An Analysis of the Effects of Residential Photovoltaic Energy Systems on Home Sales Prices in California[1]

Ben Hoen[2], Pete Cappers and Ryan Wiser
Ernest Orlando Lawrence Berkeley National Laboratory

Mark Thayer
San Diego State University

April 2011

[1] This work was supported by the Office of Energy Efficiency and Renewable Energy (Solar Energy Technologies Program) of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, by the National Renewable Energy Laboratory under Contract No. DEK-8883050, and by the Clean Energy States Alliance.

[2] Corresponding authors: Ben Hoen, LBNL, 20 Sawmill Road, Milan, NY 12571, bhoen@lbl.gov; and Ryan Wiser, LBNL, 1 Cyclotron Road, Mailstop 90R4000, Berkeley, CA 94720, rhwiser@lbl.gov.

Abstract

An increasing number of homes with existing photovoltaic (PV) energy systems have sold in the U.S., yet relatively little research exists that estimates the marginal impacts of those PV systems on the sales price. A clearer understanding of these effects might influence the decisions of homeowners, home buyers and PV home builders. This research analyzes a large dataset of California homes that sold from 2000 through mid-2009 with PV installed. Across a large number of hedonic and repeat sales model specifications and robustness tests, the analysis finds strong evidence that homes with PV systems sold for a premium over comparable homes without. The effects range, on average, from approximately \$3.9 to \$6.4 per installed watt (DC), with most models coalescing near \$5.5/watt, which corresponds to a premium of approximately \$17,000 for a 3,100 watt system. The research also shows that, as PV systems age, the premium enjoyed at the time of home sale decreases. Additionally, existing homes with PV systems are found to have commanded a larger sales price premium than new homes with similarly sized PV systems. Reasons for this discrepancy are suggested, yet further research is warranted in this area as well as a number of other areas that are highlighted.

1. Introduction

In calendar year 2010, approximately 880 megawatts (MW)[3] of grid-connected solar photovoltaic (PV) energy systems were installed in the U.S. (of which approximately 30% were residential), up from 435 MW installed in 2009, and yielding a cumulative total of 2,100 MW (SEIA & GTM, 2011). California has been and continues to be the country's largest market for PV, and is approaching 1000 MW (Ibid) or 100,000 individual PV systems installed, approximately 95% of which are residential (Barbose et al., 2010; CEC & CPUC, 2011). An increasing number of these homes have sold in California and elsewhere in the U.S., yet to date, relatively little research has been conducted estimating the existence and level of any premium to sales prices that the PV systems have likely generated.

Relatedly, one of the primary incentives for homeowners to install a PV system on their home, or for home buyers to purchase a home with a PV system already installed, is to reduce their electricity bill. However, homeowners cannot always forecast owning their home for enough time to fully recoup their PV system investment, and therefore the decision to install a PV system or purchase a home with a PV system is likely to be predicated, at least in part, on the assumption that a portion of any incremental investment in PV will be returned at the time of the home's subsequent sale. Practitioners have recognized this predication, and, in the absence of having solid research of PV premiums, have used related literature on the impact of energy efficiency investments on home prices as a proxy in making the claim that PV will increase sales prices (e.g., Black, 2010).

The basis for making the claim that an installed PV system may produce higher residential selling prices is grounded in the theory that a reduction in the carrying cost of a home will translate *ceteris paribus* into the willingness of a buyer to pay more for that home. Underlying this notion is effectively a present value calculation of a stream of savings, associated with the reduced electricity bills of PV homes, which can be capitalized into the value of the home. Along these lines, a number of studies have shown that residential selling prices are positively correlated with lower energy bills, most often attributed to energy related home improvements, such as energy efficiency investments (Johnson and Kaserman, 1983; Longstreth et al., 1984; Laquatra, 1986; Dinan and Miranowski, 1989; Horowitz and Haeri, 1990; Nevin and Watson, 1998; Nevin et al., 1999).

Those increased residential sales prices associated energy efficiency measures might be expected to apply to PV as well. Some homeowners have stated as much in surveys (CEC, 2002; McCabe and Merry, 2010), though the empirical evidence supporting such claims is limited in scope. Farhar et al. (2004a; 2008) tracked repeat sales of 15 "high performance" energy efficient homes with PV installed from one subdivision in San Diego and found evidence of higher appreciation rates, using simple averages, for these homes over comparable homes (n=12). More recently,

[3] All references to size of PV systems in this paper, unless otherwise noted, are reported in terms of direct current (DC) Watts under standard test conditions (STC). This convention was used to conform to reporting conventions outside of California. In California, most PV systems sizes are referred to using the California Energy Commission Alternating Current (CEC-AC) rating convention, which is approximately 0.83 less than DC-STC convention, but depends on a variety of factors including inverter efficiency and realistic operating efficiencies for panels. A discussion of the differences between the two conventions and how conversion can be made between them is offered in Appendix A of Barbose et al., 2010.

Dastrop et al. (2010) used a hedonic analysis to investigate the selling prices of 279 homes with PV installed in the San Diego, CA metropolitan area, finding clear evidence of PV premiums that averaged approximately 3% of the total sales price of non-PV homes, which translates into approximately $4.4 per installed PV watt (DC).

In addition to energy savings, higher selling prices might be correlated with a "cachet value" based on the "green" attributes that come bundled with energy-related improvements (e.g., helping combat global warming, avoiding increased energy price risk, impressing the neighbors etc.). A number of recent papers have investigated this correlation. Eichholtz et al. (2009, 2011) analyzed commercial green properties in the U.S, and Brounen and Kok (2010) and Griffin et al. (2009) analyzed green labeled homes in the Netherlands and Portland, Oregon, respectively, each finding premiums, which, in some cases, exceeded the energy savings (Eichholtz et al., 2009, 2011; Brounen and Kok, 2010).

Specifically related to PV, Dastrop et al. (2010) found higher premiums in communities with a greater share of Toyota Prius owners and college grads, indicating, potentially, the presence of a cachet value to the systems over and above energy savings. It is therefore reasonable to believe that buyers of PV homes might price both the energy savings and the green cachet into their purchase decisions. Relatedly, other studies have investigated whether homes with PV, which is often coupled with energy efficient features as well, sell faster than comparable homes without PV, finding evidence of increased velocity due to product differentiation (Dakin et al., 2008; SunPower, 2008; Griffin et al., 2009).

Of course there is both a buyer <u>and</u> a seller in any transaction, and the sellers of PV homes might be driven by different motivations than the buyers. Specifically, recouping the net installed cost (i.e., after available state and federal incentives) of the PV system might be a driver for sellers. In California this net cost hovered near $5/watt (DC) from 2001 through 2009 (Barbose et al., 2010). Adding slightly to the complexity, the net installed costs of PV systems vary to some degree by the type of home, with PV systems installed on *new* homes in California enjoying approximately a $1/Watt lower average installed cost than PV systems installed on *existing* homes in retrofit applications (Barbose et al., 2010). Further, sellers of *new* homes with PV (i.e., new home developers) might be reluctant to aggressively increase sale prices for installed PV systems because of the burgeoning state of the market for PV homes and concern that more aggressive pricing might slow home sales (Farhar and Coburn, 2006). Finally, as systems age, and sellers (i.e., homeowners) recoup a portion of their initial investment in energy savings (and, relatedly, a system's lifespan decreases), the need (and ability) to recoup the full investment at the time of sale might decrease. On net, it stands to reason that premiums for PV on *new* homes might be lower than those for *existing* homes, and that older systems might garner lower premiums than newer systems of the same size.

Though a link between selling prices and <u>some combination</u> of energy cost savings, green cachet, recouping the net installed cost of PV, and accounting for system age likely exists, the existing empirical literature in this area, as discussed earlier, has largely focused on either energy efficiency in residential and commercial settings, or PV in residential settings but in a limited geographic area (San Diego) with relatively small sample sizes. Therefore, to date, establishing a reliable estimate for the PV premiums that may exist across a wide market of homes has not been possible. Moreover, establishing premiums for *new* versus *existing* homes has not yet been addressed, which might be fruitful given the differences in net installed costs as well any other differences that were discussed above.

Additionally, research has not investigated whether there are increasing or decreasing returns on larger PV systems, and/or larger homes with the same sized PV systems, nor has research been conducted that investigates if older PV systems garner lower premiums *ceteris paribus*. In the case of returns to scale on larger PV systems, it is not unreasonable to expect that an increase in value for PV homes is non-linear as it relates to PV system size. For example, if larger PV systems push residents into lower electricity price tiers[4], energy bill savings could be diminished on the margin as PV system size increases. This, in turn, might translate into smaller percentage increases in selling prices as systems increase in size, and therefore a decreasing return to scale. Conversely, larger PV systems might enjoy some economies of scale in installation costs, which, in turn, might translate into lower marginal premiums at the time of home sale as systems increase in size – a decreasing return to scale. Additionally, "cachet value", to the degree that it exists, is likely to be insensitive to system size, and therefore might act as an additional driver to decreasing returns to scale. Somewhat analogously, PV premiums may be related to the number of square feet of living area in the home. Potentially, as homes increase in size, energy use can also be expected to increase, leading homeowners to be subjected to higher priced electricity rate tiers and therefore greater energy bill savings for similarly sized PV systems. Finally, as discussed previously, as PV systems age, and both a portion of the initial investment is recouped and the expected life and operating efficiency of a system decreases, premiums for the PV system might decrease.

To explore all of these possible relationships, we investigate the residential selling prices across the state of California of approximately 2,000 homes with existing PV systems against a comparable set of approximately 70,000 non-PV homes. The sample is drawn from 31 California counties, with PV home sales transaction dates of 2000 through mid-2009. We apply a variety of hedonic pricing (and repeat sales) models and sample sets, as discussed in the next section, to test and bound the possible effects of PV on residential sales prices and to increase the confidence of the findings. Using these tools, we also explore whether the effects of PV systems on home prices are impacted by whether the home is *new* or *existing*, by the size of either the PV system or the home itself, and finally by how old the PV system is. It should be stated that this research is <u>not</u> intended to disentangle the specific effects of energy savings, green cachet, recovery of the cost of installation, but rather to establish a credible estimate of the aggregate PV residential sale price effect.

The paper begins with a discussion of the data used for the analyses (Section 2). This is followed by a discussion of the empirical basis for the study (Section 3), where the variety of models and sample sets are detailed. The paper then turns to a discussion of the results and their potential implications (Section 4), and finally offers some concluding remarks with recommendations for future research (Section 5).

[4] Many California electric utilities provide service under tiered residential rates that charge progressively higher prices for energy as more of it is used.

2. Data Overview

To estimate the models described below, a dataset of California homes is used that joins five different sets of data. Those data include: (1) PV home addresses and system information from three organizations that have offered financial incentives to PV system owners in the state; (2) real estate information that can be matched to those addresses and that also include the addresses of and information on non-PV homes nearby; (3) index data that allow inflation adjustments for sale prices for conversion to 2009 dollars; (4) locational data to map the homes; and (5) elevation data to be used as a proxy for "scenic vista". Each of these is described below, as are the data processing steps employed and the resulting sample dataset.

2.1. Data Sources

The California Energy Commission (CEC), the California Public Utilities Commission (CPUC), and the Sacramento Municipal Utility District (SMUD) each provide financial incentives under different programs to encourage the installation of PV systems at residential structures, and therefore have addresses for virtually all of those systems as well as accompanying data on the PV systems. [5] Through these programs, Berkeley Laboratory was provided information on approximately 42,000 homes where PV was installed, only a fraction of which (approximately 9%) subsequently sold with the PV system in place. The data provided included: Address (street, street number, city, state and zip); incentive application, and PV system install and operational dates; PV system size; and delineations as to whether the home was *new* or *existing* at the time the system was installed (where available).

These addresses were then matched to addresses as maintained by Core Logic (CL)[6], which they aggregate from both the California county assessment and deed recorder offices. Once matched, CL provided real estate information on each of the California PV homes, as well as similar information on approximately 150,000 non-PV homes that were located in the same block group and/or subdivision as the matched PV homes. The data for both of these sets of homes included:

- address (e.g., street, street number, city, state and zip+4 code);
- most recent ("second") sale date and amount;
- previous ("first") sale date and amount (if applicable);
- home characteristics (where available) (e.g., acres, square feet of living area, bathrooms, and year built);
- assessed value;
- parcel land use (e.g., commercial, residential);
- structure type (e.g., single family residence, condominium, duplex);

[5] The CEC and CPUC have both been collecting data on PV systems installed on homes in the utility service areas of investor owned utilities (e.g., PG&E, SCE, SDG&E) for which they have provided incentives, as have some of California's publicly owned utilities (e.g. SMUD) that offer similar incentives. The CEC began administering its incentive program in 1998, and provided rebates to systems of various sizes for both residential and commercial customers. The CPUC began its program in 2001, initially focusing on commercial systems over 30 kW in size. In January 2007, however, the CEC began concentrating its efforts on new residential construction through its New Solar Home Partnership program, and the CPUC took over the administration of residential retrofit systems through the California Solar Initiative program. Separately, SMUD has operated a long-standing residential solar rebate program, but of smaller size than the efforts of the CEC and CPUC.

[6] More information about this product can be obtained from http://www.corelogic.com/. Note that Core Logic, Inc. was formerly known as First American Core Logic.

- subdivision name (if applicable)[7];
- census tract and census block group.

These data, along with the PV incentive provider data, allowed us to determine if a home sold after a PV system was installed. 3,657 such homes were identified in total, and these homes, therefore, represent the possible sample of homes on which our analysis focused. A subset of these data, for which first sale information was available and for which a PV system had not yet been installed, were culled out. These "repeat sales" were also used in the analysis, as will be discussed below.

In addition to the PV and real estate data, Berkeley Laboratory obtained from Fiserv a zip code level weighted repeat sales index of housing prices in California from 1970 through mid 2009 by quarter. These indexes, where data were available, were differentiated between low, middle, and high home price tiers, to accommodate the different appreciation/depreciation rates of market segments. Using these indexes, all sale prices were adjusted to Q1, 2009 prices. [8]

From Sammamish Data, Berkeley Laboratory purchased x/y coordinates for each zip+4 code, which allowed the mapping of addresses to street level accuracy.[9] Additionally, Berkeley Laboratory obtained from the California Natural Resources Agency (via the California Environmental Resources Evaluation System) a 30 meter Digital Elevation Map (DEM) for the state of California.[10] Combining these two sets of data, a street level elevation could be obtained for each home in the dataset, which allowed the construction of a block group relative elevation. This relative elevation served as a proxy for "scenic vista", a variable used in the analysis.

2.2. Data Processing

Data cleaning and preparation for final analysis was a multifaceted process involving selecting transactions where all of the required data fields were fully populated, determining if sales of PV homes occurred after the PV system was installed, matching the homes to the appropriate index, ensuring the populated fields were appropriately coded, and finally, eliminating outliers. Initially provided were a total of 150,000 detached single family residential sale records without PV and a total of 3,657 with PV. These totals, however, were substantially reduced (by approximately 65,000 records, 1,400 of which were PV sales) because of missing/erroneous core characteristic data (e.g., sale date, sale price, year built, square feet).[11] Additionally, the final dataset was reduced (by approximately 14,000 records, 300 of which were PV sales) because some sales occurred outside the range of the index that was provided (January 1970 to June

[7] In some cases the same subdivisions were referred to using slightly different names (e.g., "Maple Tree Estates" & "Maple Trees Estates"). Therefore an iterative process of matching based on the names, the zip code, and the census tract were used to create "common" subdivision names, which were then used in the models, as discussed later.

[8] The inflation adjustment instrument used for this analysis is the Fiserv Case-Shiller Index. This index is a weighted repeat sales index, accumulated quarterly at, optimally, the zip code level over three home price tiers (e.g., low, middle and high prices). More info can be found at: http://www.caseshiller.fiserv.com/indexes.aspx

[9] More information about this product can be obtained from http://www.sammdata.com/

[10] More information about this product can be obtained from http://www.ceres.ca.gov/

[11] Examples of "erroneous" data might include a year built or sale date that is in the future (e.g., "2109" or "Jan 1, 2015", respectively), or large groups of homes that were listed at the same price in the same year in the same block group that were thought to be "bulk" sales and therefore not valid for our purposes.

2009). Moreover, to focus our analysis on more-typical California homes and minimize the impact of outliers or potential data-entry errors on our results, observations not meeting the following criteria were screened out (see Table 1 for variable descriptions):

- the inflation adjusted most recent (second) sale price (asp2) is between $85,000 and $2,500,000;[12]
- the number of square feet (sqft) is greater than 750;
- asp2 divided by sqft is between $40 and $1,000;
- the number of acres is less than 25 and greater than sqft divided by 43,560 (where one acre equals 43,560 sqft);[13]
- the year the home was built (yrbuilt) is greater than 1900;
- the age of the home (in years) at the time of the most recent sale (ages2) is greater than or equal to negative one;
- the number of bathrooms (baths) is greater than zero and less than ten;
- the size of the PV system (size) is greater than 0.5 and less than 10 kilowatts (kW);
- each block group contained at least one PV home sale and one non-PV home sale; and
- the total assessed value (avtotal), as reported by the county via Core Logic, is less than or equal to the predicted assessed value (pav), where pav = sp2*1.02^(2010-year of sale).[14]

In addition, the repeat sales used in the analysis had to meet the following criteria:

- the difference in sale dates (sddif) between the most recent (second) sale date (sd2) and the previous (first) sale date (sd1) is less than 20 years;
- PV is not installed on the home as of sd1; and
- the adjusted annual appreciation rate (adjaar) is between -0.14 and 0.3 (where adjaar = ln(asp2/asp1)/(sddif/365), which corresponds to the 5th and 95th percentile for the distribution of adjaar.[15]

[12] An alternative screen was tested that limited the data to homes under 1 million (leaving 90% of the data) and $600,000 (leaving 75%) with no change to the results.

[13] An alternative screen which incorporated the number of stories for the home with the numbers of square feet in calculating the "footprint", and therefore, allowed smaller parcels to be used, was also explored, with no change in results.

[14] This screen was intended to help ensure that homes that had improvements since the most recent sale, which would be reflected in a higher assessed value than would otherwise be the maximum allowable, were removed from the dataset. This screen was not applied to homes that sold in 2009, however, because, in those cases, assessed values often had not been updated to reflect the most recent sale.

[15] This final screen was intended to remove homes that had unusually large appreciation or deprecations between sales, after adjusting for inflation, which would indicate that underlying home characteristics between the two sales likely changed (e.g., an addition was added, the condition of the home dramatically worsened, etc.)

Table 1: Variable Descriptions

Variable	Description
acre	size of the parcel (in acres)
acregt1	number of acres more than one
acrelt1	number of acres less than one
adjaar	adjusted annual appreciation rate
ages2	age of home as of sd2
ages2sqr	ages2 squared
asp1	inflation adjusted sp1 (in 2009 dollars)
asp2	inflation adjusted sp2 (in 2009 dollars)
avtotal	total assessed value of the home
bath	number of bathrooms
bgre_100	relative elevation to other homes in block group (in 100s of feet)
elev	elevation of home (in feet)
lasp1	natural log of asp1
lasp2	natural log of asp2
pav	predicted assessed value
pvage	age of the PV system at the time of sale
sd1	first sale date
sd2	second sale date
sddif	number of days separating sd1 and sd2
size	size (in STC DC kW) of the PV system
sp1	first sale price (not adjusted for inflation)
sp2	second sale price (not adjusted for inflation)
sqft	size of living area
sqft_1000	size of living area (in 1000s of square feet)
yrbuilt	year the home was built

2.3. Data Summary

The final full dataset includes a total of 72,319 recent sales, 1,894 of which are PV homes and 70,425 of which are non-PV (see Table 2). The homes with PV systems are distributed evenly between *new* (51%) and *existing* (49%) home types, while the non-PV homes are weighted toward *existing* homes (62%) over *new* (38%) (see Table 5). The final repeat sales dataset of homes selling twice total 28,313 homes, of which 394 are PV and 27,919 are non-PV (see Table 3).

As indicated in Table 2, the average non-PV home in the full sample (not the repeat sales sample) sold for $584,740 (unadjusted) in late 2005, which corresponds to $480,862 (adjusted) in 2009 dollars.[16] This "average" home is built in 1986, is 19 years old at the time of sale, has 2,200 square feet of living space, has 2.6 bathrooms, is situated on a parcel of 0.3 acres, and is located at the mean elevation of the other homes in the block group. On the other hand, the average PV home in the full sample sold for $660,222 in early 2007, which corresponds to $537,442 in 2009 dollars. Therefore this "average" PV home, as compared to the "average" non-PV home, is higher in value. This difference might be explained, in part, because it is slightly younger at the time of sale (by two years), slightly bigger (by 200 square feet), has more

[16] The adjusted values, which are based on a housing price index, demonstrate the large-scale price collapse in the California housing market post 2005; that is, there has been significant housing price depreciation.

bathrooms (by 0.3), is located on a parcel that was slightly larger (by 0.06 acres), and, of course, has a PV system (which was 3,100 watts and was 1.5 years old).

The repeat sale dataset, as shown in Table 3, shows similar modest disparities between PV and non-PV homes, with the "average" PV homes selling for more (in 2009 $) in both the first and second sales. Potentially more tellingly though, non-PV homes show a slight depreciation (of -1.4%) between sales after adjusting for inflation, while PV homes show a modest appreciation (of 3.2%), yet PV homes are slightly bigger (by 100 square feet), and occupy a slightly larger parcel (by 0.2 acres), conversely, are older (by 10 years), and of course have a PV system (which is 4,030 watts and is 2.5 years old).[17]

Focusing on the full dataset geographically (see Table 4 and Figure 1), we find that it spans 31 counties with the total numbers of PV and non-PV sales ranging from as few as nine (Humboldt) to as many as 11,991 (Placer). The dataset spans 835 separate block groups (not shown in the table), though only 162 (18.7%) of these block groups contain subdivisions with at least one PV sale. Within the block groups that contain subdivisions with PV sales there are 497 subdivision-specific delineations. As shown in Table 5, the data on home sales are fairly evenly split between *new* and *existing* home types, are located largely among four utility service areas, with the largest concentration in PG&E, and occurred over eleven years, with the largest concentration of PV sales occurring in 2007 and 2008.

In summary, both the full dataset and the repeat sale dataset show higher values and positive appreciation between sales, respectively, for PV homes as compared to non-PV homes. This seems to indicate a premium exists, yet, without taking into account the differences between these homes (e.g., square feet), their neighborhoods, and the market conditions surrounding the sales, a determination as to whether a premium actually exists is imprudent. The models, that will be discussed next, are designed to do just that.

[17] This disparity, as is discussed later, is accounted for, to some degree, by one of the robustness tests to the DD model, which limits the maximum number of years between the second and first sale to five years.

Table 2: Summary Statistics of Full Dataset

	Non-PV Homes				
Variable	***n***	**Mean**	**Std. Dev.**	**Min**	**Max**
acre	70425	0.3	0.8	0.0	24.8
acregt1	70425	0.1	0.7	0.0	23.8
acrelt1	70425	0.2	0.2	0.0	1.0
ages2	70425	19	23.3	-1	108
ages2sqr	70425	943	1681	0	11881
asp2	70425	$ 480,862	$ 348,530	$ 85,007	$ 2,498,106
avtotal	70425	$ 497,513	$ 359,567	$ 10,601	$ 3,876,000
bath	70425	2.6	0.9	1	9
bgre_100	70425	0.0	1.2	-18.0	19.0
elev	70425	424	598	0	5961
lasp2	70425	12.9	0.6	11.4	14.7
pvage	70425	0	0	0	0
sd2	70425	9/30/2005	793 days	1/7/1999	6/30/2009
size	70425	0	0	0	0
sp2	70425	$ 584,740	$ 369,116	$ 69,000	$ 4,600,000
sqft_1000	70425	2.2	0.9	0.8	9.3
yrbuilt	70425	1986	23	1901	2009

	PV Homes				
Variable	***n***	**Mean**	**Std. Dev.**	**Min**	**Max**
acre	1894	0.4	1.0	0.0	21.6
acregt1	1894	0.1	0.9	0.0	20.6
acrelt1	1894	0.2	0.2	0.0	1.0
ages2	1894	17.3	24.5	-1	104
ages2sqr	1894	937	1849	0	11025
asp2	1894	$ 537,442	$ 387,023	$ 85,973	$ 2,419,214
avtotal	1894	$ 552,052	$ 414,574	$ 23,460	$ 3,433,320
bath	1894	2.9	1	1	7
bgre_100	1894	0.2	1.3	-10.0	17.9
elev	1894	414	584	0	5183
lasp2	1894	13.0	0.6	11.4	14.7
pvage	1894	1.5	2.0	-1.0	9.0
sd2	1894	3/28/2007	622 days	8/1/2000	6/29/2009
size	1894	3.1	1.6	0.6	10.0
sp2	1894	$ 660,222	$ 435,217	$ 100,000	$ 3,300,000
sqft_1000	1894	2.4	0.9	0.8	11.0
yrbuilt	1894	1989	25	1904	2009

Table 3: Summary Statistics of Repeat Sale Dataset

	Non-PV Homes				
Variable	***n***	**Mean**	**Std. Dev.**	**Min**	**Max**
acre	27919	0.3	0.7	0.0	23.2
acregt1	27919	0.1	0.6	0.0	22.2
acrelt1	27919	0.2	0.2	0.0	1.0
ages2	27919	23.6	22.7	0	108
ages2sqr	27919	1122.0	1775.0	1.0	11881.0
asp1	27919	$ 488,127	$ 355,212	$ 85,398	$2,495,044
asp2	27919	$ 481,183	$ 347,762	$ 85,007	$2,472,668
avtotal	27919	$ 498,978	$ 360,673	$ 35,804	$3,788,511
bath	27919	2.5	0.8	1	9
bgre_100	27919	0.0	1.3	-17.7	19.0
elev	27919	426	588	0	5961
lasp1	27919	12.9	0.6	11.4	14.7
lasp2	27919	12.9	0.6	11.4	14.7
pvage	27919	0	0	0	0
sd1	27919	5/5/2001	1780 days	11/1/1984	12/11/2008
sd2	27919	5/14/2006	786 days	3/11/1999	6/30/2009
sddif	27919	1835	1509	181	7288
size	27919	0	0	0	0
sp1	27919	$ 444,431	$ 287,901	$ 26,500	$2,649,000
sp2	27919	$ 577,843	$ 371,157	$ 69,000	$3,500,000
sqft_1000	27919	2.1	0.8	0.8	7.7
yrbuilt	27919	1982	23	1901	2008

	PV Homes				
Variable	***n***	**Mean**	**Std. Dev.**	**Min**	**Max**
acre	394	0.5	1.4	0.0	21.6
acregt1	394	0.2	1.3	0.0	20.6
acrelt1	394	0.2	0.2	0.0	1.0
ages2	394	34.6	25.6	1	104
ages2sqr	394	1918.0	2336.0	4.0	11025.0
asp1	394	$ 645,873	$ 417,639	$ 110,106	$2,339,804
asp2	394	$ 666,416	$ 438,544	$ 91,446	$2,416,498
avtotal	394	$ 682,459	$ 478,768	$ 51,737	$3,433,320
bath	394	2.6	0.9	1	7
bgre_100	394	0.1	1.6	-5.5	17.9
elev	394	479	581	3	3687
lasp1	394	13.2	0.6	11.6	14.7
lasp2	394	13.2	0.6	11.4	14.7
pvage	394	2.5	1.6	-1.0	9.0
sd1	394	11/22/1999	1792 days	11/30/1984	1/7/2008
sd2	394	1/9/2007	672 days	8/1/2000	6/29/2009
sddif	394	2605	1686	387	7280
size	394	4.03	1.94	0.89	10
sp1	394	$ 492,368	$ 351,817	$ 81,500	$2,500,000
sp2	394	$ 800,359	$ 489,032	$ 121,000	$3,300,000
sqft_1000	394	2.2	0.8	0.8	5.3
yrbuilt	394	1972	26	1904	2008

Table 4: Frequency Summary by California County

CA County	Non-PV	PV	Total
Alameda	4,826	153	**4,979**
Butte	457	12	**469**
Contra Costa	5,882	138	**6,020**
El Dorado	938	85	**1,023**
Humboldt	7	2	**9**
Kern	2,498	53	**2,551**
Kings	134	5	**139**
Los Angeles	3,368	82	**3,450**
Marin	1,911	61	**1,972**
Merced	48	2	**50**
Monterey	10	2	**12**
Napa	36	1	**37**
Orange	1,581	44	**1,625**
Placer	11,832	159	**11,991**
Riverside	4,262	87	**4,349**
Sacramento	10,928	483	**11,411**
San Bernardino	2,138	50	**2,188**
San Diego	1,083	30	**1,113**
San Francisco	407	16	**423**
San Joaquin	1,807	20	**1,827**
San Luis Obispo	232	1	**233**
San Mateo	2,647	92	**2,739**
Santa Barbara	224	7	**231**
Santa Clara	6,127	157	**6,284**
Santa Cruz	90	1	**91**
Solano	2,413	39	**2,452**
Sonoma	1,246	32	**1,278**
Tulare	774	14	**788**
Ventura	1,643	42	**1,685**
Yolo	16	1	**17**
Yuba	860	23	**883**
Total	**70,425**	**1,894**	**72,319**

Figure 1: Map of Frequencies of PV Homes by California County

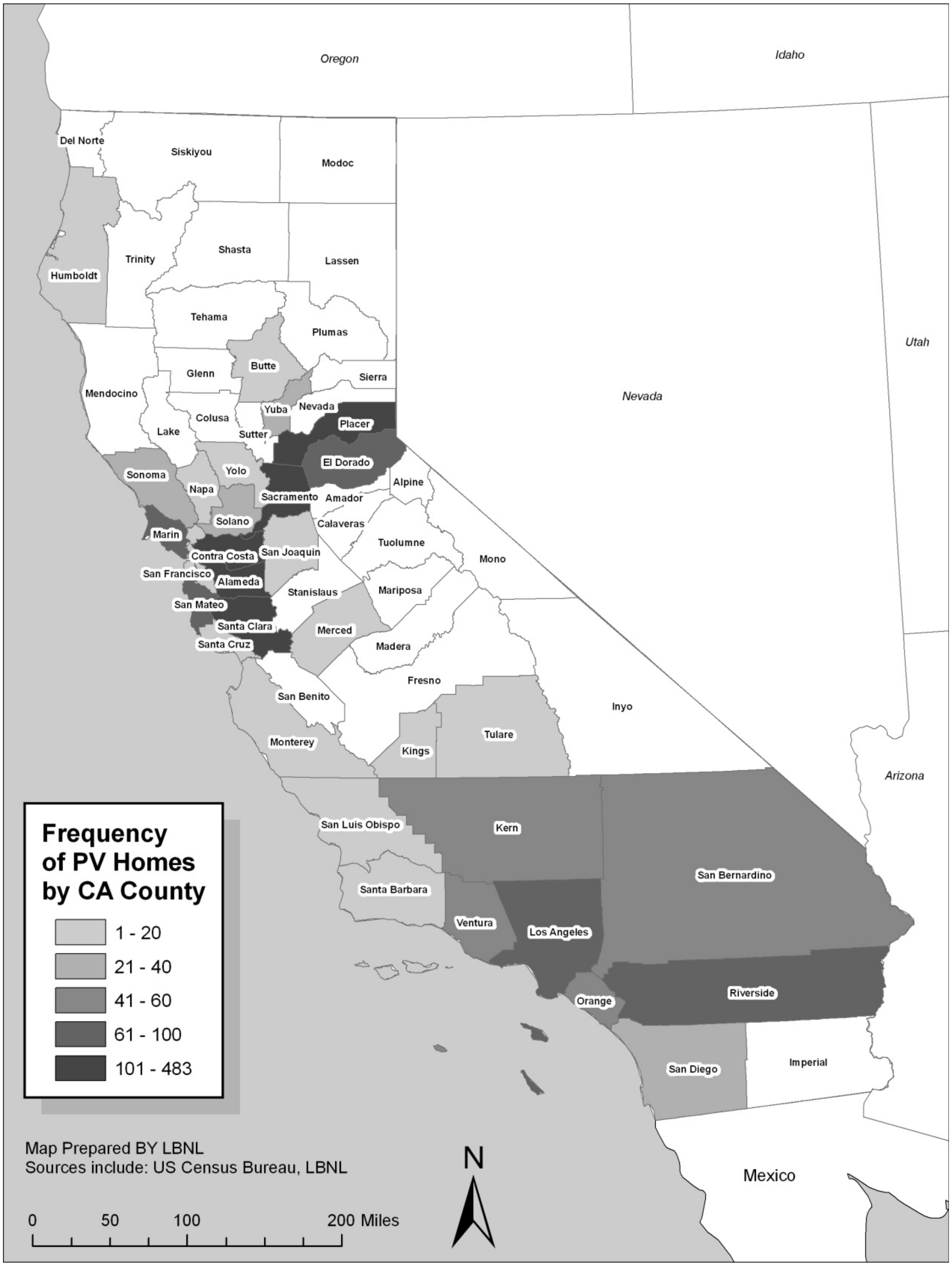


Table 5: Frequency Summary by Home Type, Utility and Sale Year

Home Type *	**Non-PV**	**PV**	**Total**
New Home	26,938	935	27,873
Existing Home	43,487	897	44,384
Utility **	**Non-PV**	**PV**	**Total**
Pacific Gas & Electric (PG&E)	36,137	1,019	37,156
Southern California Edison (SCE)	14,502	337	14,839
San Diego Gas & Electric (SDG&E)	8,191	35	8,226
Sacramento Municipal Utility District (SMUD)	11,393	498	11,891
Other	202	5	207
Sale Year	**Non-PV**	**PV**	**Total**
1999	110	0	110
2000	379	1	380
2001	1,335	10	1,345
2002	6,278	37	6,315
2003	8,783	63	8,846
2004	10,888	153	11,041
2005	10,678	168	10,846
2006	9,072	173	9,245
2007	8,794	472	9,266
2008	9,490	642	10,132
2009	4,618	175	4,793

** A portion of the PV homes could not be classified as either new or existing and therefore are not included in these totals*

*** Non-PV utility frequencies were estimated by mapping block groups to utility service areas, and then attributing the utility to all homes that were located in the block group*

3. Methods and Statistical Models

3.1. Methodological Overview

The data, as outlined above, not only show increased values and appreciation for PV homes (in 2009 $), but also important differences between PV and non-PV homes as regards home, site, neighborhood and market characteristics that could, potentially, be driving these differences in value. A total of 18 empirical models, with a high reliance on the hedonic model, are used in this paper to disentangle these potentially competing influences to determine if PV homes sell for a premium *ceteris paribus*. The basic theory behind the hedonic model starts with the concept that a house can be thought of as a bundle of characteristics. When a price is agreed upon between a buyer and seller there is an implicit understanding that those characteristics have value. When data from a number of sales transactions are available, the average individual marginal contribution to the sales price of each characteristic can be estimated with a hedonic regression model (Rosen, 1974; Freeman, 1979). This relationship takes the basic form:

Sales price = *f* (home and site, neighborhood, and market characteristics)

"Home and site characteristics" might include, but are not limited to, the number of square feet of living area, the size of the parcel of land, and the presence of a PV system. "Neighborhood" characteristics effects such as the crime rate, school district and distance to the central business district. Finally, "market [i.e. temporal] characteristics" might include, but are not limited to, temporal effects such as housing market inflation/deflation.

A variant of the hedonic model, a repeat sales model, as discussed briefly before, holds constant many of the characteristics discussed above, and compares inflation adjusted selling prices of homes that have sold twice, both before a condition exists (e.g., before a PV system is installed on the home) and after the condition exists (e.g., after a PV system is installed on the home), and across PV and non-PV homes. This "repeat sales" model, in the form used in this paper, is referred to as a difference-in-difference (DD) model, and is discussed in more detail later.

To test for the impact of PV systems on residential selling prices, a series of "base" hedonic models, a "base" difference-in-difference model, a series of robustness models, and two "other" models are estimated for this research.[18] As discussed later, these models are used to test for fixed (whether the home has a PV systems) and continuous (the size of the PV system) effects using the full dataset of PV homes. As well, they are also used to test for any differences that exist between new and existing PV homes, homes with PV systems of different ages, and to test for the possibility of non-linear returns to scale based on the size of the PV system or the home itself. Before describing these models in more detail, however, a summary of the variables to be included in the models is provided.

3.2. Variables Used in Models

In each base model, be it hedonic or difference-in-difference, four similar sets of parameters are estimated, namely coefficients on the variables of interest and coefficients for three sets of

[18] As will be discussed later, each of the "base" models is coupled with a set of two or three robustness models. The "other" models are presented without "robustness" models.

controls that include home and site characteristics, neighborhood (census block group) fixed effects, and temporal (year and quarter) fixed effects. The variables of interest are the focus of the research, and include such variables as whether the home has a PV system installed or not, the size of the PV system, and interactions between these two variables and others, such as the size of the home or the age of the PV system. To accurately measure these variables of interest (and their interactions), other potentially confounding variables need to be controlled for in the models. The base models differ in their specification and testing of the variables of interest, as discussed later, but use the same three sets of controls.

The first of these sets of control variables accounts for differences across the dataset in home and site-specific characteristics, including the age of the home (linear and squared), the total square feet of living area, and the relative elevation of the home (in feet) to other homes in the block group, which serves as a proxy for "scenic vista" a value influencing characteristic (see e.g., Hoen et al., 2009).[19] Additionally, the size of the property in acres was entered into the model in a spline form to account for the different valuations of less than one acre and greater than one acre.

The second set of controls, the geographic fixed effects variables, includes dummy variables that control for aggregated "neighborhood" influences, which, in our case, are census block groups.[20] A census block group generally contains between 200 and 1,000 households[21], and is delineated to never cross boundaries of states, counties, or census tracts, and therefore, in our analysis, serves as a proxy for "neighborhood". To be usable, each block group had to contain at least one PV home and one non-PV home. The estimated coefficients for this group of variables capture the combined effects of school districts, tax rates, crime, distance to central business district and other block group specific characteristics. This approach greatly simplifies the estimation of the model relative to determining these characteristics for each home, but interpreting the resulting coefficients can be difficult because of the myriad of influences captured by the variables. Because block groups are fairly small geographically, spatial autocorrelation[22] is also, to some degree, dealt with through the inclusion of these variables.

Finally, the third set of controls, the temporal fixed effect variables, includes dummy variables for each quarter of the study period to control for any inaccuracies in the housing inflation adjustment that was used. The housing inflation adjustment, as will be discussed in more detail

[19] Other home and site characteristics were also tested, such as the condition of the home, the number of bathrooms, the number of fireplaces, and if the home had a garage and/or a pool. Because these home and site characteristics were not available for all home transactions (thus reduced the sample of homes available), did not add explanatory power to the model, nor did they affect the results substantively, they were not included.

[20] For a portion of the dataset, a common subdivision name was identified, which, arguably, serves as a better proxy for neighborhood than block group. Unfortunately, not all homes fell within a subdivision. Nonetheless, a separate combined subdivision-block group fixed effect was tested and will be discussed later.

[21] Block groups generally contain between 600 and 3,000 people, and the median household size in California is roughly 3.

[22] Spatial Autocorrelation - a correlation between neighbor's selling prices - can produce unstable estimates yielding unreliable significance tests in hedonic models if not accounted for. One reason for this autocorrelation is omitted variables, such as neighborhood characteristics (e.g., distance to the central business district), which affect all properties from the same area similarly. Having micro-spatial controls, such as block groups or subdivisions, help control for autocorrelation.

below later, adjusts the sales prices throughout the study period to 2009 prices at a zip code level across as many as three price tiers. Although this adjustment is expected to greatly improve the model - relative to using *just* a temporal fixed effect with an unadjusted price - it is also assumed that because of the volatility of the housing market, the index may not capture price changes perfectly and therefore the model is enhanced with the additional inclusion of these quarterly controls.[23]

3.3. Fixed and Continuous Effect Hedonic Models

The analysis begins with the most basic model comparing prices of all of the PV homes in the sample (whether new or existing) to non-PV homes across the full dataset. As is common in the literature (Malpezzi, 2003; Sirmans et al., 2005b; Simons and Saginor, 2006), a semi-log functional form of the hedonic model is used where the dependent variable, the (natural log of) sales price (P), is measured in zip code-specific inflation-adjusted (2009) dollars. To determine if an average-sized PV system has an effect on the sale price of PV homes (i.e., a fixed effect) we estimate the following base fixed effect model:

$$\ln(\mathrm{P_{itk}}) = \alpha + \beta_1(\mathrm{T_t}) + \beta_2(\mathrm{N_k}) + \sum_{\mathrm{a}} \beta_3(\mathrm{X_i}) + \beta_4(\mathrm{PV_i}) + \varepsilon_{\mathrm{itk}} \tag{1}$$

where
P_{itk} represents the inflation adjusted sale price for transaction *i*, in quarter *t*, in block group *k*,
α is the constant or intercept across the full sample,
T_t is the quarter in which transaction *i* occurred,
N_k is the block group in which transaction *i* occurred,
X_i is a vector of *a* home characteristics for transaction *i* (e.g., acres, square feet, age, etc.),
PV_i is a fixed effect variable indicating if a PV system is installed on the home in transaction *i*,
β_1 is a parameter estimate for the quarter in which transaction *i* occurred,
β_2 is a parameter estimate for the block group in which transaction *i* occurred,
β_3 is a vector of parameter estimates for home characteristics *a*,
β_4 is a parameter estimate for the PV fixed effects variable, and
ε_{itk} is a random disturbance term for transaction *i*.

The parameter estimate of primary interest in this model is β_4 which represents the marginal percentage change in sale price with the addition of an average sized PV system. If differences in selling prices exist between PV and non-PV homes, we would expect the coefficient to be positive and statistically significant.

An alternative to equation (1) is to interact the PV fixed effect variable (PV_i) with the size (in kW) of the PV system as installed on the home at the time of sale ($SIZE_i$) therefore producing an estimate for the differences in sales prices as a function of size of the PV system. This base continuous effect model takes the form:

$$\ln(\mathrm{P_{itk}}) = \alpha + \beta_1(\mathrm{T_t}) + \beta_2(\mathrm{N_k}) + \sum_{\mathrm{a}} \beta_3(\mathrm{X_i}) + \beta_4(\mathrm{PV_i} \cdot \mathrm{SIZE_i}) + \varepsilon_{\mathrm{itk}} \tag{2}$$

[23] A number of models were tested both with and without these temporal controls and with a variety of different temporal controls (e.g., monthly) and temporal/spatial controls (e.g., quarter and tract interactions). The quarterly dummy variables were the most parsimonious, and none of the other approaches impacted the results substantively.

where $SIZE_i$ is a continuous variable for the size (in kW) of the PV system installed on the home prior to transaction i, β_4 is a parameter estimate for the percentage change in sale price for each additional kW added to a PV system, and all other terms are as were defined for equation (1). If differences in selling prices exist between PV and non-PV homes, we would expect the coefficient to be positive and statistically significant, indicating that for each additional kilowatt added to the PV system the sale price increases by β_4 (in % terms).

This "continuous effect" specification may be preferable to the PV "fixed effect" model because one would expect that the impact of PV systems on residential selling prices would be based, at least partially, on the size of the system, as size is related to energy bill savings.[24] Moreover, this specification allows for a direct estimate of any PV home sales premium in dollars per watt ($/watt), which is the form in which other estimates – namely average net installed costs – are reported. With the previous fixed effects specification, a $/watt estimate can still be derived, but not directly. Therefore, where possible in this paper, greater emphasis is placed on the continuous effect specification in place of the fixed effect estimation.

As mentioned earlier, for each base model we explore a number of different robustness models to better understand if and to what degree the results are unbiased. In the present research, two areas of bias are of particular concern: omitted variable bias and sample selection bias.

The omitted variables that are of specific concern are any that might be correlated with the presence of PV, and that might affect sales prices. An example is energy efficiency (EE) improvements, which might be installed contemporaneously with a PV energy system. If many homes with PV have EE improvements, whereas the comparable non-PV homes do not, then estimates for the effects of PV on selling prices might be inclusive of EE effects and, therefore, may be inappropriately high. Any other value-influencing home improvements (e.g., kitchen remodels), if correlated with the presence of PV, could similarly bias the results if not carefully addressed.

With respect to selection bias, the concern is that the distribution of homes that have installed PV may be different from the broad sample of homes on which PV is not installed. If both sets of homes are assumed to have similar distributions but are, in point of fact, dissimilar due to selection, then the estimates for the effects of PV on the selling price could be inclusive of these underlying differences but attributed to the existence of PV, thereby also potentially biasing the results.

To address the issue of omitted variable bias, one robustness model uses the same data sample as the base model but a different model specification. Specifically, a combined subdivision-block group fixed effect variable can be substituted, where available, in place of the block group fixed effect variable as an alternative proxy for "neighborhood." Potentially omitted variables are

[24] Ideally, the energy bill savings associated with individual PV systems could be entered into the model directly, but these data were not available. Moreover, estimating the savings accurately on a system-by-system basis was not possible because of the myriad of different rate structures in California, the idiosyncratic nature of energy use at the household level, and the different PV system designs and orientations.

likely to be more similar between PV and non-PV homes at the subdivision level than at the block level, and therefore this model may more-effectively control for such omitted variables.[25]

To address the issue of selection bias, one robustness model uses the same model specification as the base model but with an alternative (subset) of the data sample. Specifically, instead of using the full dataset with equations (1) and (2), a "coarsened exact matched" dataset can be used (King et al., 2010).[26] Because the PV and non-PV datasets are statistically equal on their covariates after the matching process, biases related to selection are minimized.

Finally, specific to equation (2), a robustness model to address both omitted variable and selection bias was constructed in which the sample was restricted to *only* include PV homes (in place of the full sample of PV *and* non-PV homes). In other words, because this model does not include non-PV "comparable" homes, sales prices of PV homes are "compared" against each other based on the size of the PV systems, while controlling for the differences in the home via the controlling characteristics (e.g., square feet of living space). Because PV system size effects are estimated without the use of non-PV homes in this model, it provides an important comparison to the base models, while also directly addressing any concerns about the inherent differences between PV and non-PV homes and therefore omitted variable and sample selection bias.

3.4. New and Existing Home Models

Although equations (1) and (2) are used to estimate whether a PV system, on average, effects selling prices across the entire data sample, they do not allow one to distinguish any such effects as a function of house type, specifically whether the home is *new* or *existing*. As discussed earlier, *new* homes with PV might have different premiums than *existing* homes. To try to tease out these possible differences two base hedonic models are estimated using equation (2), one with only *new* homes and the other with only *existing* homes.[27] Comparing the coefficient of the

[25] Subdivisions are often geographically smaller than block groups, and therefore more accurately control for geographical influences such as distance to central business district. Moreover, homes in the same subdivision are often built at similar times using similar materials and therefore serve as a control for a variety of house specific characteristics that are not controlled for elsewhere in the model. For example, all homes in a subdivision will often be built using the same building code with similar appliances being installed, both of which might control for the underlying energy efficiency (EE) characteristics of the home. For homes not situated in a subdivision, the block group delineation was used, and therefore these fixed effects are referred to as "combined subdivision-block group" delineations.

[26] The procedure used, as described in the referenced paper, is coarsened exact matching (cem) in Stata, available at: http://ideas.repec.org/c/boc/bocode/s457127.html. The matching procedure creates statistically matching sets of PV and non-PV homes in each block group, based on a set of covariates, which, for this research, include the number of square feet, acres, and baths, as well as the age of the home, its elevation, and the date at which it sold. Because this matching process excludes non-PV homes that are without a statistically similar PV match (and vice versa), a large percentage of homes (approximately 80% non-PV and 20% PV) are *not* included in the resulting dataset.

[27] *New* and *existing* homes were determined in an iterative process. For PV homes, the type of home was often specified by the data provider. It was also discovered that virtually all of the *new* PV homes (as specified by the PV data providers) had ages, at the time of sale, between negative one and two years, inclusive, whereas the *existing* PV homes (as specified by the PV data providers) had ages greater than two years in virtually every case. The small percentage (3%) of PV homes that did not fit these criteria were excluded from the models. For non-PV homes, no data specifying the home type were available, therefore, groupings were created following the age at sale criteria used for PV homes (e.g., ages between negative one and two years apply to *new* non-PV homes).

variable of interest (β_4) between these two models allows for an assessment of the relative size of the impact of PV systems across the two home types.

Additionally, the two sets of robustness models that were discussed above for the *new* and *existing* home specifications are explored, one using the coarsened exact matched dataset and the other using the combined subdivision-block group delineations. These models test the robustness of the results for selection and omitted variable bias, respectively. Although it is discussed separately as a base model in the following subsection, the difference-in-difference model, using repeat sales of *existing* homes, also doubly serves as a robustness test to the *existing* homes base model.

3.4.1. Difference-in-Difference Models

One classic alternative to estimating a hedonic model, as briefly discussed earlier, is to estimate a difference-in-difference (DD) model (Wooldridge, 2009). This model (see Table 1) uses a set of homes that have sold twice, both with and without PV, and provides estimates of the effect of adding PV to a subset of those homes as of the second sale ("DD" as noted in Table 1), while simultaneously both accounting for both the inherent differences in the PV and non-PV groups and the trend in housing prices between the first and second sales. Repeat sales models of this type are particularly effective in controlling for selection and certain types of omitted variable bias. In the former case, any underlying difference in home prices between PV and non-PV homes prior to the addition of PV is controlled for. In the latter case, PV and non-PV homes are assumed to have undergone mostly similar changes (e.g., home improvements) between sales. Changes to the home that are coincident with the installation of the PV system are not directly controlled for, but are assumed to be *di minimis*.[28] Therefore, to a large extent, the resulting inflation-adjusted sale price difference can be confidently attributed to the installation of PV.

The set of PV homes that are used in the DD model are, by default, *existing* homes (i.e., the home was not new when the PV system was installed). Estimates derived from this model, therefore, apply to - while also serving as a robustness tests for - the *existing* home models as specified above.

Table 6: Difference-in-Difference Description

	Pre PV	Post PV	Difference
PV Homes	PV_1	PV_2	$\Delta PV = PV_2 - PV_1$
Non-PV Homes	NPV_1	NPV_2	$\Delta NPV = NPV_2 - NPV_1$
			$DD = \Delta PV - \Delta NPV$
1 and 2 denote time periods			

[28] Support for this assumption comes from two sources. Although, surveys (e.g., CPUC, 2010) indicate that PV homeowners install energy efficient "measures" with greater frequency than non-PV homeowners, the differences are relatively small and largely focus on lighting and appliances. The former is not expected to impact sales prices, while the latter could. The surveys also indicate that indeed PV homeowners install other larger EE measures, such as building shell, water heating and cooling, with greater frequency too. Dastrop et al. (2010) investigated if these types of changes, which might require a permit, affect PV estimates, and found they were stable despite controlling for whether permits are taken out prior to a home's sale. Combined these give us support for our assumption, but needless to say, further research in the area of coincident improvements is warranted.

The base DD model is estimated as follows:

$$\ln(P_{itk}) = \alpha + \beta_1(T_t) + \beta_2(N_k) + \sum_a \beta_3(X_i) + \beta_4(PVH_i) + \beta_5(SALE2_i) + \beta_6(PVS_i) + \varepsilon_{itk} \qquad (3)$$

where

PVH_i is a fixed effect variable indicating if a PV system is or <u>will be</u> installed on the home in transaction *i*,

$SALE2_i$ is a fixed effect variable indicating if transaction *i* is the second of the two sales,

PVS_i is a fixed effect variable (an interaction between PVH_i and $SALE2_i$) indicating if transaction *i* is both the second of the two sales and contained a PV system at the time of sale,

α is the constant or intercept across the full sample, and represents the base value of non-PV homes as of the initial sale (i.e., "NPV_1" from Table 1),

β_4 is a parameter estimate for homes that have or will have PV installed (i.e., "$PV_1 - NPV_1$"),

β_5 is a parameter estimate if transaction *i* occurred as of the second sale (i.e., "ΔNPV"),

β_6 is a parameter estimate if transaction *i* occurred as of the second sale and the home contained PV (i.e., "ΔPV – ΔNPV" or "DD"), and all other terms are as were defined for equation (1).

The coefficient of interest is β_6, which represents the percentage change in sale price, as expressed in 2009 dollars, when PV is added to the home, after accounting for the differences between PV and non-PV homes (β_4) and the differences between the initial sale and the second sale (β_5). If differences in selling prices exist between PV and non-PV homes, we would expect the coefficient to be positive and statistically significant.[29]

To further attempt to mitigate the potential for omitted variable bias, two robustness models are estimated for the base DD model: one with the combined subdivision-block group delineations and a second with a limitation applied on the number of days between the first and second sales.[30] The first robustness model is similar to the one discussed earlier. The second robustness model accounts for the fact that the home characteristics used (in all models) reflect the most recent home assessment, and therefore do not necessarily reflect the characteristics at the time of the sale for both transactions equally. Therefore, especially worrisome are the first sales in the DD model, which can be as much as 20 years before the second sale. To test if our results are biased because of these older sales - and the large periods between sales - an additional data screen is applied in which the difference between the two sale dates is limited to five years.[31]

3.5. Age of the PV System for Existing Homes Models

As discussed above, the age of the PV system could affect the premium the systems on *existing* homes garner. This would occur because not only is the system less efficient and potentially also has a shorter expected life (and therefore a lower net present value of bill savings), it has also

[29] This is the classic model form derived from a quasi-experiment, where the installation of PV is the treatment. An alternative specification would look at the incremental effect of PV system size holding the starting differences between PV and non-PV homes as well as the time-trend in non-PV homes constant. This model form was not evaluated in the current analysis effort, but could be considered grounds for future research in this area.

[30] Ideally a matched dataset could be utilized, for reasons described earlier, but because the matching procedure severely limits the size of the dataset, the resulting dataset was too small to be useful.

[31] As was discussed above, a screen for this eventuality (using *adjaar*) is incorporated in our data cleaning, so, this test serves as an additional check of robustness of the results.

returned a portion of the original installation cost to the seller. All of these factors would indicate that premiums for older systems on *existing homes* would be lower than for newer systems *ceteris paribus*. In order to test this directly the following base model is estimated:

$$\ln(P_{itk}) = \alpha + \beta_1(T_t) + \beta_2(N_k) + \sum_a \beta_3(X_i) + \beta_4(PV_i \cdot SIZE_i \cdot AGE_i) + \varepsilon_{itk} \quad (4)$$

where AGE_i is a categorical variable for three groups of PV system age as of the time of sale of the home: 1) less than or equal to one year old; 2) between 2 and 4 years old; and, 3) more than four years old. Therefore, β_4 is a vector of parameter estimates for the percentage change in sales price for each additional kW added to a PV system for each of the three PV system age groups, and all other terms are as are defined for equation (2). The assumption is that the coefficients for β_4 will be decreasing - indicating they are valued less - as the age of the systems decrease. The sample used for this model is the same as for the *existing* home model defined previously.

Additionally, two sets of robustness models for the age of the PV system specifications are explored, one using the coarsened exact matched dataset and, the other using the combined subdivision-block group delineations, to test the robustness of the results for selection and omitted variable bias, respectively.

3.6. Returns to Scale Hedonic Models

As discussed earlier, it is not unreasonable to expect that any increases in the selling prices of PV homes may be non-linear with PV system size. In equation (2), it was assumed that estimated price differences were based on a continuous linear relationship with the size of the system. To explore the possibility of a non-linear relationship among the full sample of homes in the dataset the following model is estimated:[32]

$$\ln(P_{itk}) = \alpha + \beta_1(T_t) + \beta_2(N_k) + \sum_a \beta_3(X_i) + \beta_4(PV_i \cdot SIZE_i) + \beta_5(PV_i \cdot SIZE_i \cdot SIZE_i) + \varepsilon_{itk} \quad (5)$$

where β_5 is a parameter estimate for the percentage change in sales price for each additional kW added to a PV system squared, and all other terms are as are defined for equation (2). A negative statistically significant coefficient (β_5) would indicate decreasing returns to scale for larger PV systems, while a positive coefficient would indicate the opposite.

Somewhat analogously, as was discussed previously, premiums for PV systems may be related to the size of the home.[33] To test this directly using the full dataset, the following model is estimated:

[32] Neither this nor the following model is coupled with robustness models in this paper.

[33] Of course, PV system size is somewhat correlated with house size, because house size is correlated with energy use. If this correlation were strong estimates could be biased, but we assume this not to be the case because a variety of factors, other than size of the home, determine the ultimate size chosen for the PV system, including size of the roof, available capital, and home energy use (which, in turn, could be influenced by such things as whether the home has a pool or not, or how many appliances the home has). More tellingly, the correlation between PV house size and PV system size in the sample is 0.14.

$$\ln(P_{itk}) = \alpha + \beta_1(T_t) + \beta_2(N_k) + \sum_a \beta_3(X_i) + \beta_4(SQFT_i) + \beta_5(PV_i \cdot SIZE_i) + \beta_6(PV_i \cdot SIZE_i \cdot SQFT_i) + \varepsilon_{itk} \quad (6)$$

where

$SQFT_i$ is a continuous variable for the number of square feet for the home in transaction *i*,[34]

β_4 is a parameter estimate for the percentage change in sale price for each additional 1000 square feet added to the home,

β_5 is a parameter estimate for the percentage change in sale price for each additional kW added to a PV system,

β_6 is a parameter estimate for the percentage change in sale price for each additional 1000 square feet added to PV homes, assuming the size of the PV system does not change, and all other terms are as were defined for equation (2).

A negative statistically significant coefficient for β_6 would indicate decreasing returns to scale for PV systems as homes increase in size. Alternatively, a positive and statistically significant coefficient would indicate increasing returns to scale for PV systems installed on larger homes.

3.7. Model Summary

To summarize, the entire set of estimated models discussed herein is shown in Table 7. The following definitions of terms, all of which were discussed earlier, are relevant for interpreting the models listed in the table, and therefore are briefly reviewed again. All "base" models are coupled with a set of "robustness" models (as noted by a capital "R" in the model number). The "Other" (returns to scale) models are presented alone. Models 1 - 4 and 6 - 8 use the hedonic pricing model whereas Model 5 is based on the difference-in-difference (DD) model. "Fixed" (versus "continuous") means that the PV variable is entered into the regression as a zero-one dichotomous variable (for Models 1-1Rb and 5-5Rb), whereas "continuous" (for all other models) means that the model estimates the impact of an increase in PV system size on residential selling prices. Base Models 1, 2, 7 and 8 use the full dataset, while Models 4 and 6 are restricted to *existing* homes, Model 3 to *new* homes, and Model 5 to the repeat sales dataset. The "matched" models use the smaller dataset of coarsened exact matched (PV and non-PV) homes. "Base" models estimate neighborhood fixed effects at the block group level, whereas the "subdivision" models estimate neighborhood fixed effects at the combined subdivision-block group level.

[34] In all of the previous models the number of square feet is contained in the vector of characteristics represented by X_i, but in this model it is separated out for clarity.

Table 7: Summary of Models

Model Number	Model Name	Base Model	Robustness Model	Other Models	Dataset	Neighborhood Fixed Effects
1	Fixed - Base	**X**			Full	Block Group
1Ra	Fixed - Matched		**X**		Full Matched	Block Group
1Rb	Fixed - Subdivision		**X**		Full	Subdivision/Block Group
2	Continuous - Base	**X**			Full	Block Group
2Ra	Continuous - Matched		**X**		Full Matched	Block Group
2Rb	Continuous - Subdivision		**X**		Full	Subdivision/Block Group
2Rc	Continuous - PV Only		**X**		PV Only	Block Group
3	New Homes - Base	**X**			New	Block Group
3Ra	New - Matched		**X**		New - Matched	Block Group
3Rb	New - Subdivision		**X**		New	Subdivision/Block Group
4	Existing Homes - Base	**X**			Existing	Block Group
4Ra	Existing - Matched		**X**		Existing - Matched	Block Group
4Rb	Existing - Subdivision		**X**		Existing	Subdivision/Block Group
5	Difference-in-Difference (DD) - Base	**X**			Repeat Sales	Block Group
5Ra	Difference-in-Difference (DD) - Subdivision		**X**		Repeat Sales	Subdivision/Block Group
5Rb	Difference-in-Difference (DD) - Sddif < 5 Years		**X**		Repeat Sales w/ sddif < 5	Block Group
6	Age of System - Base	**X**			Existing	Block Group
6Ra	Age of System - Matched		**X**		Existing - Matched	Block Group
6Rb	Age of System - Subdivision		**X**		Existing	Subdivision/Block Group
7	Returns to Scale - Size			**X**	Full	Block Group
8	Returns to Scale - Square Feet			**X**	Full	Block Group

4. Estimation Results

Estimation results for all 18 models (as defined in Table 1) are presented in Tables 8-11, with the salient results on the impacts of PV on homes sales prices summarized in Figures 2-4.[35] The adjusted R^2 for all models is high, ranging from 0.93 to 0.95, which is notable because the dataset spanned a period of unusual volatility in the housing market.[36] The model performance reflects, in part, the ability of the inflation index and temporal fixed effects variables to adequately control for market conditions.[37] Moreover, the sign and magnitude of the home and site control variables are consistent with *a priori* expectations, are largely stable across all models, and are statistically significant at the 1% level in most models.[38] Each additional 1000 square feet of living area added to a home is estimated to add between 19% and 26% to its value, while the first acre adds approximately 40% to its value with each additional acre adding approximately 1.5%. For each year a home ages, it is estimated that approximately 0.2% of its value is lost, yet at 60 years, age becomes an asset with homes older than that estimated to garner premiums for each additional year in age. Finally, for each additional 100 feet above the median elevation of the other homes in the block group, a home's value is estimated to increase by approximately 0.3%.

These results can be benchmarked to other research. Specifically, Sirmans et al. (2005a; 2005b) conducted a meta-analysis of 64 hedonic studies carried out in multiple locations in the U.S. during multiple time periods, and investigated similar characteristics as included in the models presented here, except relative elevation. As a group, each of the home and site characteristic estimates in the present study differ from the mean Sirmans et al. estimates by no more than one half of one standard deviation. In summary, these results suggest that the hedonic and repeat sales models estimated here are effectively capturing many of the drivers to homes sales prices in California, increasing confidence that those same models can be used to capture any PV effects.

4.1. Fixed and Continuous Effect Hedonic Model Results

The results from the base hedonic models (equations 1 and 2) are shown in Table 8 as Models 1 and 2, respectively. These models estimate the differences across the full dataset between PV and non-PV homes, with Model 1 estimating this difference as a fixed effect, and Model 2 estimating the difference as a continuous effect for each additional kilowatt (kW) of PV added. Also shown in the table are the results from the robustness tests using the coarsened exact matching

---

[35] For simplicity, this paper does not present the results for the quarter and block group (nor combined subdivision-block group) fixed effects, which consist of more than 900 coefficients. These are available upon request from the authors.

[36] All models were estimated with Stata SE Version 11.1 using the "areg" procedure with White's correction for standard errors (White, 1980). It should also be noted that all Durbin-Watson (Durbin and Watson, 1951) test statistics were within the acceptable range (Gujarati, 2003), there was little multicollinearity associated with the variables of interest, and all results were robust to the removal of any cases with a Cook's Distance greater than 4/*n* (Cook, 1977) and/or standardized residuals greater than four.

[37] As mentioned in footnote 23, a variety of approaches were tested to control for market conditions, such as spatial temporal fixed effects (e.g., census block / year quarter) both with and without adjusted sale prices, and the models presented here were the most parsimonious and the results were robust to the various specifications, which, in turn, provides additional confidence that the effects presented are not biased by market conditions.

[38] In some models, where there is little variation between the cases on the covariate (e.g., acres), the results are non-significant at the 10% level.

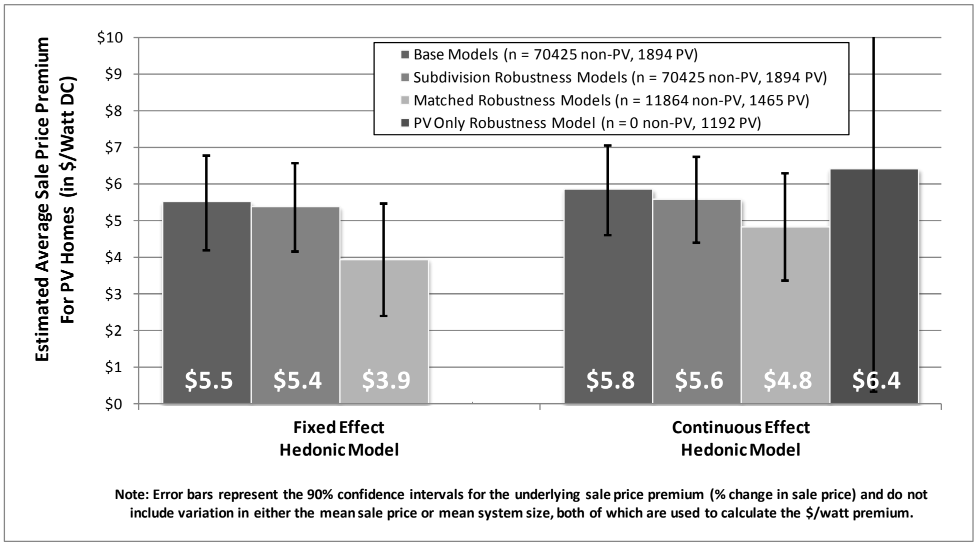
procedure and the combined subdivision-block group delineations, as shown as Models 1Ra and 1Rb for PV fixed effect models, and Models 2Ra and 2Rb for continuous effect variables. Finally, the model that derives marginal impact estimates from only PV homes is shown in the table as Model 2Rc.

Across all seven of these models (Models 1 – 2Rc), regardless of the specification, the variables of interest of PV and SIZE are positive and significant at the 10% level, with six out of seven estimates being significant at the 1% level. Where a PV fixed effect is estimated, the coefficient can be interpreted as the percentage increase in the sales price of a PV home over the mean non-PV home sales price in 2009 dollars based on an average sized PV system. By dividing the monetary value of this increase by the number of watts for the average sized system, this premium can be converted to 2009 dollars per watt ($/watt). For example, for Model 1, multiplying the mean non-PV house value of $480,862 by 0.036 and dividing by 3120 watts, yields a premium of $5.5/watt (see bottom of Table 8). Where SIZE, a continuous PV effect, is used, the coefficients reflect the percentage increase in selling prices in 2009 dollars for each additional kW added to the PV system. Therefore, to convert the SIZE coefficient to $/watt, the mean house value for non-PV homes is multiplied by the coefficient and divided by 1000. For example, for Model 2, $480,862 is multiplied by 0.012 and divided by 1000 resulting in an estimate of $5.8/watt.[39]

As summarized in Figure 2, these base model results for the impact of PV on residential selling prices are consistent with those estimated after controlling for subdivision fixed effects ($5.4/watt and $5.6/watt for fixed and continuous effects, respectively), differing by no more than $0.2/watt. On the other hand, the estimated PV premiums derived from the coarsened exact matched dataset are noticeably smaller, decreasing by 20 to 30%, and ranging from $3.9/watt to $4.8/watt for fixed and continuous effects, respectively. Alternatively, the PV only Model 2Rc estimates a higher $/watt continuous effect of $6.4/watt, although that estimate is only statistically significant at the 10% level.

[39] To be accurate the conversion is a bit more complicated. For example, for the fixed effect model the conversion is actually (EXP(LN(480,862)+0.036)-480,862)/3.12/1000, but the differences are de minimis, and therefore are not used herein.

Figure 2: Fixed and Continuous Effect Base Model Results with Robustness Tests



Though results among these seven models therefore differ to some degree, the results are consistent in finding a premium for PV homes over non-PV homes in California, which varies from $3.9 to $6.4/watt on average, depending on the model specification. These sale price premiums are very much in line with, if not slightly above, the historical mean net (i.e., after available state and federal incentives) installed costs of residential PV systems in California of approximately $5/watt from 2001 through 2009 (Barbose et al., 2010), which, as discussed above, is reasonable given that both buyers and sellers might use this cost as a basis to value the home.[40]

Additionally, the one other hedonic analysis of PV selling price premiums (which used reasonably similar models as those employed here but a different dataset, concentrating only on homes in the San Diego metropolitan area) found a similar result (Dastrop et al., 2010). In their analysis of 279 homes that sold with PV systems installed in San Diego (our model only contained 35 homes from this area[41] – See Table 5), Dastrop et al. estimated an average increase in selling price of $16,235, which, when divided by their mean PV system size of 3.2 kW, implies an effect of approximately $5/watt.[42]

[40] Although not investigated here, one possible reason why sale price premiums could be above net installed costs is that buyers of homes with PV are pricing in the opportunity cost of avoiding having to do this upgrade themselves, which might be perceived as complex. Moreover, any upgrades after the purchase would likely be financed outside the first mortgage and therefore would likely cost more on net.

[41] Even though we identified a higher number of PV homes that sold in the San Diego metropolitan area in our dataset, the home and site characteristics provided to us from the real estate data provider did not contain information on the year of the sale and therefore were not usable.

[42] In a different model, Dastrup et al. estimated an effect size of $2.4/watt but, for reasons not addressed here, this estimate is not believed to be as robust.

Table 8: Fixed and Continuous Base Hedonic Model Results with Robustness Tests

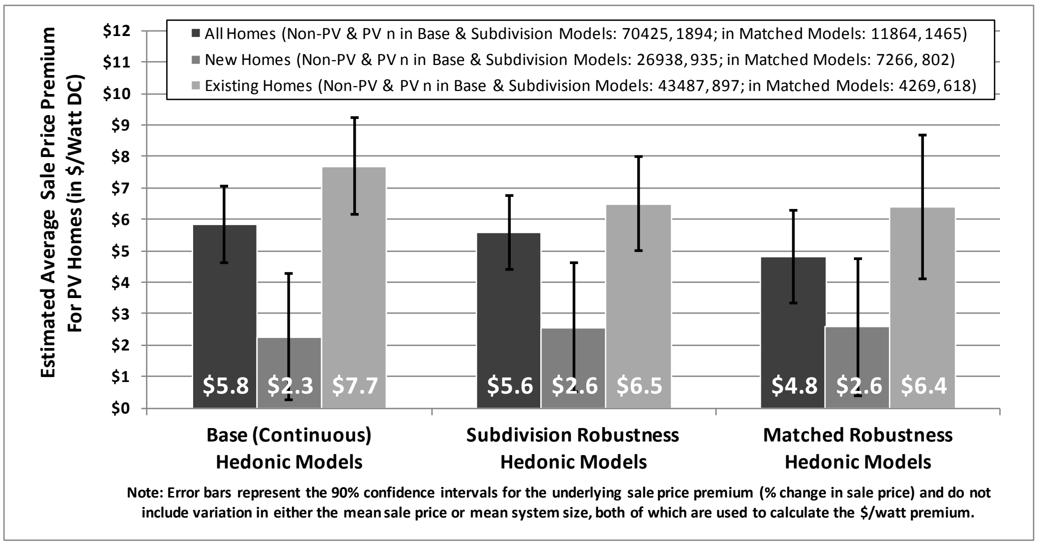
	Fixed			Continuous			
	Base	Robustness Matched	Robustness Subdivision	Base	Robustness Matched	Robustness Subdivision	Robustness PV Only
	Model 1	Model 1Ra	Model 1Rb	Model 2	Model 2Ra	Model 2Rb	Model 2Rc
pv	0.036***	0.024***	0.035***				
	(0.005)	(0.006)	(0.005)				
size				0.012***	0.010***	0.012***	0.013*
				(0.002)	(0.002)	(0.001)	(0.008)
sqft_1000	0.253***	0.205***	0.250***	0.253***	0.205***	0.250***	0.224***
	(0.001)	(0.006)	(0.001)	(0.001)	(0.006)	(0.001)	(0.010)
lt1acre	0.417***	0.514***	0.414***	0.416***	0.510***	0.413***	0.441***
	(0.009)	(0.040)	(0.010)	(0.009)	(0.040)	(0.010)	(0.066)
acre	0.016***	0.013	0.015***	0.016***	0.013	0.015***	-0.002
	(0.002)	(0.011)	(0.003)	(0.002)	(0.010)	(0.003)	(0.012)
ages2	-0.004***	-0.006***	-0.004***	-0.004***	-0.006***	-0.004***	-0.008***
	(0.0002)	(0.0012)	(0.0002)	(0.0002)	(0.0012)	(0.0002)	(0.0030)
ages2sqr	0.00003***	0.00004***	0.00003***	0.00003***	0.00004***	0.00003***	0.00004***
	(0.000003)	(0.000012)	(0.000003)	(0.000003)	(0.000012)	(0.000003)	(0.000033)
bgre_100	0.003***	0.015***	0.003***	0.003***	0.015***	0.003***	0.013***
	(0.001)	(0.004)	(0.001)	(0.001)	(0.004)	(0.001)	(0.005)
intercept	12.703***	12.961***	12.710***	12.702***	12.957***	12.710***	12.842***
	(0.010)	(0.044)	(0.012)	(0.010)	(0.043)	(0.012)	(0.073)
*Numbers in parenthesis are standard errors, *** p<0.01, ** p<0.05, * p<0.1*							
Results for subdivision, block group, and quarterly fixed effect variables are not reported here, but are available upon request from the authors							
Total n	72,319	13,329	72,319	72,319	13,329	72,319	1,192
Adjusted R^2	0.93	0.95	0.94	0.93	0.95	0.94	0.93
n (pv homes)	1,894	1,465	1,894	1,894	1,465	1,894	1,192
Mean non-pv asp2	$ 480,862	$ 480,533	$ 480,862	$ 480,862	$ 480,533	$ 480,862	$ 475,811
Mean size (kW)	3.1	3.0	3.1	3.1	3.0	3.1	2.7
Estimated $/Watt	$ 5.5	$ 3.9	$ 5.4	$ 5.8	$ 4.8	$ 5.6	$ 6.4
PV Only Model Notes: Mean non-pv asp2 amount shown is actually the mean PV asp2. Sample is limited to blockgroups with more than one PV home							

4.2. New and Existing Home Model Results

Turning from the full dataset to one specific to the home type, we estimate continuous effects models for *new* and *existing* homes (see equation (2)). These results are shown in Table 9, with Model 3 the base model for *new* homes and Model 4 the base model for *existing* homes. Also shown are the results from the robustness tests using the coarsened exact matching procedure and the combined subdivision-block group delineations, as Models 3Ra and 3Rb, respectively, for *new* homes, and as Models 4Ra and 4Rb, respectively, for *existing* homes.

The coefficient of interest, SIZE, is statistically significant at or below the 10% level in all of the *new* home models and at the 1% level in all of the *existing* home models. Estimates for the average $/watt increase in selling prices as a result of PV systems (as summarized in Figure 3, which also includes the results presented earlier for all homes, Models 2, 2Ra, and 2Rb) for *new* homes are quite stable, ranging from $2.3 to $2.6. Conversely, for PV sold with *existing* homes, not only are the selling price impacts found to be higher, but their range across the three models is greater, ranging from $ 6.4 to $7.7/watt.

Figure 3: New and Existing Home Base Model Results with Robustness Tests



The reasons for the apparent discrepancy in selling price impacts between *new* and *existing* homes are unclear, should be the focus of future research, but, might be explained, in part, by the difference in average net installed costs, which, from 2007 to 2009, were approximately $5.2/watt for *existing* homes and $4.2/watt for *new* homes (derived from the dataset used for Barbose et al., 2010). The gap in net installed costs between new and existing homes is not wide enough to fully account for these findings, however, and the model estimates for PV selling price premiums are below net installed costs for *new* homes and above net installed costs for *existing* homes. [43]

Alternative explanations for the disparity between *new* and *existing* home premiums exist, though. As discussed previously, there is evidence that builders of *new* homes might discount premiums for PV if, in exchange, having PV installed on the home differentiated the product and therefore increased the sales velocity, thus decreasing overall carrying costs (Dakin et al., 2008; SunPower, 2008; Griffin et al., 2009). Additionally, because many builders of *new* homes found offering PV as an option, rather than a standard feature, posed a set of difficulties (Farhar et al., 2004b; Griffin et al., 2009), they more regularly began installing PV as a standard feature on homes (Griffin et al., 2009). This potentially affected the valuation of the PV system for two reasons. First, because buyers of *new* homes often ranked PV as a less important than, for

[43] A small number of "affordable homes" (n = 7) are included in the *new* PV homes subset, which, as a group, appear to have a slight downward yet inconsequential effect on results, and therefore were not investigated further herein. If their numbers are significant future research though, they should be controlled for directly.

example, the location of the home or other qualities of the community (Farhar et al., 2004a; Griffin et al., 2009), they might have been less likely to pay the full price for this feature, because they might have been able to choose a home without PV in the same or similar community. Secondly, because sales agents for the *new* PV containing homes were either not well versed in the specifics of PV and felt that selling a PV system was a new sales pitch (Farhar et al., 2004b) or combined the discussion of PV with a set of other energy features (Griffin et al., 2009), up-selling the full value of the PV system might not have been possible.

Both of the downward influences for *new* homes are potentially contrasted with analogous upward influences for *existing* homes. E*xisting* homes might be less homogenous and potentially spread across a larger geographic area, therefore, replacing the *existing* PV home with a non-PV home in the same area might have been more difficult, and therefore garner a higher relative price. Secondly, in contrast to *new* home sellers, who might not be familiar with the intricacies and benefits of the PV system, *existing* home sellers are likely to be very familiar with the particulars of the system and it benefits, and therefore might be able to "up-sell" it more effectively. In summary, this set of postulates might explain some of the drivers to the premium disparity between *new* and *existing* homes, and, although reasonable, are not investigated directly here, and therefore should be the focus of future research.

Table 9: New and Existing Home Base Hedonic Model Results with Robustness Tests

	New Homes			Existing Homes		
	Base	Robustness Matched	Robustness Subdivision	Base	Robustnes Matched	Robustness Subdivision
	Model 3	Model 3Ra	Model 3Rb	Model 4	Model 4Ra	Model 4Rb
size	0.006*	0.006*	0.006**	0.014***	0.011***	0.012***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
sqft_1000	0.247***	0.190***	0.250***	0.256***	0.238***	0.251***
	(0.002)	(0.006)	(0.002)	(0.002)	(0.015)	(0.002)
lt1acre	0.536***	0.279***	0.517***	0.373***	0.426***	0.376***
	(0.019)	(0.073)	(0.024)	(0.010)	(0.046)	(0.012)
acre	-0.007	0.338***	-0.009*	0.019***	0.011	0.017***
	(0.005)	(0.027)	(0.005)	(0.002)	(0.011)	(0.003)
ages2	-0.010	0.081***	-0.010*	-0.005***	-0.006***	-0.005***
	(0.006)	(0.017)	(0.006)	(0.000)	(0.002)	(0.000)
ages2sqr	0.00768***	-0.02443***	0.00715***	0.00004***	0.00004***	0.00004***
	(0.001676)	(0.004407)	(0.001604)	(0.000003)	(0.000014)	(0.000004)
bgre_100	0.008***	0.027***	0.007***	0.002	-0.002	0.002
	(0.001)	(0.003)	(0.001)	(0.001)	(0.009)	(0.001)
intercept	12.651***	12.585***	12.627***	12.820***	13.023***	12.833***
	(0.022)	(0.066)	(0.025)	(0.013)	(0.077)	(0.014)
*Numbers in parenthesis are standard errors, *** p<0.01, ** p<0.05, * p<0.1*						
Results for subdivision, block group, and quarterly fixed effect variables are not reported here, but are available upon request from the authors						
Total n	27,873	8,068	27,873	44,384	4,887	44,384
Adjusted R^2	0.94	0.94	0.94	0.93	0.95	0.94
n (pv homes)	935	802	935	897	618	897
Mean non-pv asp2	$ 397,265	$ 399,162	$ 397,265	$ 532,645	$ 590,428	$ 532,645
Mean size (kW)	2.5	2.4	2.5	3.8	3.7	3.8
Estimated $/Watt	$ 2.3	$ 2.6	$ 2.6	$ 7.7	$ 6.4	$ 6.5

4.2.1. Difference-in-Difference Model Results

Delving deeper into PV system impacts on *existing* homes, Table 10 (and Figure 4) shows the results of the base Difference-in-Difference Model 5 as well as results from the two robustness tests (all of which can be compared to Models 4, 4Ra, and 4rb above). As a reminder, one robustness model limited the differences in sales dates between the first and second sales to five years (Model 5Rb), whereas the other robustness model used the combined subdivision-block group delineations as fixed effect variables (Model 5Rc). The variables of interest are PVH, SALE2 and especially PVS.

PVH estimates the difference in the first sale prices of homes that will have PV installed (as of the second sale date) non-PV homes. The three models are consistent in their estimates, showing approximately a 2% premium for "future" PV homes, though only two of these estimates are statistically significant, and then only at the 10% level. Regardless, this finding suggests that PV homes tend to sell for somewhat more even before the installation of PV, presumably as a result of other amenities that are correlated with the (ultimate) installation of PV (such as, potentially, energy efficiency features). SALE2 estimates the trend between the first and second sales for all homes. The coefficient for this variable is significant at the 1% level, and is fairly stable across the models, indicating a clear general trend of price increases, over and above inflation adjustments, of approximately 2% to 2.5% between the first and second sales.

Finally, and most importantly, homes with PV systems installed on them as of the second sale - after controlling for any inherent differences in first sale prices (PVH) and any trend between the first and second sales (SALE2) - show statistically significant sale price premiums of approximately 5 to 6%. These premiums equate to an increase in selling prices equivalent to approximately $6/watt for *existing* homes, closely approximating those presented earlier for the hedonic models in Table 9 and Figure 3. For comparison purposes, both sets of results are presented in Figure 4.

The premium for *existing* PV homes, as estimated in the DD Models 5, 5Ra, and 5Rb and both robustness tests for the hedonic model (using the "matched" and "subdivision" datasets, Models 4Ra and 4Rb respectively) are consistently between $6 and $6.5/watt and are in line with – though slightly higher than - the mean net installed costs of PV of approximately $5.2/watt, while the base *existing* home model estimates a premium of $7.7/watt. One possible explanation for this inconsistency is that the two robustness tests for the hedonic model and the various difference-in-difference models are less likely to be influenced by either selection or omitted variable bias than the base hedonic model, which estimates a larger premium of $7.7/W. Regardless of the absolute magnitude, a premium for *existing* PV homes over that garnered by *new* PV homes is clearly evident in these results.

Figure 4: Existing Home Hedonic and Difference-in-Difference Model Results with Robustness Tests

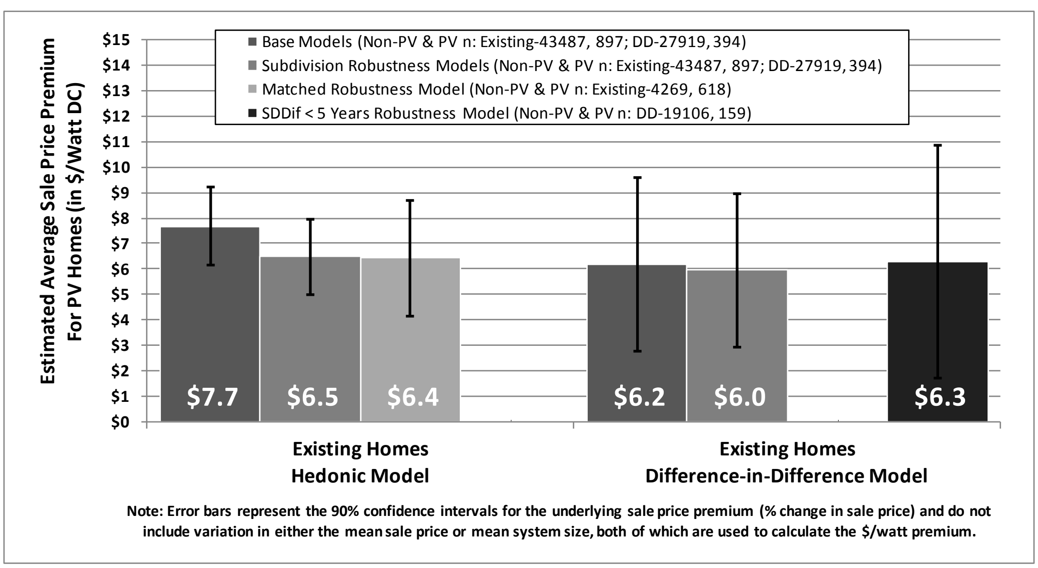


Table 10: Difference-in-Difference Model Results

	Difference-in-Difference		
	Base	Robustness	Robustness
		Subdivision	Sddif < 5
	Model 5	Model 5Ra	Model 5Rb
pvh	0.022*	0.024	0.022*
	(0.013)	(0.021)	(0.012)
sale2	0.023***	0.026***	0.019***
	(0.002)	(0.002)	(0.002)
pvs	0.051***	0.061**	0.049***
	(0.017)	(0.027)	(0.015)
size			
sizesqr			
size*sqft_1000			
sqft_1000	0.255***	0.256***	0.251***
	(0.002)	(0.002)	(0.002)
lt1acre	0.374***	0.385***	0.377***
	(0.011)	(0.013)	(0.012)
acre	0.012***	0.009**	0.011***
	(0.003)	(0.004)	(0.003)
age	-0.005***	-0.005***	-0.005***
	(0.0002)	(0.0003)	(0.0003)
agesqr	0.00004***	0.00004***	0.00004***
	(0.000003)	(0.000003)	(0.000003)
bgre_100	0.002*	0.000	0.001
	(0.001)	(0.001)	(0.001)
intercept	12.677***	12.594***	12.694***
	(0.013)	(0.015)	(0.014)
*Numbers in parenthesis are standard errors. *** p<0.01, ** p<0.05, * p<0.1. Results for subdivision, block group, and quarterly fixed effect variables are not reported here, but are available upon request from the authors*			
Total n	28,313	19,265	28,313
Adjusted R^2	0.93	0.94	0.94
n (pv homes)	394	159	394
Mean non-pv asp2	$ 488,127	$ 450,223	$ 488,127
Mean size (kW)	4.0	4.3	4.0
Estimated $/Watt	$ 6.2	$ 6.3	$ 6.0

4.3. Age of PV System for Existing Home Model Results

To this point, the marginal impacts to selling prices of each additional kW of PV added to *existing* homes have been estimated using the full dataset of *existing* homes, which has produced an average effect, regardless of the age of the PV system. As discussed previously, it is conceivable that older PV systems would garner lower premiums than newer, similarly sized systems. Therefore, to test this directly, a base model is constructed – see Equation (4) - that estimates the marginal impacts for three age groups of PV systems: no more than one year old at the time of sale; between two and four years old; and, more than four. Results from this model as well as two robustness tests, using the coarsened exact matching procedure and the combined subdivision-block group delineations, are shown in Table 11 as Models 6, 6Ra, and 6Rb, respectively.

Each model finds significant differences between PV and non-PV homes for each age group, and more importantly, premium estimates for newer systems are larger than those for older systems and are monotonically ordered between groups. This provides some evidence that older systems are being discounted by the buyers and sellers of PV homes, but because the differences between the estimates of these various groups are not significant, it is not possible to conclude this with any confidence.

4.4. Returns to Scale Hedonic Model Results

To this point, the marginal impacts to selling prices of each additional kW of PV in the continuous models have been estimated using a linear relationship. To test whether a non-linear relationship may be a better fit, a SIZE squared term is added to the model as shown in equation (5). Similarly, decreasing or increasing returns to scale might be related to other house characteristics, such as the size of the home (i.e., square feet). This hypothesis is explored using equation (6). Both model results are shown in Table 11 as Model 7 and 8, respectively.

Both models find small and non-statistically significant relationships between their interacted variables, indicating a lack of compelling evidence of a non-linear relationship between PV system size and selling price in the dataset, and a lack of compelling evidence that the linear relationship is affected by the size of the home. As such, the impact of PV systems on residential selling prices appears to be well approximated by a simple linear relationship, while the size of the home is not found to impact the PV sales price premium. In combination, these results seem to suggest that while California's tiered rate structures may impact the energy bill savings of from PV investments to vary non-linearly with PV system size, those same rate structures have not – to this point – led to any corresponding non-linear relationship between the PV premium garnered at the time of home sale and the size of the PV system.

Table 11: Age of PV System and Return to Scale Models

	Age of PV Systems for Existing Homes			Returns to Scale	
	Base	Robustness Matched	Robustness Subdivision	Size	Square Feet
	Model 6	Model 6Ra	Model 6Rb	Model 7	Model 8
size*1 year old	0.016***	0.016***	0.013***		
	(-0.004)	(-0.005)	(-0.004)		
size*2-4 years old	0.015***	0.010***	0.013***		
	(-0.002)	(-0.003)	(-0.002)		
size*5+ years old	0.012***	0.008**	0.008**		
	(-0.003)	(-0.004)	(-0.003)		
size				0.008**	0.021***
				(0.003)	(0.006)
sizesqr				0.001	
				(0.001)	
size*sqft_1000					-0.003
					(0.002)
sqft_1000	0.256***	0.238***	0.251***	0.253***	0.253***
	(0.002)	(0.015)	(0.002)	(0.001)	(0.001)
lt1acre	0.373***	0.426***	0.376***	0.416***	0.416***
	(0.010)	(0.046)	(0.012)	(0.009)	(0.009)
acre	0.019***	0.010***	0.017***	0.016***	0.016***
	(0.002)	(0.011)	(0.003)	(0.002)	(0.002)
ages2	-0.005***	-0.006***	-0.005***	-0.004***	-0.004***
	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
ages2sqr	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
bgre_100	0.002***	-0.002***	0.002***	0.003***	0.003***
	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)
intercept	12.820***	13.024***	12.834***	12.702***	12.701***
	(0.013)	(0.078)	(0.014)	(0.010)	(0.011)
*Numbers in parenthesis are standard errors. *** p<0.01, ** p<0.05, * p<0.1*					
Total n	44,384	4,887	44,384	72,319	72,319
Adjusted R^2	0.93	0.95	0.94	0.93	0.93
n (pv homes)	897	618	897	1,894	1,894
Mean non-pv asp2	$ 532,645	$ 590,428	$ 532,645	$ 480,862	$ 480,862
Mean size (kW)	3.8	3.7	3.8	3.1	3.1
Estimated $/Watt	$8.3 - $6.1	$9.3 - $4.9	$7.0 - $4.1	$ 6.3	$ 6.4

Note: $/watt estimates for Returns to Scale models include the non-statistically significant interaction coefficients and therefore should be interpreted with caution

5. Conclusions

The market for solar PV is expanding rapidly in the U.S. More than 100,000 PV systems have been installed in California alone, 95% of which are residential systems. Some of those "PV homes" have sold, yet little research exists estimating if those homes sold for significantly more than similar non-PV homes. Therefore one of the claimed incentives for solar homes – namely that a portion of the initial investment into a PV system will be recouped at the time of sale - has largely been built on speculation. Practitioners have transferred the results from past research focused on energy efficiency and, while recent research has turned to PV, that research has so far focused on smaller sets of PV homes concentrated in one geographic area. Moreover, the effect of installing PV on a *new* versus an *existing* home has not previously been the subject of research, nor have determinations if the relationship of PV system size and sale price impacts are linear, and/or are affected by either the size of the home or the age of the PV system.

This research has used a dataset of approximately 72,000 California homes, approximately 2,000 of which had PV systems installed at the time of sale, and has estimated a variety of different hedonic and repeat sales models to directly address the questions outlined above. Moreover, an extensive set of robustness tests were incorporated into the analysis to test and bound the possible effects and increase the confidence of the findings by directly addressing potential biases. The research was not intended to disentangle the various individual effects that might dictate the level of premium, such as, energy costs savings, the net (i.e., after applicable state and federal incentives) installed cost of the PV system, or the possible presence of a green cachet, but rather to establish a credible estimate of the aggregate PV residential sale price effect.

The research finds strong evidence that homes with PV systems have sold for a premium over comparable homes without PV systems. More specifically, estimates for the PV premiums range from approximately \$3.9 to \$6.4 per installed watt (DC) among a large number of different model specifications, with most models coalescing near \$5.5, which corresponds to a premium of approximately \$17,000 for a 3,100 watt PV system. To benchmark these results they can be compared to a variety of different previous findings. First, and most importantly, the results found here are similar to the average increase for PV homes found by Dastrop et al. (2010), which used similar methods but a different dataset, one that focused on homes in the San Diego metropolitan area. Secondly, the results are of similar levels of the net (i.e., after applicable state and federal incentives) installed cost of California residential PV systems from 2001-2009 (Barbose et al., 2010) of approximately \$5/watt.

Although the results for the full dataset from the variety of models are quite similar, when the dataset is split among *new* and *existing* homes, PV system premiums are found to be markedly affected, with *new* homes showing premiums of \$2.3-2.6/watt, while *existing* homes show premiums of \$6-7.7/watt. Some reasons suggested for this disparity between *new* and *existing* PV homes, might include: differences in underlying net installation costs for PV systems; a tradeoff occurring for builders of new homes between recapturing the full net installed costs and efforts to differentiate the product leading to increased sales velocity leading to decreased carrying costs; a lack of familiarity and/or interest in marketing the PV system separately from the other features of the *new* home contrasted with a likely strong familiarity with the PV system for *existing* home sellers.

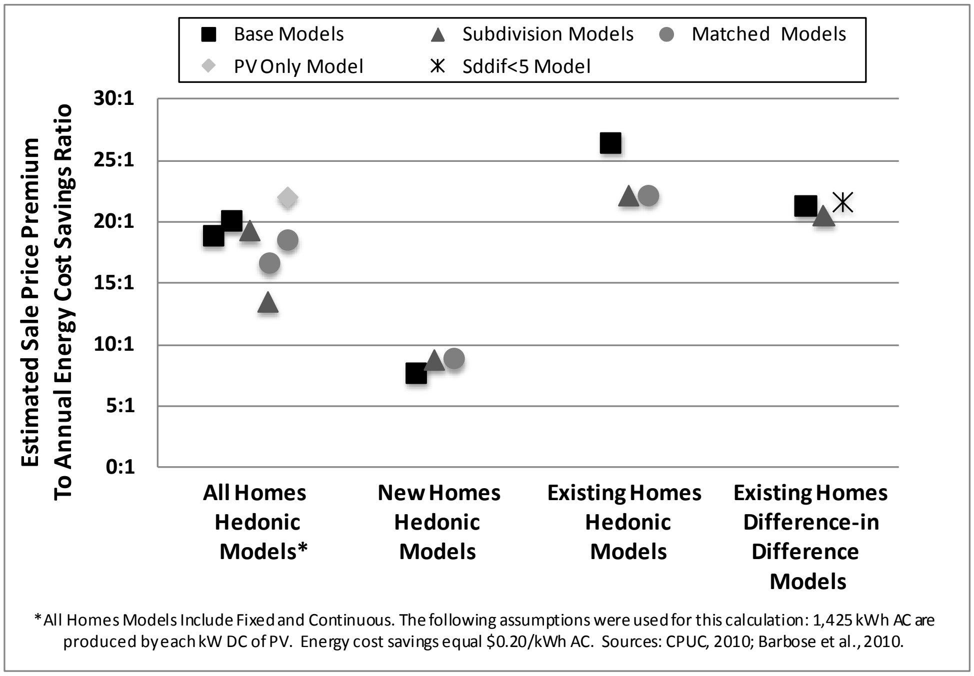
In addition to the results outlined above, the research also investigated differences in premiums for existing homes with PV systems of different ages finding some evidence that older systems are discounted in the marketplace as compared to newer systems. Finally, evidence of returns to scale for either larger PV systems or larger homes was investigated but not found.

In addition to benchmarking the results from the full model of $3.9-$6.4/watt, to previous literature investigating the sales price premiums associated with PV, they can also be compared to previous literature investigating premiums associated with energy efficiency (EE) or, more generally, energy cost savings. A number of these studies have converted this relationship into a ratio representing the relative size of the sales price premium to the annual savings expected due to energy bill reductions. These ratios have ranged from approximately 7:1 (Longstreth et al., 1984; Horowitz and Haeri, 1990), to 12:1 (Dinan and Miranowski, 1989) to approximately 20:1 (Johnson and Kaserman, 1983; Nevin et al., 1999; Eichholtz et al., 2009) to as high as 31:1 (Nevin and Watson, 1998).

Although actual energy bill savings from PV for the sample of homes used for this research were not available, a rough estimate is possible, allowing for a comparison to the previous results for EE. Specifically, assuming that 1,425 kWh (AC) are produced per year per kW (DC) of installed PV on a home (Barbose et al., 2010; CPUC, 2010)[44] that offset marginal retail electricity rates average $0.20/kWh (AC) (Darghouth et al., 2010), each watt (DC) of installed PV can be estimated to save $0.29 in energy costs a year. Using these assumptions, the $/watt PV premium estimates reported earlier can be converted to sale price to energy savings ratios, as shown in Figure 5. A $3.9 to $6.4/watt premium in selling price for an average California home with PV installed equates to a 14:1 to 22:1 sale price to energy savings ratio, respectively. For *new* homes, with a $2.3-2.6/watt sale price premium, this ratio is estimated to be 8:1 or 9:1, and for *existing* homes, with an overall sale price premium range of $6-7.6/watt, the ratio is estimated to range from 21:1 to 26:1. Without <u>actual</u> energy bill savings, these estimates are somewhat speculative, but nonetheless are broadly consistent with the previous research that has focused on EE-based home energy improvement and, therefore, further benchmark the results herein.

[44] The 1,425 kWh (AC) estimate is a combination of a 0.19 capacity factor (based on AC kWh and CEC-AC kW) from CPUC (2010), and an 0.86 conversion factor between CEC-AC kW and DC kW (Barbose et al., 2010).

Figure 5: Estimated Ratios of Sale Price Premium to Annual Energy Cost Savings



Although this research finds strong evidence that homes with PV systems have sold for a premium over comparable homes without PV systems, applicability of these results should be restricted to homes similar to those in our sample and for areas where similar net installed costs and electricity rate policies are applicable. Although applicable scenarios outside the state of California are not likely to be common, concluding that some premium exists for PV homes regardless of where the home is located is reasonable.

Finally, although this research provides a robust estimate for sale price premiums for PV homes, additional questions remain that merit further research. Perhaps most importantly, although the dataset used for this analysis consists of almost 2,000 PV homes, the study period was limited to sales occurring prior to mid-2009 and the dataset was limited to California. Future research would ideally include more-recent sales from a broader geographic area to better understand any regional/national differences that may exist as well as any changes to PV premiums that occur over time as the market for PV containing homes and/or the net installed cost of PV changes. More research should also be conducted for *new* versus *existing* homes to better understand the differential discovered in this research, which could include interviewing/surveying home builders and buyers and exploring the impact of demographic, socio-economic, and others factors on the PV premium.

Additionally, future research might compare actual home energy cost savings to sale price premiums, not only to explore the sale price to annual energy cost savings ratio directly, but also to explore if a green cachet exists over and above any sale price premiums that would be expected from energy savings alone. Further, house-by-house PV system and other information not included in the present study could be included in future studies, such as actual net installed costs of PV, rack-mounted or roof-integrated distinctions, the level of energy efficiency of the home, whether the home has a solar hot water heater, and whether the PV system was customer or 3rd party owned at the time of sale.[45] Such research could elucidate important differences in PV premiums among households and PV system designs, as well as bolster confidence in the magnitude of the PV premium estimated here. Finally, and more generally, additional research could investigate the impact of PV systems on the time homes remain on the market before sale, a factor which may be especially important for large developers and sellers of *new* homes.

[45] We assume that 3rd party owned systems would not be expected to command the same sort of premium as was discovered here. Although the level of penetration of 3rd party owners in our data was not significant (between 10 and 0%), and therefore would likely have not influenced our results in a substantive way, any future research, using more recent data, must account for their inclusion specifically.

References

Barbose, G., Darghouth, N. and Wiser, R. (2010) Tracking the Sun III: The Installed Cost of Photovoltaics in the U.S. From 1998-2009. Lawrence Berkeley National Laboratory, Berkeley, CA. December 2010. 54 pages. LBNL-4121E.

Black, A. (2010) Does It Pay? Figuring the Financial Value of a Solar or Wind Energy System. Solar Today. Fall/Winter 2010. pp. 26-27.

Brounen, D. and Kok, N. (2010) On the Economics of Energy Labels in the Housing Market. Program on Housing and Urban Policy: Working Papers Series. Prepared for Institute of Business and Economic Research and Fisher Center for Real Estate and Urban Economics, University of California, Berkeley, CA. August 2010. 34 pages. W10-002.

California Energy Commission (CEC) (2002) Consultant Report: Renewable Energy Study. The Phelps Group and ICF Consulting. Prepared for California Energy Commission (CEC). November 2001. 77 pages.

California Energy Commission (CEC) and California Public Utilities Commission (CPUC). (2011) California Solar Statistics. http://www.californiasolarstatistics.ca.gov/. (Accessed March 8, 2011).

California Public Utilities Commission (CPUC) (2010) CPUC California Solar Initiative: 2009 Impact Evaluation. Final Report. Prepared by: Itron and KEMA. Prepared for California Public Utilities Commission, Energy Division. June 2010. 632 pages.

Cook, R. D. (1977) Detection of Influential Observations in Linear Regression. *Technometrics.* 19(1): 15-18.

Dakin, W., Springer, D. and Kelly, B. (2008). Case Study: The Effectiveness of Zero Energy Home Strategies in the Marketplace. Presented at ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, California. August 17–22, 2008.

Darghouth, N., Barbose, G. and Wiser, R. (2010) The Impact of Rate Design and Net Metering on the Bill Savings from Distributed PV for Residential Customers in California. Lawrence Berkeley National Laboratory, Berkeley, CA. December 2010. 62 pages. LBNL-3276E.

Dastrop, S., Zivin, J. G., Costa, D. L. and Kahn, M. E. (2010) Understanding the Solar Home Price Premium: Electricity Generation and "Green" Social Status. Working Paper Series. Prepared for UC Center for Energy and Environmental Economics, Berkeley, CA. December 9, 2010. 29 pages. E3 WP-001.

Dinan, T. M. and Miranowski, J. A. (1989) Estimating the Implicit Price of Energy Efficiency Improvements in the Residential Housing Market: A Hedonic Approach. *Journal of Urban Economics.* 25(1): 52-67.

Durbin, J. and Watson, G. S. (1951) Testing for Serial Correlation in Least-Squares Regression. *Biometrika.* 38(1-2): 159-178.

Eichholtz, P., Kok, N. and Quigley, J. M. (2009) Doing Well by Doing Good? An Analysis of the Financial Performance of Green Office Buildings in the USA. University of California. Institute of Business and Economic Research. Berkeley Program on Housing and Urban Policy, Berkeley, CA. April 1, 2008. 49 pages. W08-001S.

Eichholtz, P., Kok, N. and Quigley, J. M. (2011) The Economics of Green Building. Working Paper Series Prepared for UC Center for Energy and Environmental Economics (UCE^3), Berkeley, CA. January 2011. 35 pages. WP-002.

Farhar, B. and Coburn, T. (2008) A New Market Paradigm for Zero-Energy Homes: A Comparative Case Study. *Environment: Science and Policy for Sustainable Development* 50(1): 18-32.

Farhar, B. and Coburn, T. C. (2006) A New Market Paradigm for Zero-Energy Homes: The Comparative San Diego Case Study. Volume 1 of 2. Prepared for National Renewable Energy Laboratory, Bolder, CO. December 2006. 413 pages. NREL/TP-550-38304-01.

Farhar, B. C., Coburn, T. C. and Murphy, M. (2004a) Comparative Analysis of Home Buyer Response to New Zero-Energy Homes. Summer Study on Energy Efficiency in Buildings, August 22-27, 2004. Prepared for American Council for an Energy-Efficient Economy, Pacific Grove, California. Preprinted in July, 2004. NREL/CP-550-35912.

Farhar, B. C., Coburn, T. C. and Murphy, M. (2004b) Large-Production Home Builder Experience with Zero Energy Homes. Summer Study on Energy Efficiency in Buildings, August 22-27, 2004. Prepared for American Council for an Energy-Efficient Economy, Pacific Grove, California. Preprinted in July, 2004. 15 pages. NREL/CP-550-35913.

Freeman, A. M. (1979) Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues. *Scandinavian Journal of Economics.* 81(2): 154-173.

Griffin, A., Kaufman, B. and Hamilton, S. (2009) Certified Home Performance: Assessing the Market Impacts of Third Party Certification on Residential Properties. Earth Advantage Institute. Prepared for Green Building Value Initiative Steering Committee, Portland, OR. May 29, 2009. 36 pages.

Gujarati, D. N. (2003) Basic Econometrics. McGraw-Hill/Irwin. Fourth Edition, New York. 1002 pages. ISBN 0-07-233542-4.

Hoen, B., Wiser, R., Cappers, P., Thayer, M. and Sethi, G. (2009) The Impact of Wind Power Projects on Residential Property Values in the United States: A Multi-Site Hedonic Analysis. Lawrence Berkeley National Laboratory, Berkeley, CA. December, 2009. 146 pages. LBNL-2829E.

Horowitz, M. J. and Haeri, H. (1990) Economic Efficiency vs. Energy Efficiency: Do Model Conservation Standards Make Good Sense? *Energy Economics.* 12(2): 122-131.

Johnson, R. C. and Kaserman, D. L. (1983) Housing Market Capitalization of Energy-Saving Durable Good Investments. *Economic Inquiry.* 21: 374 - 386.

King, G., Blackwell, M., Iacus, S. and Porro, G. (2010) Cem: Coarsened Exact Matching in Stata. *Stata Journal.* 9(4): 524-546.

Laquatra, J. (1986) Housing Market Capitalization of Thermal Integrity. *Energy Economics.* 8(3): 134-138.

Longstreth, M., Coveney, A. R. and Bowers, J. S. (1984) Conservation Characteristics among Determinants of Residential Property Value. *Journal of Consumer Research.* 11(1): 564-571.

Malpezzi, S. (2003) Hedonic Pricing Models: A Selective and Applied Review. Section in Housing Economics and Public Policy: Essays in Honor of Duncan Maclennan. Wiley-Blackwell. Hoboken, NJ. pp. 67-85 of 328 pages. ISBN 978-0-632-06461-8.

McCabe, M. and Merry, L. (2010) Resale Market Value of Residential Solar PV. *Journal of Sustainable Real Estate*. Volume 2. Industry Perspectives. Retrieved March 28, 2011 from http://www.costar.com/josre/industryPerspectives.htm

Nevin, R., Bender, C. and Gazan, H. (1999) More Evidence of Rational Market Values for Energy Efficiency. *The Appraisal Journal.* 67(4): 454-460.

Nevin, R. and Watson, G. (1998) Evidence of Rational Market Values for Home Energy Efficiency. *The Appraisal Journal.* 68: 401-409.
Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy.* 82(1): 34-55.
Simons, R. A. and Saginor, J. D. (2006) A Meta-Analysis of the Effect of Environmental Contamination and Positive Amenities on Residential Real Estate Values. *Journal of Real Estate Research.* 28(1): 71-104.
Sirmans, G. S., Lynn, M., Macpherson, D. A. and Zietz, E. N. (2005a). The Value of Housing Characteristics: A Meta Analysis. Presented at Mid Year Meeting of the American Real Estate and Urban Economics Association. May 2005.
Sirmans, G. S., Macpherson, D. A. and Zietz, E. N. (2005b) The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature.* 13(1): 3-42.
Solar Energy Industries Association (SEIA) and GTM Research (GTM) (2011) U.S. Solar Market Insight - 2010 Year in Review. GTM Research (GTM) in Boston MA. Prepared for Solar Energy Industries Association (SEIA), Washington, DC.
SunPower (2008) New Homes With SunPower Solar Systems are Bright Spot in Market. Press Release Regarding Ryness Corporation Report. June 24, 2008
White, H. (1980) A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica.* 48(4): 817-838.
Wooldridge, J. (2009) Introductory Econometrics: A Modern Approach. South Western Cengage Learning. Mason, OH. ISBN# 0324660545.