A Online Appendix

A.1 Additional Tables and Figures

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

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

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

Table A.2: Summary Statistics by Income Quintile

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

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

Table A.3: Summary Statistics by Race & Ethnicity

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

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

Table A.4: Number of Potential Installers By Distance

Panel A: Income Quintile

	Income Quintile				
	1 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 registered installers on the platform within 250 miles of the project.

Table A.5: Bidding Disparities: Names versus Neighborhoods

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

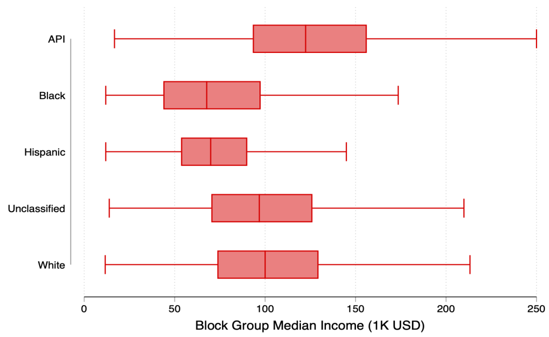
Notes: Columns 1 and 3 report the baseline regression estimates from Equation 1 using our preferred binary measure of the household race/ethnicity. Columns 2 and 4 report analogous regression results but omit the binary race/ethnicity variables as regressors and instead includes as regressors: (1) the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

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

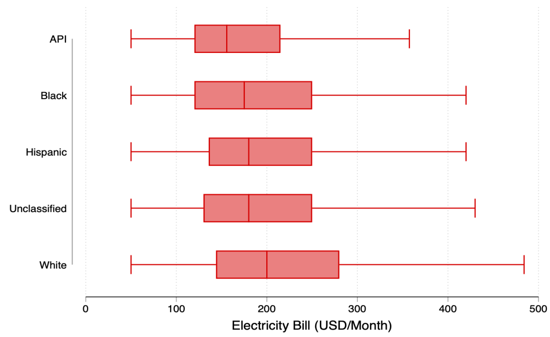
	(1) Log(Price)	(2) Log(Price)	(3) Log(Price)	(4) Log(Price)
Race/Ethnicity - Binary Measure				
Black	0.0231***	0.0169***	0.0173***	0.0120***
	(0.0032)	(0.0031)	(0.0029)	(0.0030)
Asian/Pac. Islander	0.0029***	0.0041***	-0.0036***	-0.0025***
	(0.0006)	(0.0008)	(0.0006)	(0.0007)
Hispanic	0.0017**	0.0007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

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

Figure A.1: Box Plots of Household Characteristics



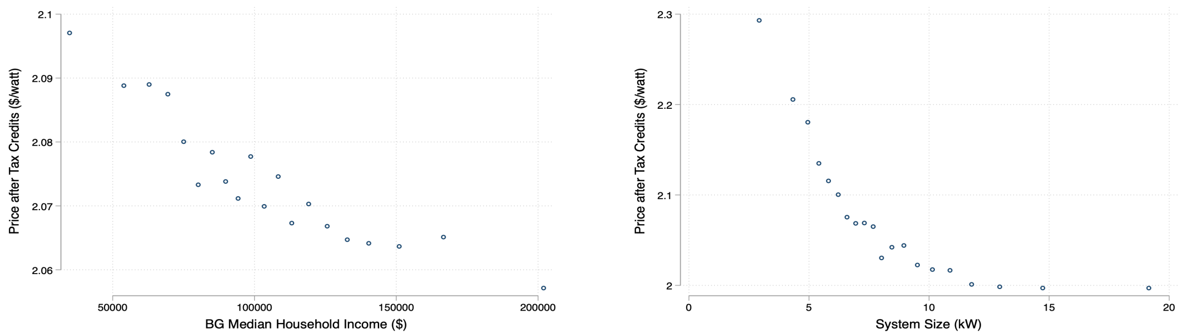
(a) Income



(b) Electricity Bill

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

Figure A.2: Household Characteristics and Prices

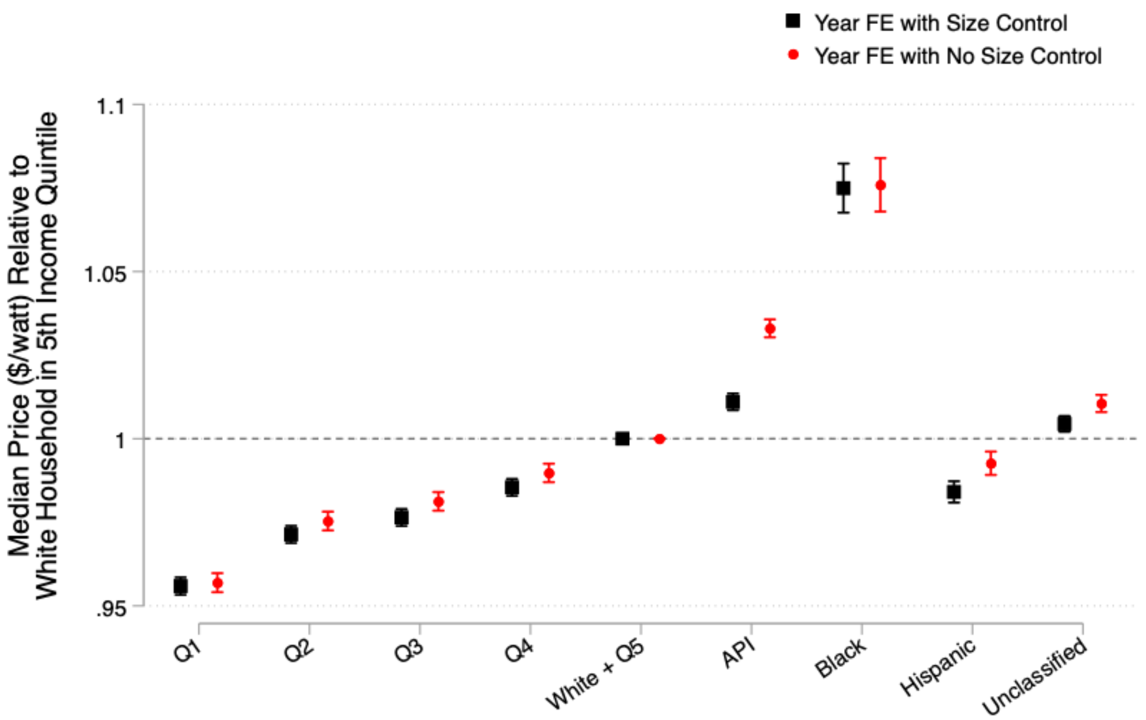


(a) Household Income

(b) System Size (kW)

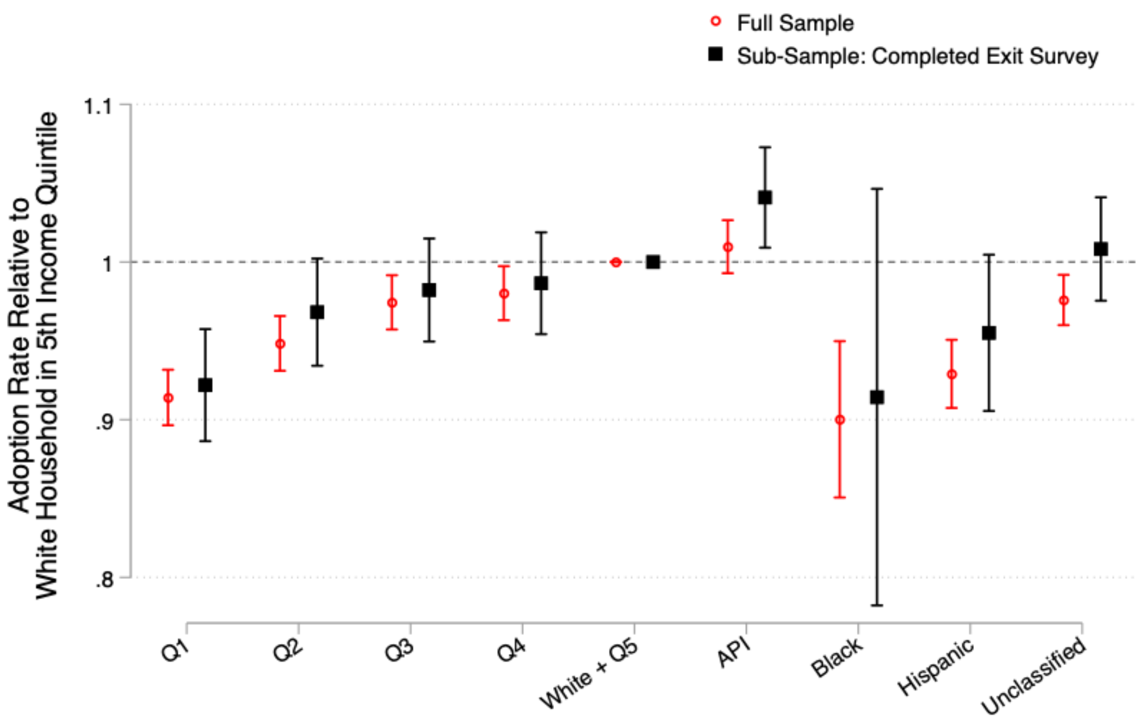
Notes: Panels (a) and (b) depict the average bid prices net of the ITC subsidy by household's block group median income and system size. Each point is calculated as the mean price among each quantile bin after controlling for CBSA fixed effects. The sample includes all bids submitted in our sample.

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



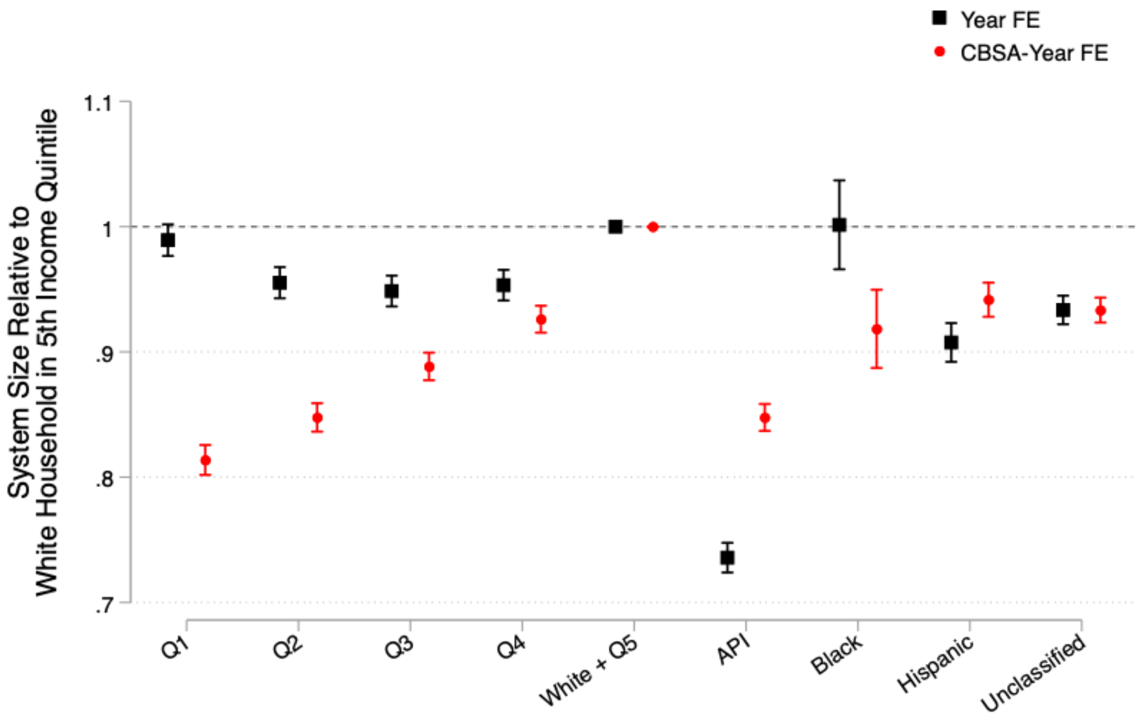
Notes: This figure compares the regression estimates from Figure 2d to an alternate specification that does not control for system size (the mean capacity bid in KW submitted to the household).

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



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

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



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

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

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

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

Table A.8: Consumer Surplus by Income and Race

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Panel A: High-Low Income Consumer Surplus Decomposition

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

Panel B: Low-High Income Consumer Surplus Decomposition

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

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

A.2 Data Appendix

A.2.1 Installer Data

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

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

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

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

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

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

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

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

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

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

A.3 Details on Installers' Bidding Problem

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

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

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

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

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

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

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

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

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

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

Recall that $Prob_{ij}$ is the probability that buyer i selects firm j's bid conditional on realized vector of competing price bids $\mathbf{B}_{i,-j}$, having a stacked vector of non-price characteristics $\mathbf{X}_{i,-j}$, and having types $\mathbf{W}_{-j}$. G represents the joint distribution function of $\mathbf{B}_{i,-j}, \mathbf{X}_{i,-j}$, and $\mathbf{W}_{-j}$ occurring in equilibrium, conditional on the project being of type $\mathbf{z}_i$.[39] Since each firm's entry draw and marginal cost draw is assumed to be *i.i.d.*, we can express dG as the product of the probabilities that each competing firm l decides to enter the auction and then bids B_{il} and has non-price characteristics $\mathbf{x}_{il}$.

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

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

[38] In practice, firms on the platform submit bids at slightly different times. Although the identities of competing bidders are not visible to auction participants, firms can see how many bids have already been submitted for a given auction. Therefore, it is possible that firms could update their expectations based on the number of bids that have already been submitted. Dorsey (2024) provides evidence that the assumption reasonably approximates firms' behavior by showing that sellers' bids do not vary systematically depending on the bid's submitted order.

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

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

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

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

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

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

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

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

A.3.1 Equilibrium

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

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

A.4 Procedure for Inferring Markups and Marginal Costs

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

References

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

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

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

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

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

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

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