0.36 (0.2)	0.66 (0.25)	0.34 (0.2)	0.31 (0.21)	0.37 (0.2)
Income Quintile 1 $\times$ Roof Age $\leq$ 20 years		-0.75 (0.42)		0.14 (0.16)	
Roof Age $>$ 20 years	0.18 (0.21)	0.55 (0.26)	0.15 (0.21)	0.2 (0.22)	0.18 (0.21)
Income Quintile 1 $\times$ Roof Age $>$ 20 years		-1.02 (0.44)		-0.14 (0.21)	
Ownership Preference: Lease/Loan	0.63 (0.21)	0.21 (0.26)	0.66 (0.21)	0.63 (0.21)	0.55 (0.23)
Income Quintile 1 $\times$ Ownership Preference: Lease/Loan		1.05 (0.44)			0.18 (0.22)
Ownership Preference: Cash Purchase	0.87 (0.21)	0.43 (0.25)	0.9 (0.21)	0.87 (0.21)	0.77 (0.22)
Income Quintile 1 $\times$ Equipment Preference: Premium		0.07 (0.16)			0.03 (0.16)
Equipment Preference: Premium	0.24 (0.08)	0.21 (0.1)	0.23 (0.08)	0.24 (0.08)	0.22 (0.1)
Income Quintile 1 $\times$ Equipment Preference: Premium		0.08 (0.17)			0.06 (0.17)
Equipment Preference: Value	0.23 (0.08)	0.2 (0.1)	0.23 (0.08)	0.23 (0.08)	0.22 (0.1)
Income Quintile 1 $\times$ Equipment Preference: Value		0.07 (0.16)			0.03 (0.16)
Panel Rating = Excellent	0.54 (0.09)	0.53 (0.11)	0.54 (0.09)	0.54 (0.09)	0.54 (0.09)
Income Quintile 1 $\times$ Panel Rating = Excellent		0.06 (0.16)			
Panel Rating = Fair/Poor	-1.01 (0.5)	-0.83 (0.65)	-1.01 (0.5)	-1.01 (0.5)	-1.02 (0.5)
Income Quintile 1 $\times$ Panel Rating = Fair/Poor		-0.44 (1.03)			
Panel Rating = Good	-0.41 (0.14)	-0.3 (0.17)	-0.42 (0.14)	-0.41 (0.14)	-0.42 (0.14)
Income Quintile 1 $\times$ Panel Rating = Good		-0.32 (0.27)			
Panel Rating = Very Good	-0.06 (0.1)	-0.05 (0.13)	-0.06 (0.1)	-0.06 (0.1)	-0.06 (0.1)
Income Quintile 1 $\times$ Panel Rating = Very Good		0 (0.19)			
Controls					
CBSA FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Permanent Installer FE	Y	Y	Y	Y	Y
Transient Installer Star Rating FE	Y	Y	Y	Y	Y
Transient Installer Number of Ratings FE	Y	Y	Y	Y	Y

Notes: The sample for all specifications includes households in the top and bottom income quintiles. All utility specifications include CBSA, year, and permanent seller fixed effects. Permanent sellers are those who submitted over 1000 total bids. For transient sellers, we include a set of fixed effects for the installer's star rating and the installer's number of reviews. All specifications allow for heterogeneity in the price coefficients across income. The second specification allows for heterogeneity in the coefficients associated with the household survey responses, which we allow to shift the utility of all of the installation bids (*e.g.* the inside options). For the survey responses that include dummy variables, the omitted group represents buyers who did not answer the survey question. For example, the "Roof Age: $<=$ 20 years" variable is relative to buyers who did not report the age of their roof. The second specification also allows for heterogeneity in tastes for panel quality across income. Columns 3-5 include selected subsets of the interaction terms between income and other variables. Standard errors are in parentheses.

Table A.11: Detailed Decomposition of Consumer Surplus Gap: High- vs. Low-Income Households

	Estimate
Difference in Mean Consumer Surplus ($)	652.01
Difference Explained by System Size	66.42
Share Explained by System Size	0.1
Difference Explained by Price Coefficient	299.77
Share Explained by Price Coefficient	0.46
Difference Explained by Household Survey Response Covariates	124.38
Share Explained by Household Survey Response Covariates	0.19
Difference Explained by CBSA Market and Year Fixed Effects	233.72
Share Explained by CBSA Market and Year Fixed Effects	0.36

Notes: The first row is the mean consumer surplus gap between the two groups of households. The middle section of each panel further decomposes the gap in consumer surplus into a portion explained by system size, price sensitivity, household survey responses that shift utility for the inside options (such as roof age, financing preference, and monthly electricity expenditure), and CBSA and year fixed effects that also shift utility for the inside options. Note that each element of these decompositions is performed independently, and therefore, the shares will not sum to one.

Table A.12: Detailed Decomposition of Consumer Surplus Gap: White vs. Black Households

	Estimate
Difference in Mean Consumer Surplus ($)	706.65
Difference Explained by System Size	1.93
Share Explained by System Size	0
Difference Explained by Price Coefficent	164.26
Share Explained by Price Coefficent	0.23
Difference Explained by Household Survey Response Covariates	52
Share Explained by Household Survey Response Covariates	0.07
Difference Explained by CBSA Market and Year Fixed Effects	66.38
Share Explained by CBSA Market and Year Fixed Effects	0.09

Notes: The first row is the mean consumer surplus gap between the two groups of households. The middle section of each panel further decomposes the gap in consumer surplus into a portion explained by system size, price sensitivity, household survey responses that shift utility for the inside options (such as roof age, financing preference, and monthly electricity expenditure), and CBSA and year fixed effects that also shift utility for the inside options. Note that each element of these decompositions is performed independently, and therefore, the shares will not sum to one.

Table A.13: Detailed Decomposition of Consumer Surplus Gap: White vs. Hispanic Households

	Estimate
Difference in Mean Consumer Surplus ($)	557.49
Difference Explained by System Size	89.32
Share Explained by System Size	0.16
Difference Explained by Price Coefficient	449
Share Explained by Price Coefficient	0.81
Difference Explained by Household Survey Response Covariates	98.93
Share Explained by Household Survey Response Covariates	0.18
Difference Explained by CBSA Market and Year Fixed Effects	48.81
Share Explained by CBSA Market and Year Fixed Effects	0.09

Notes: The first row is the mean consumer surplus gap between the two groups of households. The middle section of each panel further decomposes the gap in consumer surplus into a portion explained by system size, price sensitivity, household survey responses that shift utility for the inside options (such as roof age, financing preference, and monthly electricity expenditure), and CBSA and year fixed effects that also shift utility for the inside options. Note that each element of these decompositions is performed independently, and therefore, the shares will not sum to one.

Table A.14: Detailed Decomposition of Consumer Surplus Gap: White vs. API Households

	Estimate
Difference in Mean Consumer Surplus ($)	-109.12
Difference Explained by System Size	410.65
Share Explained by System Size	-3.76
Difference Explained by Price Coefficient	-36.84
Share Explained by Price Coefficient	0.34
Difference Explained by Household Survey Response Covariates	-17.24
Share Explained by Household Survey Response Covariates	0.16
Difference Explained by CBSA Market and Year Fixed Effects	-122.1
Share Explained by CBSA Market and Year Fixed Effects	1.12

Notes: The first row is the mean consumer surplus gap between the two groups of households. The middle section of each panel further decomposes the gap in consumer surplus into a portion explained by system size, price sensitivity, household survey responses that shift utility for the inside options (such as roof age, financing preference, and monthly electricity expenditure), and CBSA and year fixed effects that also shift utility for the inside options. Note that each element of these decompositions is performed independently, and therefore, the shares will not sum to one.

Table A.15: Decomposition of Consumer Surplus Gap - Swapping Base Group

Panel A: High-Low Income Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	652.01	(117.28)
Demand Component ($)	663.38	(111.72)
Share Explained by Demand Component	1.02	(0.05)
Supply Component ($)	44.29	(19.83)
Share Explained by Supply Component	0.07	(0.03)
Interaction Component ($)	-55.65	(24.35)
Share Explained by Interaction Component	-0.09	(0.04)

Panel B: Low-High Income Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	-652.01	(117.28)
Demand Component ($)	-617.95	(112.72)
Share Explained by Demand Component	0.95	(0.03)
Supply Component ($)	23.82	(31.25)
Share Explained by Supply Component	-0.04	(0.05)
Interaction Component ($)	-57.88	(24.48)
Share Explained by Interaction Component	0.09	(0.04)

Notes: The second row of each panel show the mean consumer surplus gap between the two groups of households (*e.g.* Black mean consumer surplus minus White mean consumer surplus). The middle section of each panel decomposes the gap in consumer surplus into a portion explained by price sensitivity and a portion that is unexplained by price sensitivity. The top row describes the reference group used to measure the decomposition. The bottom section of each panel shows the gap in the mean number of bids received across the two groups. Bootstrapped standard errors are in parentheses.

A.2 Data Appendix

A.2.1 Installer Data

We do not observe firms' exact identities in the data (*i.e.* firm names or detailed locations). However, we observe the distance between the installer and the potential buyer for each bid in the dataset, which we can use in conjunction with the household location data to infer installers' approximate locations. Given that we observe household locations at the block group level, if a given installer bids on households in three distinct census blocks, we can use this triplet of distances to infer installer locations based on trilateration. We conduct this process at the CBSA level and restrict installers within 250 miles of the household to be included in the trilateration exercise. This procedure requires a minimum of three bids across different block groups–however–to improve fit for those bidding in more than three block groups, we use non-linear least squares to find the location for each installer that minimizes the residual distance for all bids for that installer.

Lastly, we observe sellers' ratings on the EnergySage platform and use them to measure installer quality. As is standard on many online marketplaces, buyers can rate sellers based on their interactions on the EnergySage platform. EnergySage aggregates this information via star ratings between 0 and 5 and then displays these ratings to potential buyers on the platform. We observe these ratings in our data, which we use to control for installer quality throughout our analysis.

A.2.2 Inferring Buyers' Race/Ethnicity From Names and Locations

We follow the two-step approach used in Diamond et al. (2019) to determine households' race/ethnicity. In the first step, we use US Census data that provides the distribution of ethnic identities associated with thousands of common surnames to assign a probabilistic distribution of racial/ethnicity to each buyer in the data.[34] In the second step, we update this distribution based on the household's last name using the racial composition of homeowners in the buyer's census block using Bayes' rule. We calculate the probability that a buyer belongs to race or ethnicity r conditional on having name s and living in census block g as:

$$P(r \mid g, s) = \frac{P(r \mid s)P(g \mid r)}{\sum_{r' \in R} P(r' \mid s)P(g \mid r')} \tag{A.1}$$

where R denotes the set of six possible race/ethnic categories–Black, White, Asian and Pacific Islander (API), American Indian or Alaska Native (AIAN), Hispanic and other.

Since the US Census Bureau measures race and ethnicity separately in the ACS these variables are subject to overlap. To account for this data feature, we make assumptions to en-

[34] EnergySage cannot release each buyer's name based on their privacy terms and conditions; however, they did match each household by last name to the US Census database on racial and ethnic population shares.

sure our race and ethnic probabilities sum to one. In practice, we build this distribution of race and ethnicity so that the racial measures only include households who identify as that race and *not* Hispanic. However, as noted previously, we want these distributions to reflect homeowners only. Since the publicly available ACS data does not report the trivariate distribution of race-by-ethnicity-by-homeowner at the block-group level, we construct this distribution using the two bivariate distributions of race-by-ethnicity and race-by-homeowner and an assumption—that the race-by-ethnicity distribution for homeowners is the same as the entire block group (including renters). Using this assumption, we can construct the race-by-ethnicity-by-homeowner distribution and take conditional probabilities to create ethnic and racial measures that are mutually exclusive. We use this constructed distribution to *net out* any overlap between Hispanic-identifying households and each race, as in Diamond et al. (2019).[35]

Finally, and following Diamond et al. (2019) again, we use the resultant proportions for each household to create binary measures of race and ethnicity equal one if the Bayesian probability for that race or ethnicity is 0.8 or greater. If no race or ethnicity passes this threshold, we define that household as "Unclassified". Notably, we omit the American Indian or Alaskan Native group from our analysis, given that only a handful of these observations are in our sample.

A.3 Details on Installers' Bidding Problem

Each firm j solves the following problem when setting a bid price for project i:

$$\max_{B_{ij}} \quad q_{ij}[B_{ij} - c_{ij}] \cdot \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) \tag{A.2}$$

Where q_{ij} is the system capacity, B_{ij} is firm j's per-unit price bid, and c_{ij} is firm j's marginal cost.[36] $\mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)$ is the equilibrium expected probability of winning the auction conditional on placing a bid price of B_{ij}. The equilibrium expected probability of being selected is also a function of the project type $\mathbf{z}_i$, the seller's type $\mathbf{w}_j$, and the non-price characteristics of the bid $\mathbf{x}_{ij}$. The project type, $\mathbf{z}_i$, is characterized by the geographic market where the project is located, the time period, and the household's characteristics. We categorize sell-

[35] As an example, consider a block group that is 40% Hispanic with a homeowner-only racial distribution of 60% Black and 40% API. Now assume that 25% of all (*i.e.* both renter and homeowner) API households identify as Hispanic and 50% of all Black households identify as Hispanic. Then of the 40% Hispanic households–30% of this represents dual-identifying Black and Hispanic households, and the remaining 10% represents API and Hispanic households. With the constructed race-by-ethnicity-by-homeowner distribution, we simply calculate conditional probabilities to determine that this block group is 40% Hispanic, 30% non-Hispanic Black, and 30% non-Hispanic API. We omit the "non-Hispanic" qualifier for ease of exposition when discussing impacts across different racial groups.

[36] Here, c_{ij}, is installer j's average cost-per-watt from installing household i's system, relative to not performing the installation.

ers into types $\mathbf{w}_j$ using a relatively parsimonious measure using either the seller's ratings and reviews or seller-specific indicators (*i.e.,* seller fixed effects).[37]

We work with expected probabilities since the seller does not know exactly which competitors they will face nor the bids of those competitors. We note that the solution to the bid pricing problem is not a function of the system capacity realization, q_{ij} enters the expected profit function multiplicatively and, therefore, does not directly influence the optimal per-unit bid price. However, the system capacity can indirectly affect the price bid if system capacity and marginal cost are correlated.

When formulating firms' expectations, we assume that all sellers submit their bids simultaneously. Therefore, the installers do not know the exact number of bidders they will compete against nor the identities of their competitors. Thus, firms' expectations about the probability of winning will only be a function of the project characteristics, conditional on the price and non-price characteristics of their bid.[38]

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

$$\begin{aligned}&\mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) = \mathbb{E}[Prob_{ij}(B_{ij}; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)] = \\ &\int Prob_{ij}(B_{ij}; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) \cdot dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{z}_i)\end{aligned} \tag{A.3}$$

Recall that $Prob_{ij}$ is the probability that buyer i selects firm j's bid conditional on realized vector of competing price bids $\mathbf{B}_{i,-j}$, having a stacked vector of non-price characteristics $\mathbf{X}_{i,-j}$, and having types $\mathbf{W}_{-j}$. G represents the joint distribution function of $\mathbf{B}_{i,-j}$, $\mathbf{X}_{i,-j}$, and $\mathbf{W}_{-j}$ occurring in equilibrium, conditional on the project being of type $\mathbf{z}_i$.[39] 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 $\mathbf{x}_{il}$.

We define the optimal bid function as $B^*_{il}(c_{il} \mid \mathbf{x}_{il}, \mathbf{w}_l, \mathbf{z}_i)$ and the probability a potential seller

[37] Our model is static and therefore rules out dynamic incentives for both buyers and sellers. Dorsey (2024) provides a more detailed discussion on why a static demand model is likely to provide a reasonable approximation in this setting because solar installation prices have become relatively stable during the sample period relative to the early 2010s. However, a limitation of this framework is that it rules out the possibility of dynamic pricing incentives that may arise due to peer effects (Bollinger et al., 2022) or learning-by-doing (Bollinger and Gillingham, 2019).

[38] 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. Dorsey (2024) provides evidence that the assumption reasonably approximates firms' behavior by showing that sellers' bids do not vary systematically depending on the bid's submitted order.

[39] G is only a function of $\mathbf{z}_i$ because a seller's type $\mathbf{w}_j$ and non-price characteristics $\mathbf{x}_{ij}$ are private information at the time of bidding.

l of type $\mathbf{w}_l$ enters an auction of type $\mathbf{z}_i$ as $H(\mathbf{w}_l, \mathbf{z}_i)$. Thus, we obtain:

$$dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-\boldsymbol{j}} \mid \mathbf{z}_i) = \prod_{l \in \mathcal{N}(\mathbf{z}_i) \setminus \{j\}} H(\mathbf{w}_l, \mathbf{z}_i) \cdot dF_{CX|\mathbf{w}_l, \mathbf{z}_i}\big(B^{*-1}(B_{il} \mid \mathbf{x}_{il}, \mathbf{w}_l, \mathbf{z}_i), \mathbf{x}_{il} \mid \mathbf{w}_l, \mathbf{z}_i\big) \tag{A.4}$$

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

$$(B_{ij} - c_{ij}) \frac{\partial \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)}{\partial B_{ij}} + \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) = 0 \tag{A.5}$$

Given a vector of non-price characteristics, Equation A.5 implicitly defines the optimal bid function $B^*_{ij}(c_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)$.

We follow Yoganarasimhan (2015) and do not impose structural assumptions on sellers' entry decisions in estimation. In principle, it is possible to explicitly model auction entry decisions as in Dorsey (2024) and Krasnokutskaya et al. (2019). However, EnergySage changed its rules and commission structure starting in 2019 in a way that makes the entry incentives asymmetric across sellers. As such, we estimate flexible reduced-form entry probabilities for $H(\mathbf{w}_l, \mathbf{z}_i)$ instead of modeling the underlying micro foundation for these entry probabilities. This approach is appealing for tractability but does not allow us to estimate counterfactual changes in auction entry behavior.

A.3.1 Equilibrium

For each seller j, a strategy consists of a bidding function $\mathbf{w} \times \mathbf{z} \times \mathbf{x} \times c \to \mathbb{R}_+$. In particular, firms consider the project type, their seller type, their marginal cost draw, and their non-price characteristics to form a price bid. We 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 type possessing the same non-price characteristics use the same bidding strategy in equilibrium. An equilibrium requires that all firms satisfy Equation 4 given the other installer's strategies. Krasnokutskaya et al. (2019) proves the existence and uniqueness of a type-symmetric pure strategy Bayesian equilibrium of this game.[40] The next section describes the estimation procedure in detail.

[40] Krasnokutskaya et al. (2019) and Dorsey (2024) also model the firms' entry decision; there is no guarantee of a unique equilibrium in the participation stage of the game.

A.4 Procedure for Inferring Markups and Marginal Costs

1. First, we estimate each installer's entry probability for each project. An installer's entry probability depends on both the installer's own characteristics and the characteristics of the project $\{\mathbf{w}_j, \mathbf{z}_i\}$. We approximate the conditional probability of entry by estimating the following logistic regression model:

$$Prob(Enter_{ij}) = \mathbf{z}_i'\gamma_z + \mathbf{w}_j'\gamma_w \tag{A.6}$$

   We assume a seller is a potential entrant for auction i if they entered at least one auction within the same CBSA in the same year as project i.

2. Estimate the conditional probability that a seller who enters auction i offers a particular set of non-price characteristics $\mathbf{x}_{ij}$. Specifically, the non-price characteristics of the bid indicate the quality of the solar panels offered which include the following categories: "Excellent", "Very Good", "Good", "Fair/Poor", and "Missing Rating". We approximate the conditional probability of offering non-price characteristics $\mathbf{x}_{ij}$ using the following multinomial logistic regression using the full sample of observed bids:

$$Prob(\mathbf{x}_{ij}) = \mathbf{z}_i'\theta_z + \mathbf{w}_j'\theta_w. \tag{A.7}$$

3. Estimate the expected bid price that each entrant j would offer conditional on entering an auction i and having a vector of non-price characteristics $\mathbf{x}_{ij}$ using the following linear regression from the full sample of observed bids:

$$B_{ij} = \mathbf{z}_i'\psi_z + \mathbf{w}_j'\psi_w + \mathbf{x}_{ij}'\psi_x + \epsilon_{ij}. \tag{A.8}$$

4. Next, we use the conditional probabilities estimated in Step 1 to simulate the entry decisions for auction i for each potential entrant in $\mathcal{N}(\mathbf{z}_i)$.

5. Simulate the set of non-price characteristics for each of the simulated entrants using Equation A.7.

6. Simulate the bid price for each simulated entrant as the $\widehat{B_{ij}} + \widehat{\epsilon_{ij}}$. In particular, we simulate the bid price as the predicted value from Equation A.8 plus a residual drawn from the error distribution of the regression.[41]

7. 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 competitor's observed characteristics.

8. Repeat the second through fourth step S times each and take the average of all the sim-

[41] In the baseline model, we assume that ϵ_{ij} is i.i.d and normally distributed. We experimented with more flexible error distributions and found that they had little impact on the estimated markups and costs.

ulated choice probabilities, and simulated demand semi-elasticities to obtain estimates for the two expectations.[42] Let s denote the simulation iteration, we define the relevant expressions as:

$$\widehat{\mathcal{P}_{ij}} = \frac{1}{S}\sum_{s=1}^{S} Prob_{ij}^{s}, \qquad \widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}} = \frac{1}{S}\sum_{s=1}^{S} \frac{\partial Prob_{ij}^{s}}{\partial B_{ij}} \tag{A.9}$$

9. Finally, use the average choice probabilities, and average demand semi-elasticities from the previous step 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:

$$\widehat{c_{ij}} = B_{ij} + \frac{\widehat{\mathcal{P}_{ij}}}{\widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}}} \tag{A.10}$$

This procedure allows us to infer a project-specific marginal cost for every bid in the data. With these estimates, we can explicitly evaluate differences in average markups and marginal costs across projects belonging to different demographic groups.

References

Bollinger, Bryan and Kenneth Gillingham, "Learning-by-doing in solar photovoltaic installations," *Available at SSRN 2342406*, 2019.

_ , _ , A Justin Kirkpatrick, and Steven Sexton, "Visibility and peer influence in durable good adoption," *Marketing Science*, 2022, *41* (3), 453–476.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–3394.

Dorsey, Jackson, "Solar Market Frictions: The Role of Platfroms and Policies," *Working Paper*, 2024.

Krasnokutskaya, Elena, Kyungchul Song, and Xun Tang, "The Role of Quality in Internet Service Markets," *Journal of Political Economy*, 2019, *128* (1), 75–117.

Yoganarasimhan, Hema, "Estimation of Beauty Contest Auctions," *Marketing Science*, 2015, *35* (1), 27–54.

[42]We simulate 100 iterations of each auction type.