Unequal Uptake: Assessing Distributional Disparities in the Residential Solar Market

Jackson Dorsey*

Derek Wolfson†

August 15, 2023

Abstract

We examine technology adoption and consumer welfare disparities across demographic groups using data from an online solar photovoltaic (PV) marketplace. First, we show that conditional on visiting the online platform, low-income, Black, and Hispanic households are relatively less likely to adopt solar. Moreover, we estimate high-income households' expected consumer surplus ($1,755) is more than double that of low-income households ($824). Similarly, White and Asian Households obtain substantially higher consumer surplus than Black and Hispanic households. While some of these welfare differences can be attributed to varied willingness to pay across demographics, a significant portion is due to supply-side dynamics. In particular, low-income and Black households face higher prices and receive fewer bids from installers. If bid prices submitted to low-income households decreased by $0.81 per watt–approximately 38%—then low-income households would obtain the same expected consumer surplus as high-income households.

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, policymakers have earmarked billions of dollars in subsidies for electric vehicles, solar photovoltaic (PV) panels, and energy efficiency retrofits to encourage clean energy investments. These policies have accelerated adoption of these emerging technologies. However, many of these 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).[1] Consequently, the U.S. Environmental Protection Agency (EPA) recently launched a $7 billion "Solar for All" grant competition to increase access to affordable solar energy for low-income households (Environmental Protection Agency, 2023). Furthermore, the U.S. Inflation Reduction Act (IRA) formulated a provision for low- and middle-income individuals—the LMI adder—to provide supplementary renewable energy tax credits (Internal Revenue Service, 2023).[2] These programs underscore a rising emphasis on mitigating distributional inequities in clean energy investment (Reames, 2019).

A distinctive feature of many markets for new energy technologies—such as home energy efficiency retrofits, residential battery storage, and rooftop solar PV—is that contractors customize and price projects on an individualized basis. When services are quoted on a case-by-case basis, sellers possess considerable leeway in whether to serve a customer and how much to charge if they do. Inherent in these settings, is the possibility that sellers use household or neighborhood characteristics—such as income, race, or ethnicity—to determine which customers to serve or to adjust bid prices. Thus, technology adoption disparities may derive from 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. Therefore, a key component of our study is disentangling the relative importance of supply-side and demand-side factors in explaining technology adoption and welfare disparities.

In Figure 1, we present a stylized example demonstrating how supply-side and demand-side factors can 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

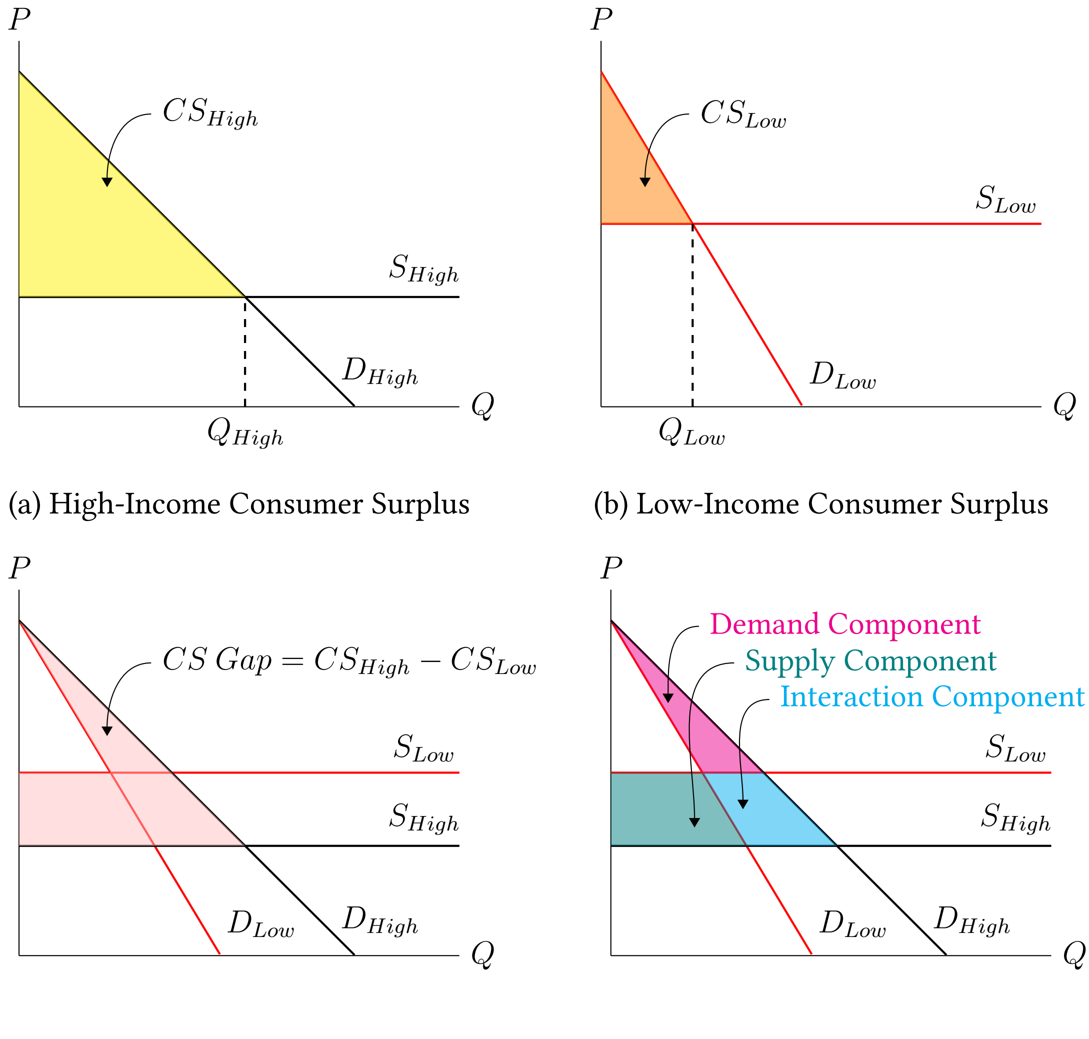
[1] In addition, Lyubich (2020) and Bednar and Reames (2020) show that low-income and minority households pay relatively more for electricity and are more likely to be burdened by energy costs.

[2] The proposed LMI adder offers low- and middle-income households with tax credits ranging between 10 to 20 percent of the purchase price of a solar PV system in addition to the 30% Investment Tax Credit (ITC) already available for residential solar.

more elastic demand and also face higher prices. In practice, low-income households may face higher prices for many reasons. For example, low-income households tend to purchase smaller solar arrays (with higher marginal costs), and they live, on average, further away from solar installers.[3]

Unsurprisingly, Figure 1a and Figure 1b demonstrate that high-income households 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*.

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.

[3]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, 2022).

To reduce distributional inequities, a thorough evaluation of consumer surplus disparities is critical for crafting effective policies. 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 the consumer surplus gap.

Figure 1d further decomposes the consumer surplus gap into three components: a *demand 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. If the demand component primarily drives the consumer surplus gap, then demand-side policies may be relatively appealing. Such policies include Pigouvian taxes, product-market subsidies, or behavioral nudges (*e.g.,* information provision, reminders, social comparisons, or default options).

In this paper, we investigate socioeconomic and demographic disparities in the residential solar PV market using detailed data on contractors' bids from a major online marketplace. 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 and make several contributions to the related literature: (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.

To analyze distributional outcomes in the solar PV market, we collect data from the EnergySage online platform. This platform serves as an intermediary between solar PV buyers and installers. In particular, solar PV shoppers can solicit bids from solar installers by uploading information about their name, monthly household energy usage, and home address through the website. The platform provides installers with information about each buyer's project, and then installers decide whether or not to submit a bid on each project. Bidding installers can customize their quote for each unique project by adjusting the price or quality of the chosen hardware components for each bid. Consequently, sellers may bid differently based on their perceptions of a buyer's income, race, or ethnicity. An advantage of using the platform data is that we can observe how each seller's bidding decisions—both the participation choice and quoted price—change across households. A key innovation in these data is our ability to observe all bid offers made to households–not just successful transactions—allowing us to account for the bidding behavior of sellers, explicitly. This is especially valuable considering that most other solar PV data sets only record transactions and do not include any information on households who ultimately decide not to adopt the technology.

We begin with a descriptive analysis documenting that solar PV purchase rates are indeed lower for low-income and minority households in our data. For instance, the platform purchase rate is 18% lower for Black households compared to White or Asian and Pacific Islander (API) households, and this same gap is nearly 30% for Hispanic households.[4] Moreover, low-income households are 25% less likely to purchase a solar PV system through the platform than high-income households.[5]

We further show that minority and low-income households obtain relatively fewer bids and higher (per-unit) bid prices. For example, compared to high-income White households, a low-income, Black household receives about 17% fewer bids and 5% higher prices. On the other hand, Hispanic households with similar income to White households do not face this same barrier–they receive a similar number of bids and prices comparable to White households. This underscores the considerable heterogeneity in effects across socioeconomic and demographic groups.

Given that bids are determined by the equilibrium behavior of both sellers and buyers, it is not clear whether disparities in the number of bids obtained across households are driven by supply-side fundamentals (*e.g.* variation in installation costs or sellers' preferences) or demand-side fundamentals (*e.g.* heterogeneity in buyers' willingness-to-pay). Namely, disparities in the number of bids obtained across households may arise even if installation costs were constant across households and the sellers were unbiased with respect to buyers' de-

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

[5]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.

mographic characteristics. If sellers with market power know that a group of buyers is more price-sensitive, they may bid less frequently to those buyers simply because bidding to them is less profitable. On the other hand, supply-side fundamentals could cause disparities in bidding participation or pricing if heterogeneity in installation costs exists across groups. Variation in costs could exist due to differences in travel, materials, labor costs, or other "implicit" costs across installation projects. Implicit costs could arise if, for example, sellers possess preferences over whether to serve specific households or neighborhoods—either because of intolerant views or due to crime rates or other variables correlated with household demographic characteristics.

We develop and estimate a structural model of the solar installation market to evaluate the welfare consequences and mechanisms explaining the aforementioned disparities in solar PV adoption. In the model, prospective solar PV buyers arrive at the platform, and then potential sellers learn about the project's characteristics—the location of the home seeking the solar installation, the household's monthly electricity use, and other household demographics, such as race and income. Bidding sellers learn their marginal cost of installing the solar system for that particular household and then submit a price bid to maximize expected profits. An optimal bid is a function of the installer's marginal cost, the household's price elasticity, and expectations about the number of competing bids for that project. After the sellers submit their bids, the household chooses one of the bidding sellers or the outside option—choosing an off-platform installer or not installing a solar system at all. 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 (2022) to estimate the model.

The structural model allows us to address several of our main research questions. First, the model allows us to quantify the disparities in expected consumer surplus across socioeconomic and demographic groups—such as the Black-White consumer surplus gap. 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. The model estimates reveal substantial disparities in consumer welfare. For instance, low-income households receive over 50% lower consumer surplus than high-income households. Similarly, White and Asian households obtain over double the expected consumer surplus of Black and Hispanic households.

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

We first decompose the consumer surplus disparity between low-income and high-income households. Low-income households are significantly more price sensitive than high-income households and tend to receive fewer bids from installers. Our decomposition shows that 43% ($405) of the consumer surplus gap between the groups is explained by the supply component, the demand component explains 37% ($340), and 20% ($186) is explained by the interaction component. We next decompose the consumer surplus gap between race and ethnic groups. The average Black household that arrives on the platform obtains a $1,010 lower expected consumer surplus than the average White household. Although Black households are substantially more price elastic than White households, we find that only 18% ($187) of the consumer surplus gap between White and Black households can be explained by differences in demand alone. In contrast, 56% ($568) of the gap arises because White households obtain preferable choice sets with more bids, lower prices, and better quality installers. Hispanic households also obtain $796 less surplus than White households. However, Hispanic households typically have choice sets that look comparable to White households, and therefore we find that 68% of the gap in consumer surplus between Hispanic and White households is explained by differences in price sensitivity.

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 38%—$0.81 per watt—to achieve

[6] While the intuition and motivation for our approach are similar to the seminal 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.

parity in consumer surplus with their high-income counterparts.

Overall, our results underscore large distributional differences in adoption and welfare among our sample of solar PV buyers. Moreover, we find that these disparities are 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 suggests that supply-side policies that reduce firms' costs of entry in underserved and disadvantaged communities may be complementary to consumer subsidies in reducing disparities in adoption and consumer welfare.[7]

Our paper relates to a broader theoretical literature on discrimination and inequality beginning with Becker (1971) and followed by Arrow (1971), Phelps (1972) and Aigner and Cain (1977). More recently, a breadth of empirical evidence indicates disparities in market outcomes among minority and low-income individuals. A recent report by Brouillette et al. (2021) estimates that Black households obtain 43% of the welfare of White households in the United States.

In addition, a wide and varied literature documents discrimination in the labor market.[8] In one influential paper that applies to our setting, 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 or Jamal Jones versus Greg Baker). More recent research corroborates this channel in the labor market (*e.g.,* Kline and Walters (2020); Jacquemet and Yannelis (2012)) and suggests this same mechanism also impacts access to housing (Diamond et al., 2019; Christensen and Timmins, 2018; Ewens et al., 2014; Hanson and Hawley, 2011) and transportation (Ge et al., 2016). We contribute to this literature by documenting whether firms' prices and bidding behavior vary across race and socioeconomic groups in an online solar PV marketplace.

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), ridesharing (Ge et al., 2016), and peer-to-peer lending (Pope and Sydnor, 2011). 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), groceries (Butters et al., 2022), vehicles (Ayres

[7]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.

[8]See Darity and Mason (1998); Altonji and Blank (1999); Rodgers (2006); Guryan and Charles (2013); Bertrand and Duflo (2016) and Neumark (2018) for relatively recent reviews of this literature.

and Siegelman, 1995), and housing (Hanson et al., 2016; Avenancio-Leon and Howard, 2019; Bayer et al., 2016, 2018; Christensen and Timmins, 2018).

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 (Mankiw, 2009; Metcalf, 2009; Grainger and Kolstad, 2010), fossil fuel extraction dynamics (Blonz et al., 2023), renewable energy policy (Reguant, 2019), and residential energy subsidies (Hahn and Metcalfe, 2021). Several previous papers have studied distributional issues in the residential solar market. For example, Nemet et al. (2017), Barbose and Darghouth (2023), and O'Shaughnessy et al. (2021) all document considerable pricing variation in solar installations, whereas Dauwalter and Harris (2023) document the distribution of environmental benefits from rooftop 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 summary statistics and descriptive analysis. We introduce our model and estimation strategy in Sections 4 and 5, and discuss the results of our model in Section 6. We conclude with conclusions and recommendations for policy in Section 7.

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 auction data we use in this research contains a set of bid prices and consumer purchase choices for solar auctions on the EnergySage 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.[9]

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

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

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.[10]

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 transaction directly with the installer or opt out and not purchase any of the offers.[11]

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 data on if 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. 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

[10] 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.

[11] 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.

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.[12] We assign the median household income to each household for the entire census block (*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.

2.3 Inferring Buyers' Race/Ethnicity From Names and Locations

One of our primary research objectives is to identify factors contributing to the ethnic or racial gap in solar PV adoption documented by Sunter et al. (2019). Absent directly measured survey questions about race and ethnicity for each household, we create a measure of race and ethnicity based on each buyer's name and the census characteristics of their block group.

In particular, we use these data to create binary measures of race and ethnicity following the two-step approach used in Diamond et al. (2019). 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.[13] 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:

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

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 ensure 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

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

[13]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.

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).[14]

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.

2.4 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.

[14] 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.

2.5 Summary Statistics

This section reports basic summary statistics regarding our sample of households and installers, in turn. Overall, the households in our sample live in areas with relatively high median income. The average household in our sample belongs to a block group with a median income of $103,000. This statistic is not surprising for a few reasons. First, our analysis focuses on the 15 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). Second, this is a sample of homeowners that have selected to shop for an expensive investment in durable capital, so we may expect them to live in relatively affluent areas. Nonetheless, the sample contains considerable variation in income.

In Appendix Table A.2, we report summary statistics separately by income quintile. For this exercise, we break our sample into income quintiles based on the classification of a household's median block-group income.[15] Median block group income ranges from $11,625 to $250,000. The average incomes within each quintile bin are $49,810, $75,390, $96,990, $121,120, and $172,750, respectively. The table shows that high-income households have relatively larger monthly electricity bills. Unsurprisingly, high-income households are more likely to report a preference for cash purchases instead of loan contracts. The non-response rate to the platform's optional onboarding questions also decreases with income.

In Appendix Table A.3, we report summary statistics across household race/ethnicity. In the bottom panel of this table, we tabulate the number of households and bids by our binary race/ethnicity measures discussed in the previous section. This sample contains data on 56,011 households and 243,120 bids. Across the households in our five race and ethnicity categories – API households account for 13.1%, Black households for 1.2%, Hispanic households for 7.2%, White households for 64.6%, and Unclassified households for the remaining 13.9%. In terms of bids, the proportions are roughly similar. API households received 14.1% of all bids, Black households received 0.8%, Hispanic households received 7.7%, White households received 63.6%, and Unclassified households collected the remaining 13.8% of all bids.

The table reveals variation in (block-group median) income across race and ethnicity. Taking White households as the reference group, API households live in areas with higher income while Black and Hispanic households live in areas with lower income. However, as shown in Appendix Figure A.1a, these distributions overlap considerably. We also see that White households have the largest monthly electricity expenditures, followed by Hispanic,

[15] We show the bounds for each quintile in first two rows of the bottom panel of Table A.2.

Black, and API households, respectively.[16] Concerning contract preference—Appendix Figure A.2 shows that White, API, and Unclassified households possess a stronger preference for purchasing, whereas Black and Hispanic households prefer loan and lease arrangements. In terms of roof age–a key determinant of the readiness for a new solar installation–63% of all households report having a roof less than 20 years old or plan to replace their roof before or coincidentally with the rooftop solar installation. This value is roughly similar for all racial and ethnic groups. About 10% report having a roof older than 20 years.[17]

2.6 Installer Locations

Finally, we summarize how proximity to solar installers varies across demographic characteristics of the installers. We display these tabulations in Panel A of Table 1 for income and Panel B for race/ethnicity.

The table first summarizes the total number of installers available to bid for each household's project. Our "All Installers" metric captures all installers located within 250 miles. Next, the table shows the number of installers within various distance bands for each group of households. Across the income groups, the total number of installers are roughly similar across all five income quintiles. For the four distance bands we consider, the mean number of installers is roughly monotonic–suggesting higher-income households are closer to more installers, on average. For example, we see 21% more installers located within 10 miles of the highest-income households relative to the lowest-income households.

The bottom panel demonstrates that Black households have fewer installers within every distance band than other race/ethnicity categories. Across our race/ethnicity categories, Hispanics have the most installers within this 250 miles at 42.57. Black households have the fewest installers with 31.74 installers within this band. White, Unclassified, and API households are in the middle with 36.11, 37.30, and 39.48 installers, respectively. We also tabulate this difference for 5, 10, 25, and 50-mile bands. Notably, for the two smallest bands–5 and 10 miles–we observe that White and API households have the most installers, whereas Black, Hispanic, and Unclassified households have fewer. At 25 miles, the distribution is quite similar for all races–except Black households. At 50 miles – it becomes even more apparent that Black households have fewer installers nearby–18.53 while all the other races are substantially higher – 25.18 for API households, 23.04 for Hispanic households, 21.25 for unclassified households, and 20.07 for White households.

[16]The box-plot we show in Figure A.1b reflects that any small deviations across the average monthly electricity spend across racial/ethnic groups are well within the bounds of statistical similarity. Since electricity rates differ, the same bill spent in one area may reflect higher or lower energy usage. We do not have detailed data on these households' electricity rates, and thus do not attempt to decode these bill amounts into monthly electricity usage.

[17]However, we note that non-response on these survey measures is higher for Black and Hispanic households and thus may be subject to selection bias.

Table 1: Number of Potential Installers By Distance

Panel A: Income Quintile

	Income Quintile				
	1 Mean (SD)	2 Mean (SD)	3 Mean (SD)	4 Mean (SD)	5 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 Mean (SD)	Black Mean (SD)	Hispanic Mean (SD)	Unclassified Mean (SD)	White 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 installers with 250 miles of the project.

In summary, lower-income households seem to be located further from installers and have a lower preference for purchasing versus financing their system. Furthermore, Black and Hispanic households are more likely to be on the lower end of the income distribution in this sample, and Black households specifically may have less access to installers. In the next section, we conduct a descriptive analysis of the market-based bidding data to analyze differences in equilibrium market outcomes across these groups directly.

3 Descriptive Analysis

In this section, we investigate heterogeneity in market outcomes across race and income, controlling for time and geography. We evaluate differences in the number of bids, prices of these bids, and adoption rate for each race and income group. We use two measures of adoption: the *close rate*, which is defined as adoption conducted within the EnergySage platform, and the *adoption rate*, which also includes households who self-report rooftop solar adoption through an off-platform source in an exit survey conducted by EnergySage.

Given that bidding and pricing behavior is a function of a general equilibrium of supply and demand, this analysis measures changes in equilibrium prices. As such, we show these results to help structure our thinking about which variables to include in our structural model but do not interpret these results as causal. We control for CBSA-by-year fixed effects to help avoid confounding these descriptive results with any fixed geographical or time-based impacts that may bias our coefficients. We explicitly model these descriptive relationships by estimating regressions of the following form:

$$Y_{icy} = \alpha + \sum_{q=1}^{q=4} \beta_d \mathbb{1}[i \in \text{Income Quintile = q}] + \sum_{r \in R} \theta_r \mathbb{1}[i \in \text{Race = r}] + \gamma_{cy} + \varepsilon_i \tag{2}$$

where Y_{icy} represents our dependent variable for buyer i in CBSA c in year y. The β coefficients represent the change in the dependent variable by income group, the θ coefficients represent the change in the dependent variable for each race/ethnicity, and the γ_{cy} coefficient represents fixed effects at the CBSA-by-year level. In these regressions, the omitted category is high-income (*i.e.* fifth income quintile) White households.

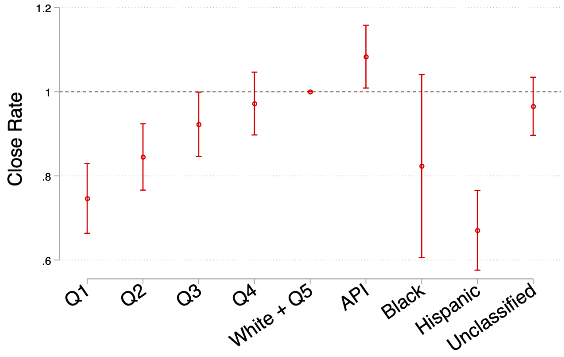
For bids and prices, we conduct this analysis using a count of the number of bids and the median price of all bids a household receives. We apply a log transformation to these dependent variables, so we interpret the β and θ coefficients as an approximate percentage change in the dependent variable. For our adoption measures—the close rate and adoption rate—we run the regressions in levels and divide the estimated coefficients by the constant parameter, α, which roughly represents the effect for the omitted group–so that these estimates can be interpreted as a percentage change relative to the reference category.[18]

We begin by documenting differences in on-platform purchase rates and survey-reported adoption rates (including offline adoption) across demographic groups. We show the results of this analysis in Figures 2a and 2b, for both the close rate and reported adoption rate, respectively.

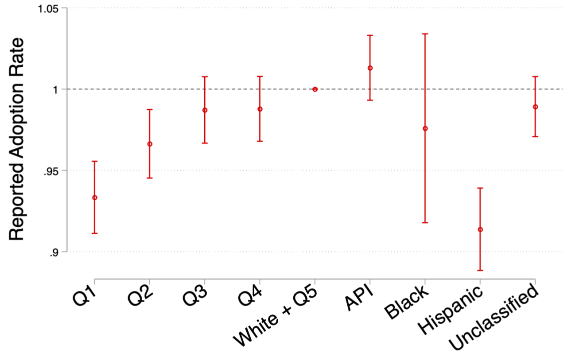
These regressions confirm that higher income is associated wither higher adoption rates–whether we analyze the platform close rate or adoption rate reported in the consumer exit survey. The lowest quintile is nearly 25% less likely to adopt via the platform and nearly 6% less likely to adopt via our exit-survey-based adoption measure, including off-platform adoption. For the second quintile, these differences are 15% and 3%, respectively. The coefficient estimates for the third and fourth quintile groups reflect lower adoption for these groups too. Regarding on-platform adoption, the third and fourth quintiles are 8% and 2%

[18]We are restricted to reporting these changes as relative percentage changes as EnergySage protects the actual close rate as a trade secret.

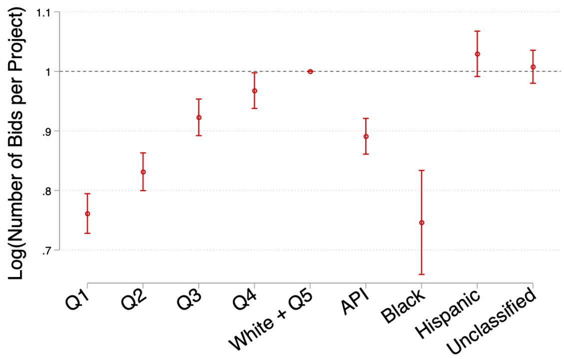
Figure 2: Descriptive Regression Estimates



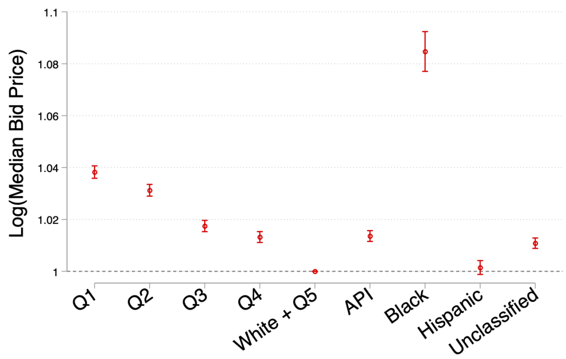
(a) Relative Close Rate by Income & Race



(b) Relative Overall Adoption by Income & Race



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



(d) Median Bid Price by Income & Race

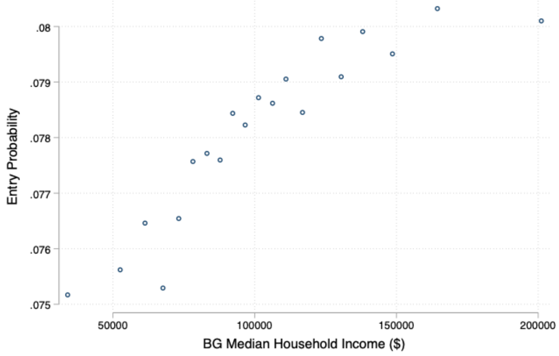
Notes: Each panel presents regression coefficients estimated from a regression of the selected variable on income and race. For example, in panel (c), the dependent variable is the logarithm of the number of bids received by a household, and the explanatory variables are income quintile and race. All coefficients are relative to white households in the 5th income quintile.

less likely to adopt, respectively–although the result for the 4th quintile is not statistically significant. Overall adoption is about 2% lower for both of these groups, but these estimates are not statistically distinguishable from zero.

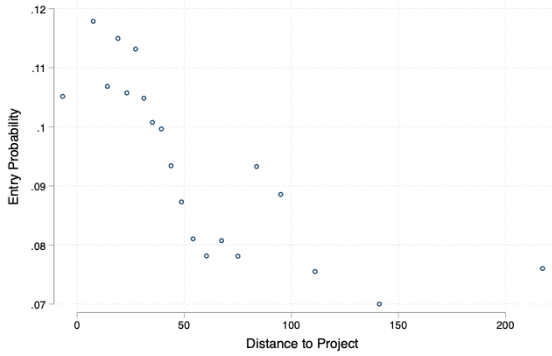
In addition, the regressions show disparities in adoption rates with respect to race and ethnicity. Compared to White households, Black households' on-platform adoption is roughly 18% lower, and overall adoption is roughly 2.5% lower, but neither result is statistically significant. API households exhibit about 8% higher adoption on-platform and about 1% higher in overall adoption than White households, although the latter of these results is not significant. Notably, Hispanic households exhibit the biggest reduction in adoption–nearly 32% lower on-platform and 7.5% lower overall. Moreover, Unclassified households are statistically indistinguishable from White households.

Having found significant differences in solar purchase rates across demographic groups, we turn to consider supply-side decisions. Notably, we investigate how the number of bids submitted to each household and the prices of those bids varies across households.

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 that 250 miles from the project.

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. Concerning race and ethnicity – we find that relative to high income (*i.e.* 5th income quintile) White households, API households receive roughly 5% fewer bids, and Black households receive roughly 9% fewer bids. Hispanic and unclassified households receive slightly more bids than high-income White households, but this effect is statistically insignificant. For income, the impacts are monotonic across quintiles. The lowest income quintile receives about 8% fewer bids, the 2nd quintile receives about 6% fewer bids, the 3rd quintile receives about 3% fewer bids and the 4th percentile receives 2% fewer bids.[19]

To better understand 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. In the previous subsection, we saw that low-income households are likelier to be farther away from potential installers. The binned scatter plot in Figure 3b indicates 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. These patterns suggest that installers' locations are a potentially important factor explaining the lower number of bids received by low-income households.

Receiving fewer bids indicates that Black and lower-income households get less attention from installers. Nonetheless, the price of these bids matters, too, since the consumer only needs to

[19] 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 about 17% fewer bids than high-income White households, on average.

contract with a single installer. If a household receives fewer bids–but these bids also come from a lower price distribution (holding quality constant)–then we may not expect receiving fewer bids to translate into lower adoption for these groups. We analyze the median price each group of households receives to capture the central tendency of the bids that *are* submitted.

In Figure 2d, we show that higher-income households receive bids with a lower price-per-watt. Notably, these coefficient estimates do not net out any economies-of-scale related to the system size. Quantitatively, these coefficients are precisely estimated but small in magnitude–the wedge between the lowest income and highest income household is about 1.5%. These point estimates are consistent with economies-of-scale associated with larger systems being quoted to richer households. Appendix Figures A.3a and A.3b 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. The scatter plots affirm 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. Larger systems tend to have lower per-unit prices because some components, such as inverters and permitting costs, are fixed. High-income households tend to have larger electricity bills and larger roofs. They, therefore, are more likely to install larger PV systems which contributes to the price disparities that we observe in the data.

Regarding race and ethnicity, Hispanic households receive roughly similar prices to high-income White households, but API, Unclassified, and Black households receive higher prices. For API and unclassified households, this effect is quite small at about 0.5%. For Black households, however, this impact is much larger–an increase of nearly 3.25% compared to high-income White households. Again, given the distribution of income and race, the price wedge for Black households is the largest, given they are more likely to be in the lower two quintiles of income.

Taken in tandem, these results suggest that low-income households receive fewer bids with higher prices than their high-income counterparts, and minority households receive fewer bids and higher prices than White households–with a larger wedge for Black households than the API, Hispanic or unclassified groups.

Lastly, we explore mechanisms explaining the bid quantity and price disparities across households' race and ethnicity. Recall that our measure of household race/ethnicity 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 2 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.4 show that neighborhoods and not names overwhelmingly explain the racial disparities in bid prices and the number of bids. Specifically, the coefficients on each variable for the racial proportion of the buyer's name are estimated as precise zeroes. In contrast, the information in the census block composition has economically significant associations with the number of bids and bid prices. We take these results as evidence that installers are primarily adjusting bidding behavior based on the locations of buyers. We do not find evidence of installers screening on buyer's names as in Bertrand and Mullainathan (2004). One caveat to these findings is that our information about buyers' names is imprecise because we only use information from the buyer's *last name* and not the first name.

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 (2022). 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, installer 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:

$$\begin{aligned} u_{ij} = B_{ij}\alpha_i + \mathbf{x}_{ij}'\beta + \mathbf{w}_j'\gamma + \delta_i + \zeta_{ig} + (1-\lambda)\varepsilon_{ij} \\ \alpha_i = \mathbf{z}_i^{(d)\prime}\tilde{\alpha}, \quad \delta_i = \mathbf{z}_i^{(s)\prime}\tilde{\delta} \end{aligned} \tag{3}$$

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 (Cardell, 1997). The nested logit model allows for more flexible substitution patterns than the standard logit model since it allows for correlation in preferences among products within pre-specified groups. 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" become 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.[20]

4.2 Supply

We model the installers' bidding decisions in a multi-attribute auction (Dorsey, 2022; 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 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 (2022), 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, 2022).[21] 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:

$$\max_{B_{ij}} \quad q_{ij}[B_{ij} - c_{ij}] \cdot \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) \tag{4}$$

[20]In practice, each installer can propose a different system capacity when bidding through the platform. In a later section, Dorsey (2022) 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.

[21]Specifically, Dorsey (2022) 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.

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. $\mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)$ 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).

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.[22]

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

$$\begin{aligned}
&\mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) = \mathbb{E}[Prob_{ij}(B_{ij}; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)] = \\
&\int Prob_{ij}(B_{ij}; \mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) \cdot dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-j} \mid \mathbf{z}_i)
\end{aligned} \tag{5}$$

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 $\mathbf{z}_i$.[23] 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

[22]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 (2022) 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.

[23]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.

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 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:

$$dG(\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}, \mathbf{W}_{-\boldsymbol{j}} \mid \mathbf{z}_i) = \prod_{l \in \mathcal{N}(\mathbf{z}_i)\setminus\{j\}} H(\mathbf{w}_l, \mathbf{z}_i) \cdot dF_{CX|\mathbf{w}_l,\mathbf{z}_i}\left(B^{*-1}(B_{il} \mid \mathbf{x}_{il}, \mathbf{w}_l, \mathbf{z}_i), \mathbf{x}_{il} \mid \mathbf{w}_l, \mathbf{z}_i\right) \tag{6}$$

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:

$$(B_{ij} - c_{ij})\frac{\partial \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i)}{\partial B_{ij}} + \mathcal{P}_{ij}(B_{ij} \mid \mathbf{x}_{ij}, \mathbf{w}_j, \mathbf{z}_i) = 0 \tag{7}$$

Given a vector of non-price characteristics, Equation 7 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 (2022) 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.

4.3 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 7 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.[24] The next

[24]Krasnokutskaya et al. (2019) and Dorsey (2022) also model the firms' entry decision; there is no guarantee of a unique equilibrium in the participation stage of the game.

section describes the estimation procedure in detail.

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 3 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 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

[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.

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 7. Notice that we cannot write a closed form for the FOC 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 identify this FOC 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 following procedure:

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:

$$Prob(Enter_{ij}) = \mathbf{z}_i'\gamma_z + \mathbf{w}_j'\gamma_w \tag{8}$$

   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:

$$Prob(\mathbf{x}_{ij}) = \mathbf{z}_i'\theta_z + \mathbf{w}_j'\theta_w. \tag{9}$$

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:

$$B_{ij} = \mathbf{z}_i'\psi_z + \mathbf{w}_j'\psi_w + \mathbf{x}_{ij}'\psi_x + \epsilon_{ij}. \tag{10}$$

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 9.
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 10 plus a residual drawn from the error distribution of the regression.[26]
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 simulated choice probabilities, and simulated demand semi-elasticities to obtain estimates for the two expectations.[27] Let s denote the simulation iteration, we define the relevant expressions as:

$$\widehat{\mathcal{P}_{ij}} = \frac{1}{S}\sum_{s=1}^{S} Prob_{ij}^s, \qquad \widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}} = \frac{1}{S}\sum_{s=1}^{S}\frac{\partial Prob_{ij}^s}{\partial B_{ij}} \tag{11}$$

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

[26]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.

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

of each project in the data:

$$\widehat{c_{ij}} = B_{ij} + \frac{\widehat{\mathcal{P}_{ij}}}{\widehat{\frac{\partial \mathcal{P}_{ij}}{\partial B_{ij}}}} \tag{12}$$

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.

6 Results

We show a selected subset of the parameter estimates pertaining to buyers' utility in Table 2. 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. Recall, 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 Asian (and Pacific Islander) 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 2: 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 that did not answer the survey question. For example, the "Roof Age: $<=$ 20 years" variable is relative to buyers that did not report the age of their roof. Standard errors are in parenthesis.

Table 3: Mean Price Elasticities Across Demographics

Income Quintile	**Mean Own-Price Elasticity**	**Avg Number of Bids**
1	-1.81	3.99
2	-1.67	4.19
3	-1.59	4.41
4	-1.58	4.52
5	-1.61	4.59

Race/Ethnicity	**Mean Own-Price Elasticity**	**Avg Number of Bids**
Asian, Pacific Islander	-1.65	4.69
Black	-1.92	2.88
Hispanic	-2.09	4.62
White	-1.6	4.28

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

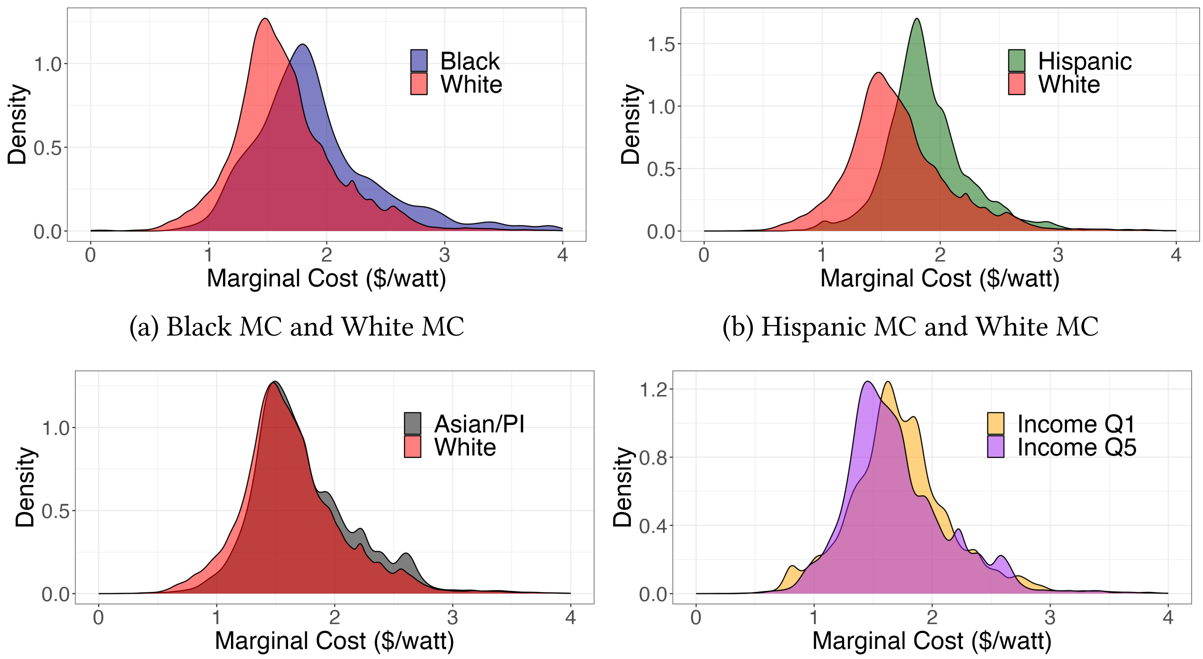
The right-hand side of Table 2 shows how the household-specific survey responses shift the utility that buyers obtain from each bid relative to the outside option. Notably, buyers with higher electricity bills obtain lower utility. Recall that utility is measured per unit of capacity and households with 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 that did not respond to that survey question (the omitted group). We do not find a statistically significant difference between buyers that report having a roof older than 20 years old relative to non-respondents. Perhaps surprisingly, we discover similar coefficients for buyers that prefer to purchase a solar installation with cash as buyers that prefer to finance their solar installation with a loan or lease. We also find similar willingness-to-pay among buyers that 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 3. We show the lowest income households are 12.5% more price elastic than the highest income households—with a mean own-price elasticity of 1.6 compared to 1.8 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 between White and minority households and the gap between low and high-income households are partly explained by differences in price sensitivity. However, Table 3 reiterates that these demographic groups receive substantially different bids from installers on average. For example, Black households obtain almost 40% fewer bids than White households (unconditionally). Therefore, households' choice sets may also partly explain the disparity in purchase rates.

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: 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.

Figure 4 compares the distribution of installation marginal costs across demographic groups. Figure (4a)-(4c) contrasts the estimated marginal costs distributions for bids made to White households, with the distribution for Black, Hispanic, and Asian 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 Table 4, 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 includ-

Table 4: 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 are 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.

ing 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 Asian 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. The coefficient estimates from the markup regressions stand in contrast with the descriptive bid price patterns that we presented in Section 3, which showed that per-unit bid prices tend to be *higher* for minority and lower-income households. Therefore, the implied marginal costs must be larger on average for Black and Hispanic households relative to similar White households. In the sixth column of Table 4, 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 Blacks than Whites, 16% higher for Hispanics than Whites, and roughly the same for APIs and Whites. 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, low-income and minority households face relatively higher solar installation costs. This indicates that underlying supply conditions contribute to the adoption disparities we observed in Section 3. There are a few reasons why marginal costs could vary systematically across household demographics. We have seen that systematic differences in labor, materials, and transportation costs across households may exist. For example, Table 1 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

Disparities in installation costs have important implications for disparities in consumer welfare and solar PV adoption. Namely, we previously show in Table 2 that Black, Hispanic, and low-income households tend to be relatively more price sensitive when shopping for solar. Furthermore, our estimates of marginal costs for these households are also higher (on average). In tandem, these results help explain the relatively low platform close rates (and survey-reported adoption rates) among these groups of households.

We use our demand estimates to evaluate disparities in expected consumer surplus obtained by the households in our sample. Specifically, we calculate the mean expected consumer surplus for a demographic group G as follows

$$\begin{aligned}\overline{CS_G} &= \frac{1}{N_G}\sum_{i\in G}\frac{\overline{q_i}}{\alpha_G}\log\Big(\sum_{j\in\mathcal{K}_i}\exp(u_{ij})\Big) \\ &= \frac{1}{N_G}\sum_{i\in G}\frac{\overline{q_i}}{\alpha_G}\log\Big(\sum_{j\in\mathcal{K}_i}\exp(B_{ij}\alpha_G+\mathbf{x}_{ij}'\beta+\mathbf{w}_j'\gamma+\delta_i)\Big)\end{aligned} \tag{13}$$

where G indicates the set of households in demographic group G, and N_G represents the number of households in that group. This formulation explicitly allows the price coefficient, α_G, to vary across demographic groups. 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.

The top panel of Table 5 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 ($824) is less than half of the surplus obtained by the highest income quintile ($1,756). Similarly, the bottom panel indicates that total consumer surplus is highest for Asian and White households ($1,567 and $1,411). In contrast, Black households obtain 72% less surplus than White households ($402) and Hispanic households obtain 56% less surplus ($615).

Appendix Table A.5 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 49% lower than the top income quintile. Moreover, the regression shows that Black households' consumer surplus is 65% lower than comparable White households and Hispanic households' surplus is 63% lower, all else equal. On the other hand, Asian households get roughly equal surplus as White households.

6.3 Decomposing Disparities in Consumer Surplus

The results in the previous subsection highlight the large disparities in consumer surplus across household demographics. However, the consumer estimates alone are not particularly informative about the mechanisms that explain these disparities. The evidence in Section 3 and 6.1 shows that low-income and minority households obtain fewer bids and face higher installation costs suggesting that supply-side factors may partly explain the income and racial

Table 5: 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	0.1	9.03	824.18
	(1.93)	(0.1)	(4.22)	(751.84)
2	4.19	0.15	8.75	1149.54
	(1.94)	(0.13)	(4.16)	(981.47)
3	4.41	0.19	8.62	1460.46
	(1.91)	(0.16)	(4.07)	(1173.61)
4	4.52	0.22	8.61	1622.55
	(1.86)	(0.17)	(4.11)	(1261.6)
5	4.59	0.22	8.86	1755.63
	(1.77)	(0.16)	(4.26)	(1292.55)

Panel B: Race/Ethnicity

Race/Ethnicity	Bids	CS/watt ($/watt)	Mean System Size (kW)	Total CS ($)
Asian, Pacific Islander	4.69	0.24	7.1	1567.36
	(1.78)	(0.17)	(3.5)	(1195.34)
Black	2.88	0.05	9.18	401.97
	(1.64)	(0.05)	(4.07)	(460.44)
Hispanic	4.62	0.08	8.38	615.06
	(1.86)	(0.07)	(3.89)	(537.28)
White	4.28	0.18	9.08	1411.47
	(1.9)	(0.15)	(4.23)	(1171.7)

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.

disparities in consumer surplus. However, the demand estimates in Table 2 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 inspired by 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:

$$\widehat{\Delta CS} = \overline{CS_A} - \overline{CS_B}. \tag{14}$$

where $\overline{CS_A}$ and $\overline{CS_B}$ represent the mean expected consumer surplus for group A households and group B households, respectively. Analogous to Figure 1, we decompose the differences $\widehat{\Delta CS}$ into three components: (1) the *demand component*—the share of the CS gap explained by differences in price sensitivity 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*. More explicitly, we expand $\widehat{\Delta CS}$ as:

$$
\begin{aligned}
\widehat{\Delta CS} = & \underbrace{[\frac{1}{N_B}\sum_{i \in B}\frac{\overline{q_i}}{\alpha_A}\log\big(\sum_{j \in \mathcal{K}_i}\exp(\alpha_A B_{ij} + \mathbf{X}'_{ij}\theta)\big) - \frac{1}{N_B}\sum_{i \in B}\frac{\overline{q_i}}{\alpha_B}\log\big(\sum_{j \in \mathcal{K}_i}\exp(\alpha_B B_{ij} + \mathbf{X}'_{ij}\theta)\big)]}_{\text{Demand Component}} \\
& + \underbrace{[\frac{1}{N_A}\sum_{i \in A}\frac{\overline{q_i}}{\alpha_B}\log\big(\sum_{j \in \mathcal{K}_i}\exp(\alpha_B B_{ij} + \mathbf{X}'_{ij}\theta)\big) - \frac{1}{N_B}\sum_{i \in B}\frac{\overline{q_i}}{\alpha_B}\log\big(\sum_{j \in \mathcal{K}_i}\exp(\alpha_B B_{ij} + \mathbf{X}'_{ij}\theta)\big)]}_{\text{Supply Component}} \\
& + \underbrace{[\overline{CS_A} - \overline{CS_B}] - \text{Demand Component} - \text{Supply Component}}_{\text{Interaction Component}}.
\end{aligned}
\tag{15}
$$

In the above expression, we write the portion of utility that household i obtains from the non-price characteristics of option j concisely as $\mathbf{X}'_{ij}\theta \equiv \mathbf{x}'_{ij}\beta + \mathbf{w}'_j\gamma + \delta_i$. In addition, $\overline{q_i}$ is the mean system capacity bid submitted to household i—this term scales utility to adjust for differences in system size across households. The first line of Equation 15 represents the portion of the CS gap explained by differences in price sensitivity between demographic group A and demographic group B. More specifically, the first term in the first line is *counterfactual* mean consumer surplus if all households in group A observed the same choice set as group B. The second term in the first line reflects the mean expected CS across all households in group B evaluated based on the observed choice set of group B. The supply component of the consumer surplus gap (second line of Equation 15) represents the difference in expected consumer surplus obtained from Group A's and Group B's observed choice sets, holding price sensitivity fixed at Group B's price sensitivity. Intuitively, the supply component encompasses the disparities created by differences in the number of bids obtained by the two groups, differences in bid prices, and differences in installer quality. The last term on the third line represents the residual portion of the consumer surplus gap which arises from changing both the price sensitivity and choice sets of the two groups simultaneously.

We first use Equation 15 to decompose the disparity in consumer surplus between high- and low-income households and then to decompose the gap between White and minority households. Specifically, the top panel of Table 6 decomposes the CS gap between households in the highest income quintile and households in the lowest income quintile. We see that the expected consumer surplus is $931 larger for the high-income group than the low-income group. We see that 37% ($340) of the consumer surplus gap is explained by low-income households being more price sensitive. Interestingly, a larger 43% share ($405) of the consumer surplus gap is explained by the supply component—differences in the choice sets that each group obtains. Recall, the highest income quintile received additional bids compared to the lowest quintile, which contributes to the observed gap in consumer surplus. The remaining 20% ($186) of the consumer surplus gap is explained by the interaction component.

In the next panel of Table 6, we decompose the gap between White and Black households. The average White households obtain $1,010 more consumer surplus than the average Black Household. Although Black households are more price sensitive than White households, our decomposition shows the difference in price sensitivity between Black and White Households accounts for only 18% ($187) of the gap in consumer surplus across these groups; thus, the price sensitivity alone cannot explain the majority of the gap. The supply component explains the majority ($568) of the White-Black consumer surplus gap. 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. Notably, the White-Black decomposition estimates are less precise due to the relatively small number of Black households in our sample.

Table 6: Decomposition of Consumer Surplus Gap by Income and Race

Panel A: High-Low Income Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	931.45	(170.57)
Demand Component ($)	340.54	(100.69)
Share Explained by Demand Component	0.37	(0.06)
Supply Component ($)	404.77	(73.64)
Share Explained by Supply Component	0.43	(0.08)
Interaction Component ($)	186.13	(50.86)
Share Explained by Interaction Component	0.2	(0.03)

Panel B: White-Black Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	1009.5	(215.75)
Demand Component ($)	186.66	(155.65)
Share Explained by Demand Component	0.18	(0.17)
Supply Component ($)	568.04	(218.9)
Share Explained by Supply Component	0.56	(0.41)
Interaction Component ($)	254.81	(213.9)
Share Explained by Interaction Component	0.25	(0.24)

Panel C: White-Hispanic Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	796.41	(138.85)
Demand Component ($)	542.99	(109.1)
Share Explained by Demand Component	0.68	(0.04)
Supply Component ($)	123.6	(27.03)
Share Explained by Supply Component	0.16	(0.04)
Interaction Component ($)	129.83	(28.93)
Share Explained by Interaction Component	0.16	(0.01)

Panel D: White-Asian/PI Consumer Surplus Decomposition

	Estimate	SE
Difference in Mean Consumer Surplus ($)	-155.89	(99.6)
Demand Component ($)	25.67	(106.75)
Share Explained by Demand Component	-0.16	(4.98)
Supply Component ($)	-178.43	(35.15)
Share Explained by Supply Component	1.14	(4.39)
Interaction Component ($)	-3.12	(13.22)
Share Explained by Interaction Component	0.02	(0.59)

Notes: The second row of each panel show 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 carry out the analogous decomposition of the gap in CS between White and Hispanic Households in Panel C of Table 6. We find a sizeable $796 gap in expected consumer surplus between White and Hispanic households on average. However, in contrast to the White-Black gap, we find differences in demand account for the majority (68%) of the White-Hispanic CS gap. Recall, that Hispanic households are more price elastic than White households, so if the two groups obtain equivalent choice sets, White households will obtain higher consumer surplus. Importantly, Hispanic and White households obtain a similar amount of bids from installers. Hence, differences in price sensitivities between White and Hispanic households explain most of the CS gap. The supply component and the interaction component each explain 16% of the consumer surplus gap.

Finally, Panel D of Table 6 shows the decomposition in the CS gap between White and Asian (and Pacific Islander) households. Asian and White Households obtain nearly the same mean CS, with Asian households obtaining $156 more surplus than White households on average. The point estimates from the demand model show that Asian households are slightly more price sensitive than White households, but the difference is small in magnitude and is not statistically significant. Moreover, because the overall gap in consumer surplus between the two groups is quite small, all of the decomposition estimates are also very noisy and statistically indistinguishable from one another.

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.[28] 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 Development, 2016). A common theme across all of these policies is

[28]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.

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).

Table 7: 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	1755.63
Comparison: Low Income (Q1)	-0.81	1.32	1755.63
Base: White	0	2.07	1411.47
Comparison: Black	-1.27	0.79	1411.47
Base: White	0	2.07	1411.47
Comparison: Hispanic	-0.77	1.3	1411.47
Base: White	0	2.07	1411.47
Comparison: Asian/PI	0.13	2.2	1411.47

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 7 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.[29] We see that if each bid submitted to a low-income household decreased by $0.81 per watt–approximately 38%—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 massive price reduction of $1.27 per watt (62%) would be necessary to eliminate the gap in consumer between White and Black households, and a $0.77 per watt (37%) decrease would be required to eliminate the gap between White and Hispanic households. Finally, prices submitted to Asian and Pacific Islander households would need to rise by $0.13 per watt to equate the expected consumer surplus with White households, since Asian households obtain slightly higher consumer surplus than White households.

The above results underscore that in the short-run relatively large subsidies or grants may be necessary to eliminate disparities in consumer surplus and solar PV adoption that we cur-

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

rently observe. For instance, our results suggest that the IRA LMI adder of 10-20% may not be sufficient to 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 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.[30]

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 showing that low-income and other disadvantaged households are heavily underrepresented 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 show that the adoption disparities are partially explained by demographic differences in willingness to pay for the technology, however, much of this gap stems from supply-side factors that materialize as higher prices and fewer bids to disadvantaged groups.

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. However, we find that these subsidies would likely need to be relatively large

[30]Importantly, in all of the calculations in Table 7 we maintain the normalization that buyers obtain zero utility from choosing the outside option.

to eliminate the consumer surplus disparities across income groups. Our decomposition results are suggestive that supply-side government policies aimed at reducing the barriers to entry for firms in underserved and disadvantaged communities may complement consumer subsidies, leading to a reduction in the gaps in adoption and consumer well-being. Similar supply-side instruments may be available to 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 that 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 and higher per-unit bid prices. However, we are not able to fully determine the extent that 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 pricing across racial and income groups presents an important opportunity for further research.

References

Aigner, Dennis J. and Glen G. Cain, “Statistical Theories of Discrimination in Labor Markets,” *Industrial and Labor Relations Review*, 1977, *30* (2), 175–187.

Altonji, Joseph G. and Rebecca M. Blank, “Race and Gender in the Labor Market,” in “Handbook of Labor Economics,” Vol. 3, Elsevier, 1999, pp. 3143–3259.

Arrow, Kenneth, “The Theory of Discrimination,” 1971.

Avenancio-Leon, Carlos and Troup Howard, “The Assessment Gap: Racial Inequalities in Property Taxation,” 2019.

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,” 2023.

Bayer, Patrick, Fernando Ferreira, and Stephen L. Ross, “The Vulnerability of Minority Homeowners in the Housing Boom and Bust,” *American Economic Journal: Economic Policy*, 2016, *8* (1), 1–27.

_ , _ , and _ , “What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders,” *The Review of Financial Studies*, 2018, *31* (1), 175–205.

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 Esther Duflo, “Field Experiments on Discrimination,” 2016.

_ 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.

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

Brouillette, Jean-Felix, Charles Jones, and Peter Klenow, "Race and Economic Well-Being in the United States," 2021.

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

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

Cardell, N. Scott, "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 1997, *13* (2), 185–213.

Christensen, Peter and Christopher Timmins, "Sorting or Steering: Experimental Evidence on the Economic Effects of Housing Discrimination," 2018.

Darity, William A. and Patrick L. Mason, "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender," *Journal of Economic Perspectives*, 1998, *12* (2), 63–90.

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, "Access to Alternatives: Increasing Rooftop Solar Adoption with Online Platforms," 2022.

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.

Ewens, Michael, Bryan Tomlin, and Liang Wang, "Statistical Discrimination or Prejudice? A Large Sample Field Experiment," *The Review of Economics and Statistics*, 2014, *96* (1), 119–134.

Ge, Yanbo, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf, "Racial and Gender Discrimination in Transportation Network Companies," 2016.

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

Grainger, Corbett A and Charles D Kolstad, "Who pays a price on carbon?," *Environmental and Resource Economics*, 2010, *46*, 359–376.

Guryan, Jonathan and Kerwin Kofi Charles, "Taste-based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots," *The Economic Journal*, 2013, *123* (572), F417–F432.

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

Hanson, Andrew and Zackary Hawley, "Do Landlords Discriminate in the Rental Housing Market? Evidence From an Internet Field Experiment in US Cities," *Journal of Urban Economics*, 2011, *70* (2), 99–114.

_ , _ , Hal Martin, and Bo Liu, "Discrimination in Mortgage Lending: Evidence from a Correspondence Experiment," *Journal of Urban Economics*, 2016.

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.

Jacquemet, Nicolas and Constantine Yannelis, "Indiscriminate discrimination: A correspondence test for ethnic homophily in the Chicago labor market," *Labour Economics*, 2012, *19* (6), 824–832.

Kline, Patrick M and Christopher R Walters, "Reasonable Doubt: Experimental Detection of Job-Level Employment Discrimination," 2020.

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," 2020.

Mankiw, N Gregory, "Smart taxes: An open invitation to join the pigou club," *Eastern Economic Journal*, 2009, *35*, 14–23.

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.

Neumark, David, "Experimental Research on Labor Market Discrimination," *Journal of Economic Literature*, 2018, *56* (3), 799–866.

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

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.

Phelps, Edmund S., "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 1972, *62* (4), 659–661.

Pope, Devin G. and Justin R. Sydnor, "What's in a Picture?: Evidence of Discrimination from Prosper.com," *Journal of Human Resources*, 2011, *46* (1), 53–92.

Reames, Tony G., "Addressing Equity with State-level Residential Solar Energy Policies | Center for Local, State, and Urban Policy," 2019.

_ , "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.

Rodgers, William M., "Handbook on the Economics of Discrimination," 2006.

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.

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

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: Means 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

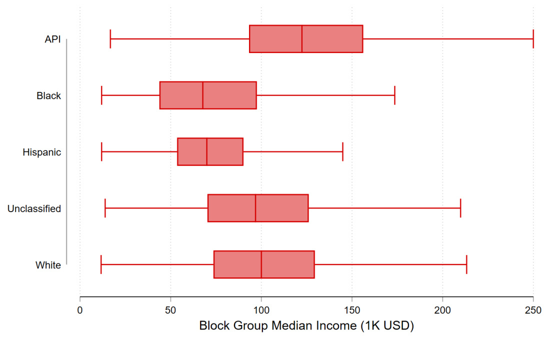
Notes: Means reported for each group with standard deviation in parentheses below.

Table A.4: Bidding Disparities: Names versus Neighborhoods

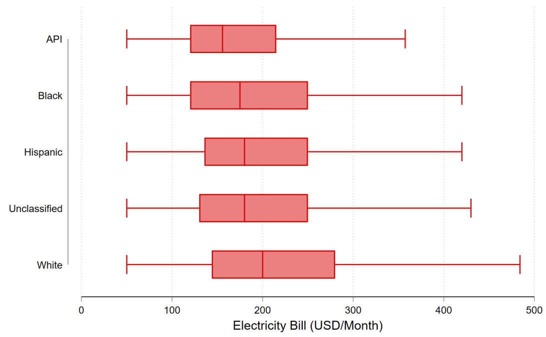
	(1) Log(Price)	(2) Log(Price)	(3) Log(Bids)	(4) Log(Bids)
Race/Ethnicity - Binary Measure				
Black	0.0319***		-0.0609***	
	(0.0040)		(0.0165)	
Asian/Pac. Islander	0.0051***		-0.0206***	
	(0.0008)		(0.0036)	
Hispanic	0.0006		0.0100**	
	(0.0009)		(0.0044)	
Unclassified	0.0041***		0.0031	
	(0.0008)		(0.0034)	
Block Group Race/Ethnicity Prop.				
Black		0.0504***		-0.0467***
		(0.0031)		(0.0127)
Asian/Pac. Islander		0.0035*		-0.0283***
		(0.0018)		(0.0082)
Hispanic		0.0010		0.0196**
		(0.0019)		(0.0090)
Name Race/Ethnicity Prop.				
Black		-0.0001***		0.0002
		(0.0000)		(0.0001)
Asian/Pac. Islander		0.0000***		-0.0001***
		(0.0000)		(0.0000)
Hispanic		-0.0000		0.0001**
		(0.0000)		(0.0000)
Income Quintiles				
1st Quintile	0.0144***	0.0123***	-0.0490***	-0.0522***
	(0.0009)	(0.0010)	(0.0041)	(0.0045)
2nd Quintile	0.0118***	0.0103***	-0.0306***	-0.0330***
	(0.0009)	(0.0009)	(0.0039)	(0.0040)
3rd Quintile	0.0066***	0.0054***	-0.0101***	-0.0116***
	(0.0008)	(0.0008)	(0.0036)	(0.0037)
4th Quintile	0.0050***	0.0041***	-0.0031	-0.0039
	(0.0008)	(0.0008)	(0.0035)	(0.0035)
Observations	243120	243120	243120	243120
R-Sq	0.579	0.581	0.428	0.428
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 2 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.

Figure A.1: Box Plots of Household Characteristics



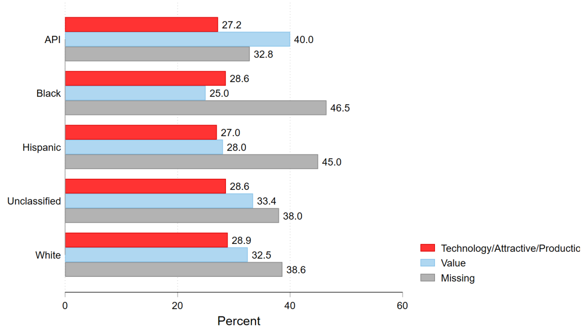
(a) Income



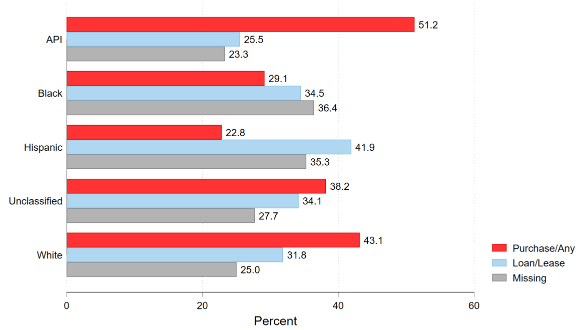
(b) Electricity Bill

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

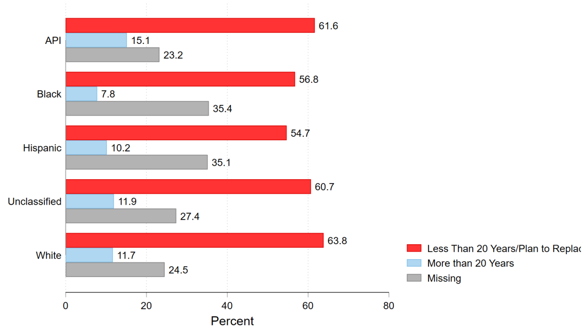
Figure A.2: Household Characteristics & Project Preferences by Race & Ethnicity



(a) Equipment Preference



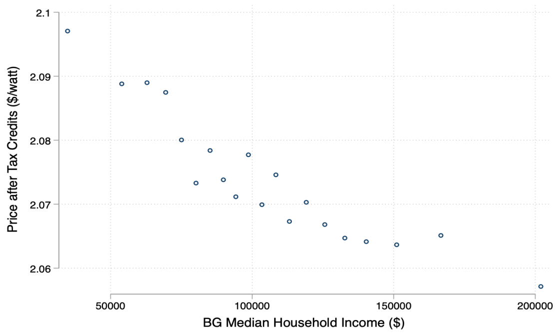
(b) Contract Preference



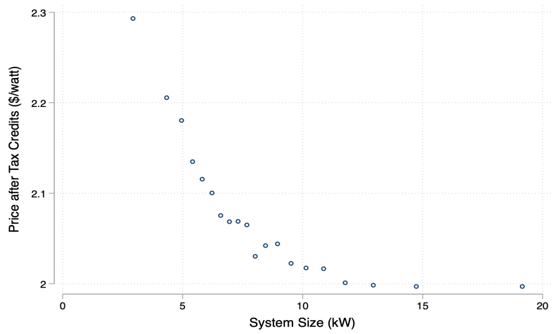
(c) Roof Age

Notes: Panels (a) and (b) compare household preference for equipment characteristics and type of contract by race category. Panel (c) shows the distribution of household's roof age by race category.

Figure A.3: 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.

Table A.5: 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 regressions include 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.