INSTALLATION PRICE, DEGRADATION, AND VALUE IN RESIDENTIAL PHOTOVOLTAIC SYSTEMS

by

Andrew Lick

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Environment and Resources: Nelson Institute for Environmental Studies

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

La Follette School Working Paper No. 2017-003
http://www.lafollette.wisc.edu/research/publications

October 2017

Abstract

Solar photovoltaic (PV) systems are becoming price competitive with conventional electricity sources. Their adoption is predicated on both private (electricity cost savings) and public (climate and air quality) benefits, which are obscured by wide variation in PV system price. System quality may be an important source of that variation, but it remains poorly understood. Here, I used degradation as a proxy for system quality, and studied degradation of small (15<kW) California Solar Initiative (CSI) systems to explore the hypothesis that high price reflects high quality. I analyzed data for 386 mature systems generated by the Expected Performance Based Buydown (EPBB) portion of the May 2016 CSI Working Dataset, the National Solar Radiation Database (NSRDB), and the Tracking the Sun (TTS) dataset. These systems showed a median annual degradation rate of 1.0% based on year-on-year (YOY) differencing. Using multiple linear regression, I found no support for the hypothesis that high-cost residential solar PV systems avoid annual degradation differently than low-cost systems. In general, the model explains little variation in the data, likely due to either data quality issues in the components of the degradation rate calculation and/or to significant but unexplored variables. Additionally, by estimating the value of a PV system with median annual degradation relative to one with no degradation, I demonstrated that the value of degradation represents a non-trivial cost to system owners. Despite a large range (+32% to 0%), median degradation adds 11% to the $/kWh cost of residential solar. These results demonstrate that policy interventions targeting degradation are an important area for transparency and financial risk reduction in residential PV markets.

Acronyms and Abbreviations

AC	Alternating Current
C	Celsius
CSI	California Solar Initiative

DC	Direct Current
DHI	Diffuse Horizontal Irradiance
DNI	Direct Normal Irradiance
EPBB	Expected Performance Based Buydown
GHI	Global Horizontal Irradiance
HIT	Heterojunction with Intrinsic Thin layer silicon
m	Meter
Mono-C-Si	Monocrystalline silicon
Multi-C-Si	Multi crystalline silicon
km	Kilometer
kW	Kilowatt
kWh	Kilowatt hour
LCOE	Levelized cost of energy
NSRDB	National Solar Radiation Database
OLS	Ordinary Least Squares
PG&E	Pacific Gas and Electric
POA	Plane of Array Irradiance
PR	Performance Ratio
PV	Photovoltaic
SCE	Southern California Edison
SDG&E	San Diego Gas and Electric
STC	Standard Test Conditions
TTS	Tracking the Sun dataset
W	Watt
YOY	Year-on-year method of calculating degradation

Contents

List of Figures

List of Figures: Appendix A

List of Figures: Appendix B

List of Tables

List of Tables: Appendix A

Introduction

Solar photovoltaic (PV[1]) cells convert sunlight directly into electricity. Today PV makes up only about 0.9% of U.S. electricity (U.S. Energy Information Administration 2017) but it is a potentially important source of low-emission supply. Solar energy has the technical potential to supply 69% of all electricity and 35% of the total energy in the U.S. PV by 2050 (Fthenakis, Mason, and Zweibel 2009), representing an important component in a transition to a low-emission electricity sector, especially given that global energy consumption is predicted to rise 48% by 2040 (U.S. Energy Information Administration 2016). Any transition to a low-emission electricity system requires scalable technology and competitive price.

Up from 3.1 gigawatts (GW) in 2015, the U.S. added 9.5 GW in 2016 making solar the number one source of new electric generation brought online (U.S. Energy Information Administration 2016) while sustained growth has exceeded 40 percent per year for most of the last decade (Kann et al. 2013). Historically, each doubling of deployment has resulted in 20 percent module price declines (Swanson 2006) and from 2009 to 2013 prices fell roughly 50 percent (Barbose et al. 2015)**.** However, as more systems are deployed for longer, reliability plays a larger role for public opinion and financial risk. Therefore, the interplay between price and reliability is critical to solar PV’s continued expansion as a source of low-emissions energy. The duration of PV system production aligning with the expectations of customers represents an important challenge for the industry.

A common saying among electricians who install solar PV systems is that “I don’t have enough money to buy cheap tools.” No great leap requires one to conclude that a

[1] Here and throughout I used PV as a shorthand for “rooftop” solar, which generally refers to PV systems installed on residential roofs and correspond to small residential PV applications.

relationship between the cost of components and the quality of the work and design in a PV system follows the same logic; building a robust system is somewhat at odds with building a low-cost system because it requires high-quality inputs and skilled labor. The solar industry has limited experience with PV systems deployed for long periods of time, making it easy to overlook the role reliability and particularly, the role degradation plays in the viability and value of a PV system. This work brings empirical evidence to the oft-asked question: “Do residential PV customers get what they pay for?”

To further illustrate the point, consider a hypothetical example. Two solar PV systems of similar size and purchase price, both installed in the same year in adjacent neighborhoods in Oakland, California. The installations are only blocks apart, where they have a similar solar resource, the same financial incentives, the same installers operate in the area, the houses are about the same age and have similar load profiles, even the roofing materials are of similar age and material. The main difference is that one system has been generating energy close to its rated capacity continuously for six years while the other has been dogged with problems that have reduced energy output. Although degradation connotes a slow decline in power this analysis defines degradation as explicitly inclusive of all factors that lead to loss of energy because performance and value are the results of both slow decline and catastrophic failure.

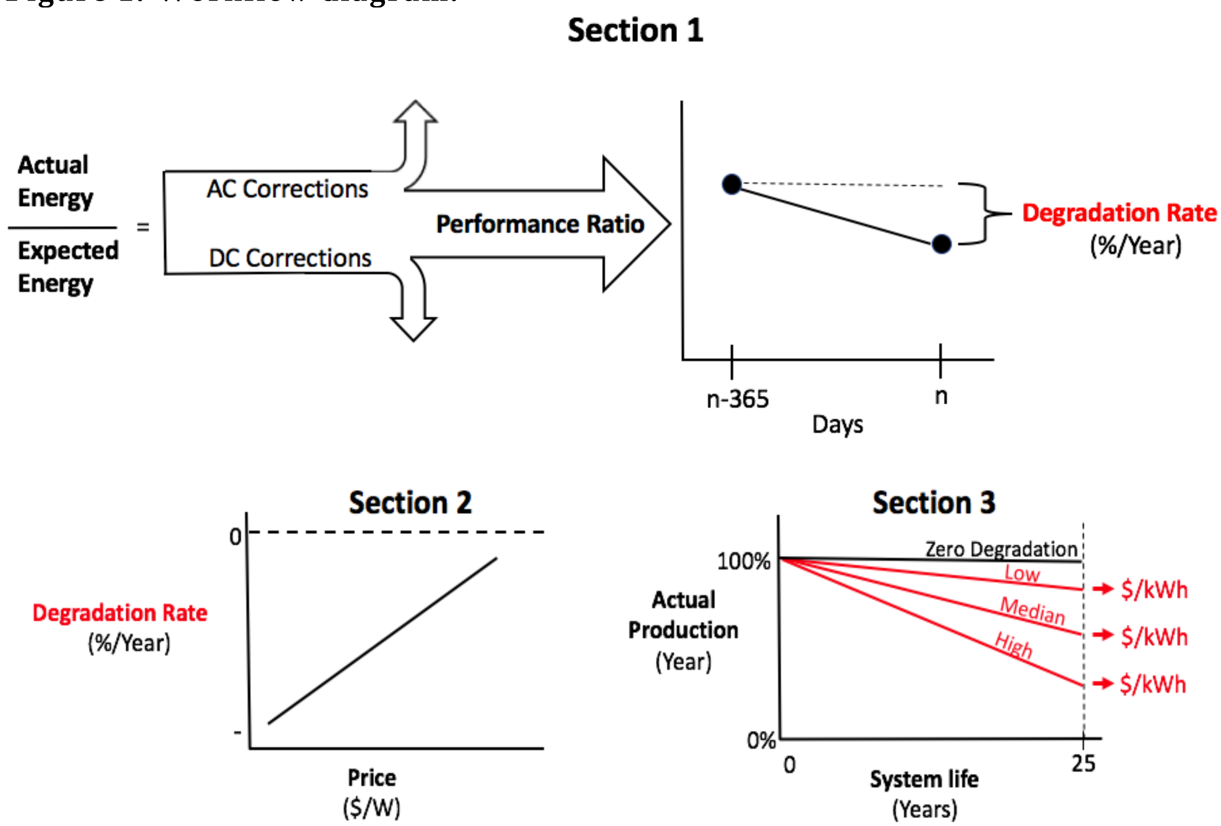
I measured reliability as degradation and show the results regarding energy delivery while controlling for intervening weather factors begging the questions: Were the quality of the installations different? Do the differences result from the quality of the components? Zooming out of the neighborhood to the state and national scale, policy makers are keenly aware that there is a difference in price, but to what extent can one predict if low-priced systems result from lower-quality components installed using sloppy techniques? This example encompasses the research question of this paper: **Do higher priced residential solar PV systems more reliably generate more energy over their lifetimes and thus represent more value to customers?**

In system hardware and installation, PV systems contain the attributes of a durable good and a service where the service constitutes over half of the system costs (Nemet et al. 2017). Since the quality of the service affects reliability, understanding degradation of whole PV systems is central to wider market penetration and the perception of financial risk. Consumer benefits are predicated on the performance of the investment and falling prices encourage consumption. Commonly, efficiency is thought to represent one of the major barriers to falling prices but while efficiency improvements are necessary, performance may be of equal importance (Woodhouse et al. 2016; Jones-Albertus et al. 2015). A focus on the relationship between the degradation and system price builds on previous work (Nemet et al. 2017) that identifies drivers of price variation in PV systems while questioning the magnitude of variation in installation quality. Although not treated explicitely in this exploratory analysis, this is further complicated by the indsutry-wide falling prices. The trade-off between price and value is so far understudied in residential solar PV.

This study makes two main contributions to solar degradation studies; it increases knowledge about degradation rates and the use of degradation rates to describe the residential PV market in California, by far the nation's largest. First, since it is impossible to comprehensively simulate degradation mechanisms that naturally occur in the field through accelerated indoor testing (Phinikarides et al. 2014), estimated degradation rates for field deployed systems in themselves further the body of knowledge of the reliability of PV systems. It also builds accuracy onto a previous degradation study of CSI PV systems (Itron 2012) by including more granular irradiance data specific to within 4km of each system. This study takes advantage of systems that were installed 6 to 10 years ago and continuously monitored for 3 to 5 years. Second, it combines performance, weather, and market data to apply degradation as a proxy for system quality which may be an important source of price variation.

Central to this paper is an investigation of the relationship of system price and degradation rate to assess whether low-priced systems are associated with higher degradation. Degradation rates quantifying the performance of entire systems also facilitate estimating the value of avoided degradation. The remainder of this paper is organized as follows. Section 1 estimates degradation rates. Section 2 uses multiple linear regression to estimate the magnitude of the association between installation price and the degradation observed in individual systems and across all systems. Section 3 shows the value of a high, median, and low degradation system relative to a system with none and shows the lost revenue over a system’s lifetime. Figure 1 describes the conceptual organization of this work from Section 1 through Section 3.

Figure 1. Workflow diagram.



Throughout Figure 1 red text indicates the role degradation rates play in this analysis. The Section 1 portion of Figure 1 shows the creation of performance ratios (PR), which consist of AC power measured on-site in the numerator over the energy expected given the weather conditions of the same system in the denominator. Data filters or “AC Corrections” eliminate erroneous AC measurements and are explained below. Expected energy is the nameplate DC rating of the system at hourly intervals with weather corrected data shown as “DC Corrections”, also explained below. A degradation rate is calculated by subtracting a PR in year-one from a PR in year-two PR; a difference of 365 days. The degradation rate is output from Section 1 and input into Section 2 use as a dependent variable to investigate the degradation system price relationship. That relationship is captured as a positively sloped line in the bottom left and is the central topic in Section 2. Section 3 uses degradation rates to generate differences in production caused by degradation. The area above each wedge of degradation represents additional cost per kWh and reduced revenue from lost kWh produced over the life of a PV system.

Section 1: Degradation

Solar PV Degradation

A degradation rate is a quantified decline in power output over time defined by a rate of change over time. Degradation implies decline so is a negative rate of change over time. On the other hand, negative degradation is a positive rate of change resulting in increasing power output over time for a given PV system. To avoid confusion regarding the signs involved, I refer to a negative rate of change in performance as faster degradation and a positive rate of change in performance over time as slower degradation.

There are two general types of solar PV degradation studies: whole system degradation and equipment degradation. To date, most degradation studies are either a mixture of whole system and equipment degradation rates (Dirk C Jordan and Kurtz 2015), system-level for single solar PV systems (Kymakis, Kalykakis, and Papazoglou 2009), exclusively equipment degradation rates (Sharma and Chandel 2013), single and multiple-measurement degradation rate studies (Dirk C. Jordan et al. 2016) with only a single known multi-year, multiple system study using the year-on-year (YOY) method (Anderson, Defreitas, and Hasselbrink, Jr. 2013). The most comprehensive review to date encompasses nearly 11,000 degradation rates in 200 studies from 40 countries (Dirk C. Jordan et al. 2016) and includes both module and system-level degradation.

This study focuses on whole system degradation, defined as the overall change in system performance over time including module and inverter degradation, soiling[2], shading, and inadequate maintenance, and explicitly includes system outages because their inclusion helps identify some of the performance variation caused by system malfunctions from equipment and poor installations practices. The variability in solar resource must be cleanly separated from losses from component deterioration, malfunction, and installation quality to assess performance through degradation.

Data

Degradation rate estimation utilizes data from individual system characteristics, alternating current (AC) production, irradiance, ambient temperature, wind speed, solar zenith angle, and system location. System characteristics come from publicly available California Solar Initiative[3] (CSI) data, based on installer provided details. Production data is also publicly available from metered generation measurements of 504 CSI systems starting in 2010 and ending in 2016 and time stamped in fifteen-minute intervals. Systems come from the three largest utility service areas in California Pacific Gas & Electric, Southern California Edison, and San Diego Gas & Electric. Irradiance, ambient temperature, wind speed, angle of incidence, and solar zenith angle data come from the National Solar Radiation Database[4] (NSRDB) Physical Solar Model. This database has a spatial resolution of 4x4 kilometers, at half-hour intervals, and includes years 1998 to 2014. Additional system characteristics and

---

[2] Soiling is the deposition of fine materials on the surface of the PV modules that inhibit incident radiation.
[3] Available from: http://www.californiadgstats.ca.gov/
[4] Available from: https://nsrdb.nrel.gov/

system location are available from the Tracking the Sun[5] (TTS) dataset, updated to 2014. These data sources facilitate analysis throughout Sections 1, 2, and 3.

Extraction of irradiance data from the NSRDB relies on the latitude and longitude of each system. The location of each system has a maximum of 3 km distance from the exact location of the NSRDB grid point. This should increase the accuracy of the performance ratios (PRs) relative to the previous degradation rate calculation (Itron 2012) that used zip code level irradiance data. Data from the NSRDB includes global horizontal (GHI), direct normal (DNI), and diffuse horizontal irradiance (DHI), ambient temperature (°C), wind speed (at a height of 10m), and solar zenith angle (degrees).

All production data represents systems incented by the Energy and Performance Based Buydown (EPBB) portion of the CSI rebate program. Of the 504 systems measured, I analyzed 386. I discarded systems with incomplete data by visually inspecting time-series plots (See Appendix A, Examples of Discarded Systems for more detail) and other criteria outlined in Table 1. Along with AC system production data, the time-series dataset represents 5 years of performance evaluation (2010-2014). To maintain time-series equivalence between AC production and NSRDB data, all time stamps were averaged to hourly intervals.

Various guides exist for acceptable ranges of data to calculate PRs and degradation rates. Table 1 filters screen data that would lead to erroneous PR values and degradation rates. I eliminated solar zenith angles greater than 90 degrees, which indicate nighttime data and are eliminated along with shifted data revealed through time-series plots. A minimum irradiance limit of 100 W/m^2 ensures low light conditions from anything nearby that could act as a shade structure do not affect rates. This filter does not, however, account for changing shade conditions over time. For instance, a system may not suffer from shade in

[5] Available from: https://openpv.nrel.gov/

certain seasons from the presence of nearby deciduous trees and their leaves or a tree may grow larger and shade later in the life of the system. Metered system production is limited to the sum of the inverter maximum power ratings as any system producing more than is feasible based on inverter output capacity is erroneous. I discarded any rate showing a yearly change greater than +/-100 percent because I considered the doubling of a PR from the previous year before to represent an aberration in the data. I corrected spells of negative AC production where they mirrored positive values but in the negative direction. A minimum limit of 0 kilowatts (kW) AC excludes negative observations occurring randomly. Production data from CSI carries flags and those indicating erroneous data are discarded. Data points affected by inverter clipping[6] can bias degradation rates and are excluded by discarding the top 2% of production from each system and dropping any observations indicating power clipping where the inverter operates at maximum power output. To limit bias in the combined degradation rate across all systems, I removed any year of data with less than 2000 observations. The final sample includes exclusively fixed-tilt, grid-tied systems. Table 1 lists all filters, their ranges, and the amount of data discarded.

[6] Inverter clipping is a situation in which the ratio of DC to AC in a system design is greater than one. On a cool sunny day, the modules deliver more power than can be inverted resulting in an artificially low PR.

Table 1. Data filters for limiting erroneous PRs.

Filter	Values Eliminated	Observations Eliminated
Nighttime	Solar Zenith Angle > 90°	7,013,774
Visually Inspected	Invalid due to shifts, portions missing, etc.	250,789
Irradiance	POA* < 100 W/m2	982,199
AC power	kW > Maximum Rated AC Capacity	59,390
Nameplate Rating	kW > Maximum Rated DC Capacity	602
High/Low Degradation	A rate is eliminated +/-100% from the previous year	2,060,819
Negative Production	kW < 0	91,130
Invalid Data	Flagged as Invalid	1,924
Inverter Clipping	Eliminate top 2% of kW from each system	111,479
Insufficient Data	Discard years with fewer than 2000 observations	497,887
Total		11,069,993

*POA stands for plane-of-irradiance, which is a plane perpendicular to incident radiation.

Methods

While there are many estimation methods for degradation rates of systems, I selected the year-on-year (YOY) change in PR because it requires only AC production, weather, and system capacity data. The YOY change differences each observation of PR if a data point is available for the same day one year earlier. The difference is calculated by subtracting the year-one data point from the year-two data point, repeated for each hour of the year. Five years of data generate four degradation rates data points for each hour of the year for 386 systems before filtering. Because system measurements started at different times in 2010 for different systems with few measurements relative to subsequent years, after filtering, no systems achieved more than 2000 measurements in 2010. The result is that no system has a degradation data point for 2011 and instead of 4 data points for each hour of the year systems have a maximum of three and a minimum of one. Therefore, because no system here was monitored starting at installation, these systems can be thought of as mature systems as opposed to young systems. From these, I calculated a median degradation rate from a distribution of all systems counted together and a median rate from distributions associated with individual systems, where each system is counted only once.

A similar study of 445 PV systems (Anderson, Defreitas, and Hasselbrink, Jr. 2013) used the same approach making results comparable across studies. One shortcoming of the YOY approach is that it does not explicitly capture soiling. However, the YOY approach is accurate to the extent that soiling is seasonal (Maghami et al. 2016). The rain patterns in the climate zones represented in this analysis are highly seasonal (Department of Water Resources, State of California 2014) lending validity to this approach.

The performance ratio is a dimensionless ratio of observed to expected PV system output at a point in time normalized to irradiance and system size and quantifies the overall effect of losses from inverter inefficiency, wiring, cell and module mismatch, incomplete conversion of irradiance, soiling, snow, downtime, shading, and component failures and all other losses when converting from DC to AC power (Marion et al. 2005). The PR is for use in performance guarantees and contractual agreements because it holistically captures performance. An additional benefit of the PR is that it is especially useful for comparing the performance of existing systems (Sengupta et al. 2015) and is used in almost a third of all published degradation rates facilitating comparisons between studies (Jordan and Kurtz 2013).

For PR calculations, I transposed irradiance into the plane-of-array (POA) based on the azimuth and tilt of each system. Together GHI, DNI, and DHI, along with solar zenith angles, provide the components necessary to calculate the total amount of irradiance incident on a flat plate solar module. Among the various transposition model, the Hay/Davies (Hay 1979) and Perez (Perez et al. 1990) are the most popular and are seen as an industry standard (Hansen 2015). The difference in transposition models is in their treatment of DHI. The Hay/Davies model separates DHI into two separate components, the circumsolar diffuse and a rest-of-sky diffuse irradiance. The Perez model differs by adding a third relative to Hay/Davies, the near-horizon component. See Appendix A, Plane of Array Radiation for more details. Since previous degradation evaluations of CSI (Itron 2012) used the Perez model and it has the smallest associated errors at different locations across the continental U.S. (Lave 2015), I used it to calculate POA irradiance values for use in the PR.

However, using POA with the exclusion of temperature and wind speed can result in PR calculations that suffer from bias due to the effects of local weather conditions at a PV site (Dierauf et al. 2013). Irradiance is the major driver of system production while ambient temperature, through a series of transformations to PV module cell temperature, negatively affect system production. Therefore, I used a weather-corrected PR to minimize any bias in the resulting system degradation estimates. Equation 1 defines the weather-corrected PR as in Dierauf et al. (2013). The PR is independent of location and system size. Equation 1 shows how the weather-corrected PR controls for system size in the denominator by multiplying the nameplate system rating by the available POA irradiance at hourly time stamp *i*.

$$PR_{i,j} = \frac{\sum kW_{AC_{i,j}}}{\sum kW_{STC\_j} \left(\frac{POA_{i,j}}{G_{STC}}\right)\left(1 - \frac{\delta_j}{100}\left(T_{cell\_typ\_avg} - T_{cell_{i,j}}\right)\right)} \qquad \textbf{(1)}$$

Where:

$PR_{i,j}$ = Temperature-corrected performance ratio for system at weighted average for hour *i*, system *j*
$kW_{AC\_i,j}$ = Metered AC electrical generation for hour *i*, system *j* (kW)
kW_{STC_j} = System power rating at STC, system *j* (kW)
$POA_{i,j}$ = Irradiance measured plane of array for hour *i*, system *j* (kW/m^2)
G_{STC} = Irradiance at standard test conditions (STC) (1,000 W/m^2)
$T_{cell\_i,j}$ = Cell temperature computed from meteorological data for hour *i*, system *j* (°C)
$T_{cell\_typ\_avg}$ = Average cell temperature from current year of weather data (°C)
δ_j = Temperature coefficient for power (%/°C) that corresponds to the installed modules.
The summations are defined for hourly time steps[7].

Equation 1 shows irradiance ($POA_{i,j}$) divided by irradiance at standard test conditions (G_{STC}). See Appendix A, Standard Test Conditions for more detail. Subtracting module cell temperature from the average annual irradiance-weighted cell temperature creates a cell temperature difference multiplied by the temperature coefficient for power, represented as a percent loss or gain depending on whether the cell temperature is higher or lower than the annual average cell temperature. Applying cell temperature corrections relies on PV array mounting. Some arrays are mounted on rails attached to the roof via stanchions holding them off the roof, close enough to the roof that heat dissipation can be altered by radiant heat from the structure. Others are pole-mounted in basically any tilt and not adjacent to large thermal masses. Installation location and the composition of the modules allows for the wind to dissipate heat from modules. The data do not include specific mounting techniques so I chose the most conservative convective heat transfer coefficients available for estimating module

---

[7] Equation 1 comes from Dierauf et al. (2013), page 3.

cell temperature (King et al. 2004). See Appendix A, Convective Heat Transfer Equations and Coefficients for more detail. The effect of wind and sunlight on module cell temperature and therefore system output are corrections in the denominator of Equation 1 allowing for PRs with actual over normalized AC output for comparison across systems and time.

Results

The columns in Table 2 represent the 386 solar PV systems used in the analysis showing descriptive statistics of the overall degradation results. The column headed “All systems” shows a single median degradation rate across the entire fleet of systems in the dataset using the YOY approach with hourly data. Also using YOY, the column headed “Individual systems” shows the result of calculating a median degradation rate for each system, indicated by drastically fewer observations.

I calculated the median degradation rates and their standard errors by sampling with replacement (bootstrap) with 1000 repetitions at the 95% confidence level. The median values are both significantly different from zero at the 95% confidence level.

Table 2. Descriptive statistics for degradation rates (%/year)

Statistic	Pooled hours of all systems	Individual systems
Mean	-1.4	-1.0
Standard Deviation	27	1.3
Median Degradation Rate	-1.0	-0.9
Standard Error of the Median	0.00	0.05
95% Confidence Interval	+/- 0.01	+/- 0.1
Minimum Degradation Rate	-100	-8.8
Maximum Degradation Rate	100	7.8
Number of Observations	2,878,039	386

Using pooled hours of all systems to calculate a fleet degradation rate for the sample of residential systems in the CSI EPBB rebate program results in a degradation rate of 1.0%/year. Figure 2 shows this distribution, which uses all 2,878,039 hourly observations in the dataset. The histogram has color-coded vertical bars showing the variation in degradation rates from different years around a black vertical line at the fleet median. Also, shown in Figure 2 are the wide tails of the degradation rates distribution that end at +/-100 percent. The distribution also shows a roughly normal distribution of rates on the positive and negative side of zero with the bulk concentrated near zero. The slightly negative median is of similar magnitude to rates in the literature (Dirk C. Jordan et al. 2016).

Figure 2. Year-on-Year with Hourly Data: All Systems

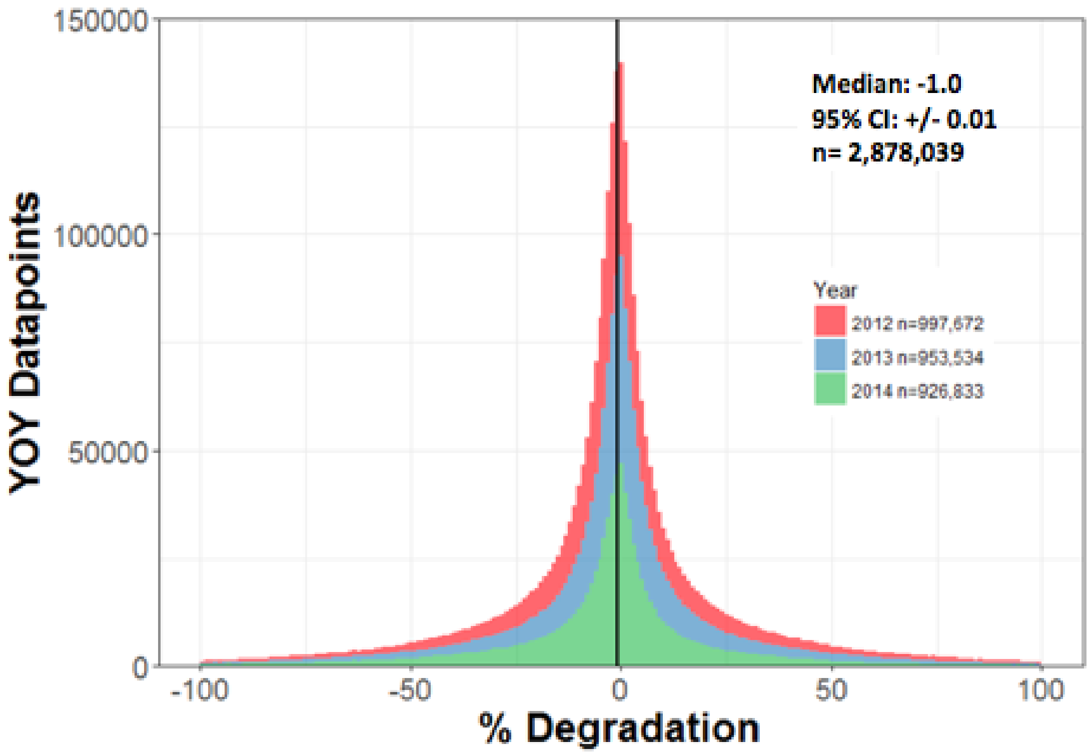


Figure 2 shows that degradation rate measurements calculated from 2011-2012, 2012-2013, and 2013-2014 appear to vary symmetrically around the median across all years. However, important to discerning if there are systematic differences from degradation rates in different years, are the results of three Wilcoxon signed-rank tests to determine if the median degradation rates from different years are statistically different from each other. Summary statistics in Table 3 show differences in both the median and mean degradation rates across year but with similar amounts of variation as evidenced by the standard deviations and similar confidence intervals.

Table 3. Summary statistics for degradation rates across years

Year	Median (%/year)	Mean (%/year)	Std. Dev.	95% CI
2012	-1.4	-1.4	27	+/- 0.03
2013	-0.3	-1.3	27	+/- 0.03
2014	-1.2	-1.7	26	+/- 0.03

Table 4 shows the results of the Wilcoxon signed-rank tests for the median of each year's degradation rate compared to the other years. The null hypothesis is that the difference between medians is zero. The test result column indicates that between 2012 and 2013 the probability that the observed test statistic would be as extreme as observed is approximately zero (p-value = 0.00) if the difference between the medians was zero. The result is identical for other combinations of years. Therefore, a statistically significant difference between the median degradation rates exists in 2012, 2013, and 2014 compared to each other.

Table 4. Test results for significant difference between degradation medians for years 2012, 2013, and 2014

Difference Between Years	Test Result (p-value)	Statistically Significant
2012 – 2013	0.00	Yes
2013 – 2014	0.00	Yes
2012 - 2014	0.00	Yes

While a full analysis is beyond the scope of this inquiry, the second column of Table 3, median degradation, shows a large difference between 2013 and 2012 as well as 2013 and 2014. I investigated the both the association of geography by utility service territory and rain across months to explain the difference. Table 5 shows median degradation rates by year for each utility—Pacific Gas and Electric(PG&E), Southern California Edison (SCE), and San Diego Gas and Electric (SDG&E)—in the dataset. Here, 2013 also shows less degradation across utilities but is most pronounced in the SDG&E systems. The year 2013 is anomalous especially for systems in the SDG&E service territory.

Table 5. Median annual degradation by utility

Year	PGE Median (%/year)	SCE Median (%/year)	SDG&E Median (%/year)
2012	-1.47	-1.36	-1.45
2013	-0.35	-0.47	0.02
2014	-1.13	-1.27	-1.12
Mean	-0.98	-1.03	-0.85

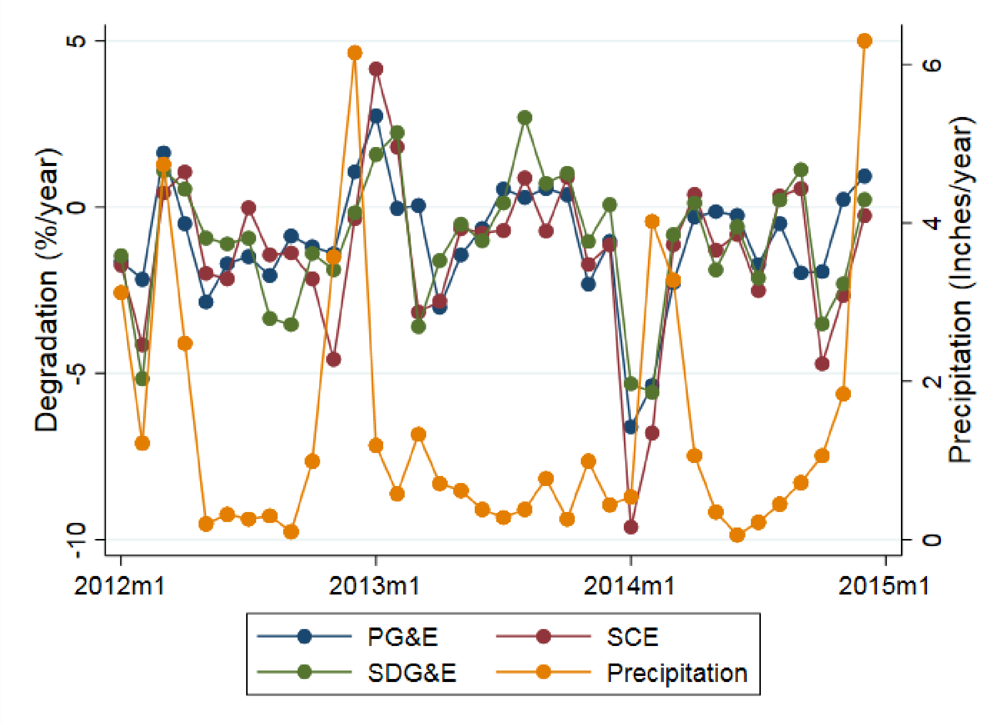
As previously mentioned, using the YOY differencing approach to calculate degradation accounts for soiling where it is roughly seasonally equivalent. Where it is not, soiling introduces unaccounted variation into the degradation rates. The interplay between rain and soiling is therefore a likely candidate for a source of the low degradation show in Table 3 and Table 5 for 2013. Rain cleans the surface of the modules making incident radiation a more effective driver of the photovoltaic effect. It impacts PV system performance loss as a function of both the amount and frequency of rainfall (Kimber et al. 2006), primarily dependent upon the time since the previous rainfall (Mejia and Kleissel 2013). While a full analysis of soiling is possible with established techniques (Deceglie et al. 2016) it remains beyond the scope of this report. Therefore, to offer a preliminary explanation of the results in Table 3 and Table 5, I combined monthly rain data from the

National Oceanic and Atmospheric Administration (NOAA)[8] and plotted mean monthly rain and degradation rates from for each utility in the dataset from 2012-2014.

The gold-colored line and dots in Figure 3 represent the monthly rain pattern in inches per month across California on the right vertical axis. The blue, green, and red show monthly median degradation rates for the three utility service areas in this analysis on the left vertical axis. Three characteristics stand out. First, degradation between utility areas follows roughly similar patterns across the time series with the more southern utilities, SCE and SDG&E showing slightly more monthly volatility. Second, 2013 shows a more prolonged period or low-rain interspersed with more small spikes compared to 2012 or 2014. At the monthly level, this data appears to explain some of the variation in the lower degradation due to more frequent cleaning of soiled modules. However, this does not differentiate monthly rain patterns according to the geography of the utility service area, which would be important for future study.

[8] For more information see: https://www.ncdc.noaa.gov/

Figure 3. Monthly degradation and rain in California

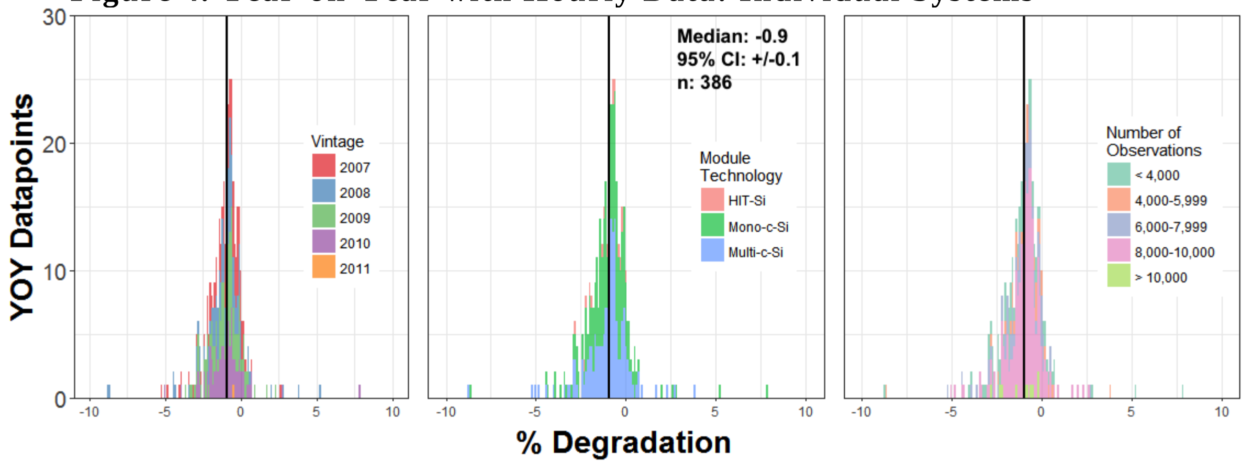


Third, there is a spike in the positive direction for all three utilities near January 2013 with a similar occurrence although more pronounced and in the negative direction in January 2014. Degradation rates calculated with the YOY approach measure the difference in PR from one year to the next so a large increase in rates (less degradation) in January 2013 has a large and opposite impact on January 2014. Given that soiling depends on the time since previous rainfall, the large rain month that spikes in the positive direction in December of 2013 is likely to have a large impact on the performance in the subsequent month, January 2013. This plot supports that explanation. Similarly, the negative spike in rates in January 2014 may be explained by the lack of a large-rain month preceding it—shown by the increase in rain following instead of preceding January 2014. Therefore, the large rain month preceding January 2013 may have caused both the low degradation in that month and the lack of a large rain month preceding January 2014 may have caused in high degradation in January 2014.

A full analysis requires more granularity in both temporal and geographic rain data to explain the difference in years and utility service areas from soiling. However, monthly rain data combined with the lessons from previous research suggest an effect of rain on soiling and degradation at the monthly level, especially because they are not controlled for otherwise. Rain events interspersed across 2013 and a large rain month in December of 2013 with no such rain month in December 2014 offer a plausible yet preliminary explanation for less observed degradation in 2013 relative to 2012 and 2014 and the most pronounced degradation spikes in Figure 3, respectively. As yearly degradation differences are important to investigate, so are the characteristics for systems and their distributions.

Using the same hourly time-series AC production data, Figure 4 shows distributions of degradation rates for 386 individual systems instead of all systems together. Here, each system's median is counted only once. Rates range from -8.8 to 7.8 percent per year with a vertical black line indicating the median of -0.9 percent per year. The standard error of the median at the 95% confidence interval is 0.05, showing that the uncertainty in this estimate is small relative to the

Figure 4. Year-on-Year with Hourly Data: Individual Systems



magnitude of the median. However, there is a large spread in the degradation rates. While not necessarily a physical improbability, negative degradation rates (less degradation or a positive rate of change in performance) manifest from uncertainty in the rate calculation. For example, in each system that operates continuously for three years, a system that underperforms between year one and year two, reverting to the mean or even above the mean in a subsequent year, generates slopes with different signs—some positive and some negative. Depending on the magnitude of those slopes, the positive slopes between year two and year three could result in a positive median over the range of the time-series dataset. Given the results in Table 4, this is a distinct possibility because 2013 rates are significantly lower than 2014 rates and where the increase from 2013 to 2014 is larger than the decrease from 2012 to 2013 negative degradation can occur. There are 27 individual systems with median degradation rates that have a positive rate of change over time, representing roughly 7% of the systems in this sample. See Appendix A, Negative Degradation for more detail.

Compared to Figure 2, in which a fleet distribution rather than a distribution of individual systems is shown, Figure 4 degradation rate medians for individual systems are considerably less dense on the positive side of zero. This distribution has 3 panels of degradation rates at the system level overlaid with characteristics of the systems. The left panel shows how vintage, or year of system installation is distributed, the middle panel breaks out the same distribution by module technology, and the right panel by number of observations occurring at different areas in the distribution.

The left panel shows no obvious systematic differences in degradation by vintage. Even though early life degradation data does not exist for most systems and Table 6 shows that almost 80% of systems are of vintage 2007 to 2009, two points stand out about the middle panel. First, as in Table 6, most of the systems in this dataset have either monocrystalline (Mono-c-Si) or multicrystalline (Multi-c-Si) and many fewer with Heterojunction with Intrinsic Thin Layer (HIT-Si) silicon modules. Second, the degradation rates for the three different module types represented in the data similarly show no obvious pattern for degradation associated with module technology. The right panel shows the distribution of the number of observations around the median degradation rate for the individual systems. Keeping in mind that a year of data was discarded when a system had fewer than 2,000 observations per year, this panel shows no obvious pattern of the density of observations for either less or more degradation. Across the panels, vintage module technology, and the number of observations for a given system appear not to map onto the degradation rate distribution at the individual system level in any systematic way.

Table 6. Number of occurrences of the vintage, module technology and number of observations for individual system distribution

Vintage	N	Module Technology	N	Number of Observations	N
2007	96	HIT-Si	17	< 4,000	65
2008	97	Mono-c-Si	199	4,000 – 5,999	28
2009	106	Multi-c-Si	170	6,000 – 7,999	69
2010	86			8,000 – 9,999	214
2011	1			> 10,000	10
Total	**386**		**386**		**386**

It is likely that component and workmanship failures occur more in the early years of a system's life, get discovered and repaired, and then continue working properly. Investigation into this possibility is difficult because most of the systems in the dataset were installed before monitoring began in 2010 and 2011. Moreover, because so few observations were collected during the first year of monitoring in 2010, no degradation rates were calculated for 2011 and do not appear in Figure 4. Future study could investigate the potential for non-linear degradation and high early-life degradation in this data. Additionally, light-induced degradation during the first days of installation and other sources of beginning-of-life degradation have been observed but this also remains outside the scope of this study.

In sum, degradation rate estimates based on this sample of residential PV systems incented through the CSI show at their medians, increasing degradation that is statistically different from zero at the 95 percent confidence level with statistically significant differences between years. As above, a full analysis of the source of the yearly difference is beyond the scope of this work but different amounts of soiling due to different rain patterns across years is a likely source of this difference that would not be captured by the YOY approach. Vintage, module technology, and the number of observations per system appear broadly distributed across degradation rates by system. These estimates comprise the dependent variables in the analyses in Section 2.

Section 2: Price and Degradation

Price-Degradation Relationship

Referring to the workflow in Figure 1, Section 2 estimates how the relationship of system price scales with degradation rates calculated in Section 1. Continuing with the hypothetical example of two residential PV systems in Oakland, CA, much like other situations where parties engage in contracts with specialized information, understood to different degrees by the parties involved, thoroughly understanding the underlying drivers of installation cost is difficult for a typical residential PV consumer. Recent developments[9] make transparency more attainable but the quality of the installation and other unobservable factors remain hidden. Degradation of whole systems includes quality differences that are difficult for consumers to understand before or even after making a purchase. In purchasing solar PV for one's home, it remains difficult to know what one should pay when the final product, a standard kilowatt-hour (kWh), has a wide price range and degradation is known to vary.

Research on the heterogeneity of PV system price has highlighted the possibility that observable characteristics alone fail to explain the large variation in installed price (Gillingham et al. 2014). One uncovered unobservable factor is that evidence of increased installer experience can decrease the non-hardware portion of system cost (Bollinger and Gillingham 2014). Therefore, at the very least, PV price reflects experience in addition to the cost and quality of system components. There are still other likely unobserved characteristics driving PV system prices. For instance, even when controlling for the effects of shading, array orientation, outages, and cloud cover, power production within a group of 80 systems using the same inverters and modules varied more than that observed between groups of systems with different modules (Lonij et al. 2012). Therefore, unobservable factors influencing system production not reflected in equipment may be significant sources of variation. Although shading, changing shade conditions, and soiling at each PV site remain

[9] For more information see: https://www.energysage.com and https://pickmysolar.com.

uncontrolled, Section 1 detailed the degradation rates of 386 systems as a proxy for system quality to characterize its relationship with the price of PV systems.

Another reason to explore this relationship is that a better understanding of system quality is also critical for positive public opinion of PV systems. At least in part, solar PV adoption is the result of social influences called peer effects that rely on social approval instead of price alone (Rode and Weber 2011; Bollinger and Gillingham 2012; Noll, Dawes, and Rai 2014; Rai and Beck 2015). The connection between degradation and installation price is therefore important not only as a potential source of variation in system price but also for the credibility of PV as an emergent low-emission technology largely predicated on private economic benefits to system owners. The perception of private benefits leading to adoption requires credible performance. Reliable energy production is part of transparent prices and is important for potential PV customers. A primary aim of the analysis in this section is to uncover what relationship, if any, exists between installation price and degradation. The working hypothesis of this section is then:

Hypothesis
High-priced systems are associated with lower rates of degradation while low-priced systems are associated with higher rates of degradation.

Methods

I used two estimation approaches to capture the degradation response to installation price controlling for various factors. First, using the median hourly degradation rates for individual systems, I captured the relationship between the price of each system and degradation rates by assigning each system a price with a single degradation rate (n=386). Due to the cross-sectional nature of the data, the approach takes advantage of weather-corrected performance ratios (PRs) that help to minimize the uncertainty in the degradation rates. However, a tradeoff exists in using a smaller dataset. Therefore, a second approach uses the entirety of the data because a larger sample (n=2,878,039) can more easily detect an effect. I used the degradation rates with nonweather-corrected or simple PRs to show the response of price on degradation utilizing the entire set of hourly rates without distinguishing individual systems. This is important because of the possibility that medians hide uncertainty in the larger dataset. Although installation price is fixed for each system, this model introduces a fixed effect offering the ability to control for time variation in the price-degradation response. See Appendix B, Guass Markov Assumptions for more detailed model assumptions.

Individual Systems Model and Results

I modeled the relationship between degradation and installation price by fitting a linear equation to the observed data where the median of each system is counted once. The empirical specification uses a cross-section of individual systems regressing degradation rate (in percent per year) of each system on several covariates hypothesized to influence degradation. These variables are listed and described in Table 7. I modeled the degradation rate (R_{d_i}) for individual system i, in Equation 2.

$$R_{d\_i} = \boldsymbol{\beta_1 Pr_i} + \boldsymbol{\beta_2 Rat_i} + \boldsymbol{\beta_3 Eff_i} + \boldsymbol{\beta_4 TPO_i} + \boldsymbol{\beta_5 VOS_i} + \boldsymbol{\beta_6 Mtech_i} + \boldsymbol{\beta_7 Vin_i} + \epsilon_i \qquad \textbf{(2)}$$

Expressed as real 2014 dollars per watt before incentives calculated in the TTS dataset, the system installation price (Pr_i) is the variable of interest and can be interpreted as the effect of system price on degradation (R_{d_i}). The nameplate rating of the system (Rat_i) represents the PV system capacity based on standard test conditions in kW because larger systems in terms of DC capacity have higher costs. Module efficiency (*Eff_i*) is a proxy for the quality of the modules and is used in the industry as a heuristic for quality components helping to separate the effect of build quality. For example, SunPower modules are generally considered a premium product and show lower degradation than some other manufacturers (Anderson, Defreitas, and Hasselbrink, Jr. 2013). Although customer-owned systems do not lack the incentive to perform regular maintenance on PV systems they may lack both the knowledge to recognize a potential threat to prolonged performance or through limited experience, not realize that the system is underperforming. The difference in degradation between customer-owned and third-party ownership (TPO_i) likely entails owners with dedicated service personnel to perform maintenance without any homeowner action. Based on retail electricity rates and the amount of insolation for a given customer's PV system, the value of solar variable (VOS_i) is an estimate of the total value of all incentives and savings on utility bills. As a variable that captures the financial attractiveness of PV, customers in areas with a higher value of solar tend to see higher system prices (Gillingham et al. 2014). I also controlled for different module technologies ($Mtech_i$) because they have different degradation patterns and different costs. There are three module technologies in this dataset; Mono-c-Si, Multi-c-Si, and HIT-Si. Lastly is vintage **(**Vin_i) or the year in which the system was installed which controls for the declining cost of systems over time. Table 7 provides descriptive statistics for these variables. The TTS dataset includes information on labor, inverter, module, and balance of system costs but none were available for the systems I analyzed. Therefore, component cost cannot be controlled here.

Table 7. Descriptive statistics for individual system regression variables

Variable	Mean	SD	Min	Max	Description and units
Degradation Rate	-1.0	1.3	-8.8	7.8	Median degradation rate/system **(%/year)**
Installation Price	9.2	2.3	3.5	22	Purchase price of system **(2014 real $/W)**
Nameplate Rating	4.8	2.3	1.3	14	DC capacity at STC **(kW)**
Module Efficiency	0.15	0.02	0.12	0.19	Conversion efficiency **(%)**
Third Party Ownership	N/A	N/A	N/A	N/A	Third party owned versus customer owned **(categorical)**
Value of Solar	9.2	2.4	5.1	15	Combined value of incentives and utility bill savings **(real 2014 $/W)**
Module Technology	N/A	N/A	N/A	N/A	Mono-c-Si, Multi-c-Si, a-Si, HIT-Si **(categorical)**
Vintage	2008	1	2007	2011	Year of system construction
N= 386					

Table 8 shows the results from regressing degradation rate on the covariates specified in Equation 2. First, as evidenced by an R-squared value of 0.015 in model seven, almost all the variation in degradation remains unexplained by the model. Reasons include measurement error in the calculated degradation rates such as site conditions not captured by modeled irradiance, temperature, and wind data, partial or intermittent shading not captured by data filters, snow, and soiling. Although steps were taken to diminish the error in the estimated degradation rates by controlling for wind, and temperature the degradation rates used in the individual systems regression in Table 7 carry their own standard errors. Moreover, a trend exists in which larger rates both in the highly positive and highly negative direction show higher standard error than rates closer to zero; uncertainty surrounding rate medians likely plays a role. See Appendix B, Uncertainty in Degradation Rate Medians for more detail. Remembering that whole system degradation includes all aspects of the performance of a system, perhaps equally the installation practices as equipment, further research is required to demonstrate non-equipment degradation pathways for inclusion in models attempting to explain degradation. This data currently lacks those variables.

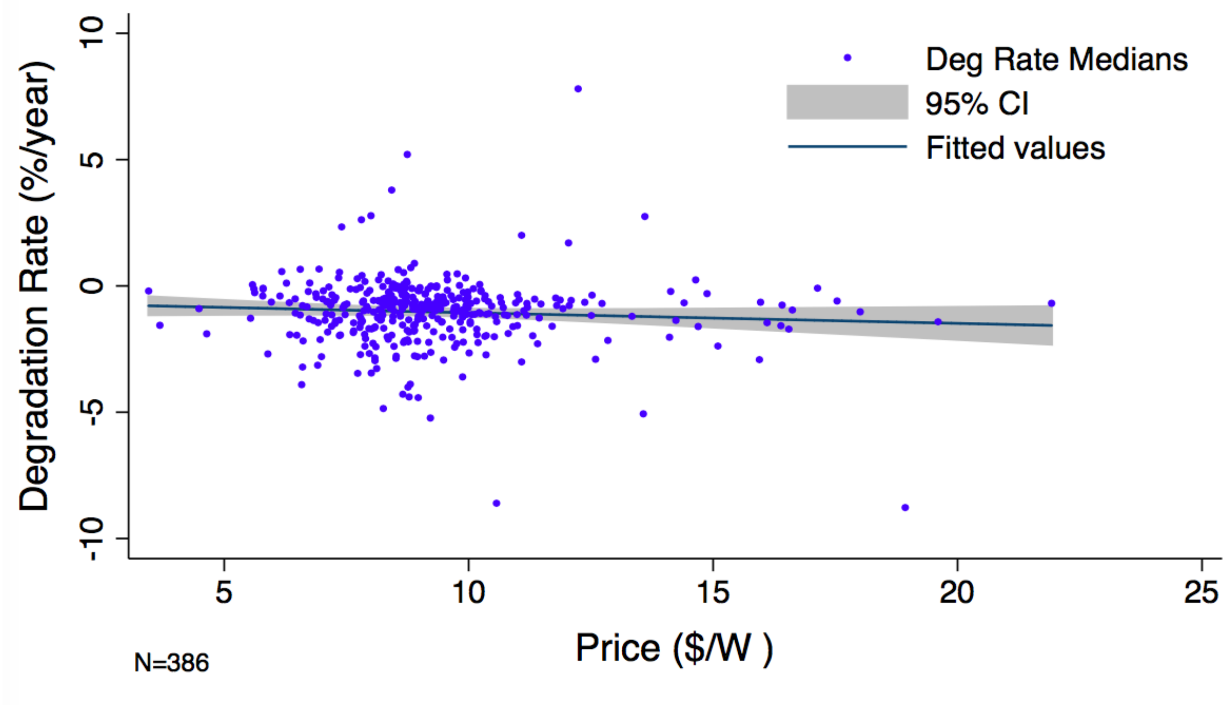
Table 8. Regressions results for individual systems

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Installation Price	**-0.0425**	**-0.0548**	**-0.0551**	**-0.0555**	**-0.0566**	**-0.0554**	**-0.0534**
	(0.0429)	**(0.0429)**	**(0.0435)**	**(0.0436)**	**(0.0433)**	**(0.0430)**	**(0.0441)**
Nameplate Rating		-0.0425	-0.0427	-0.0420	-0.0429	-0.0439	-0.0418
		(0.0311)	(0.0313)	(0.0312)	(0.0313)	(0.0312)	(0.0308)
Module Efficiency			0.305	0.156	0.204	-2.702	-2.945
			(3.159)	(3.163)	(3.176)	(4.794)	(4.890)
Third Party Owned				-0.0992	-0.129	-0.164	-0.170
				(0.221)	(0.231)	(0.238)	(0.240)
Value of Solar					-0.0162	-0.0156	-0.0199
					(0.0254)	(0.0255)	(0.0275)
Mono-c-Si						0.0800	0.118
						(0.214)	(0.227)
Multi-c-Si						-0.101	-0.0706
						(0.271)	(0.271)
Vintage							0.0390
							(0.0680)
Observations	386	386	386	386	386	386	386
R-squared	0.005	0.010	0.010	0.011	0.012	0.014	0.015

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8 shows that installation price on degradation is shown to have a slightly positive bias in the absence of nameplate rating. Module efficiency, third party ownership, and the value of solar all show similar effects on the magnitude of installation price but the addition of none of these covariates changes the probability of rejecting the null hypothesis that installation price has no impact on degradation at conventional confidence levels. The effect of module technology on degradation is variable relative to the base category, HIT-Si but not statistically significant. The addition of vintage also has no significant effect on the price-degradation relationship. It should be noted that at lower confidence levels (75-80%) the effect of installation price and nameplate rating have consistent effects on faster degradation across models. At these lower levels, a dollar increase in installation price is associated with a 0.05% faster annual degradation, suggesting higher cost systems may be less reliable or that the effect is simply more apparent in higher cost systems. Similarly, a 1 kW increase in nameplate rating is associated with 0.04% faster annual degradation. This could be an indication that degradation is more apparent in larger systems. Although the model still explains an extremely limited amount of the variation in degradation and statistical significance is not achieved at conventional levels, there is some indication that price may be an important indicator of degradation. Areas for future investigation include using additional variables to explain more variation in degradation, which can also decrease the potential that omitted variables bias regression coefficients and include additional methods to account for uncertainty in rate medians.

Figure 5. Individual system regression



Highlighting this relationship, Figure 5 shows faster yet statistically insignificant association of degradation to increasing installation price. Overall, with individual system degradation rates shown on the vertical axis and installation price on the horizontal axis, the clear trend is faster degradation for increased installation price. Wide scatter of degradation rate medians away from the sloped line and outside of the 95 percent confidence interval indicate low model fit.

All Systems Model and Results

Taking a different view of the data by modeling all systems without differentiating degradation rates for individual systems allows three important changes; it drastically increases the number of data points used to estimate the response of degradation to price, hides none of the uncertainty around the rate medians of individual systems, and allows for the addition year fixed effects to account for unobserved characteristics in each year not explained by the other variables. As with individual systems, Equation 3 uses a linear model to fit the observed data. I modeled the degradation rate (R_{d_ij}) indexed by system *i* at hourly identifier *j* in Equation 3.

$$R_{d\_ij} = \boldsymbol{\beta_1 Pr_i} + \boldsymbol{\beta_2 Rat_i} + \boldsymbol{\beta_3 Eff_i} + \boldsymbol{\beta_4 TPO_i} + \boldsymbol{\beta_5 VOS_i} + \boldsymbol{\beta_6 Mtech_i} + \boldsymbol{\beta_7 Vin_i} + \boldsymbol{\beta_8 Temp_{ij}} + \boldsymbol{\theta_t} + \epsilon_{ij} \qquad \textbf{(3)}$$

Compared to Equation 2, Equation 3 explores different sources of variation that include mean ambient temperature ($Temp_{ij}$) at hourly intervals at each system location. Because Section 1 showed that there are statistically significant differences in degradation between years, I also included a fixed effect (θ_t) to control for any unobserved time-varying factors within years. Note that this model was coded with dummy variables for the year fixed effect and no coefficient is displayed. Without year fixed effects, coefficients are estimated based on installation price and the other covariates cross-sectionally *and* over time. Such an addition aids the ability to make an inference about the relationship. This ability is not available in the approach specified in Equation 2. I also added a cluster robust standard error to account for error correlation between the observations of each system (Cameron and Miller 2015). See Appendix B: Degradation and Price, Gauss-Markov Assumptions for OLS for more details. The variables in Equation 3 are listed and described in Table 9.

Table 9. Descriptive statistics for all-system regression variables

Variable	Mean	SD	Min	Max	Description and units
Degradation Rate	-1.4	27	-100	100	Degradation rate w/o weather correction **(%/year)**
Installation Price	9.2	2.3	3.5	22	Purchase price of system **(2014 real $/W)**
Nameplate Rating	4.8	2.3	1.3	14	DC capacity at STC **(kW)**
Module Efficiency	.15	.02	.12	.19	Conversion efficiency **(%)**
Third Party Ownership	N/A	N/A	N/A	N/A	Third party owned versus customer owned **(categorical)**
Value of Solar	9.2	2.4	5.1	15	Value of incentives/utility bill savings **(real 2014 $/W)**
Module Technology	N/A	N/A	N/A	N/A	Mono-c-Si, Multi-c-Si, a-Si, HIT-Si **(categorical)**
Temperature	22	7	-10	48	Ambient temperature **(°C)**
Vintage	2008	1	2007	2011	Year of system construction
N = 2,878,039					

The regression results using the entire dataset and including the year fixed effects are shown in Table 10. The top row contains the variable of interest, installation price. There are nine models specified with the last using a fixed effect. Like the individual systems regression results, these results show that this model explains an extremely small amount of the variation in the data. At best, the R-squared value in column 8 is 0.002. As in the individual system regressions, this is likely due to potential data quality issues in the dependent variable and lack of additional covariates and even with a drastically increased number of observations the model fails to explain much variation in degradation rates. From left to right across columns, I added covariates. The response of degradation to price is initially of similar magnitude as in the individual systems but increases slightly with the addition of nameplate rating and is then statistically significant at all conventional levels except in the presence of year fixed effects.

Table 10. Regression results for all systems

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9) FE-year
Installation Price	**-0.0472**	**-0.0689**	**-0.0650**	**-0.0649**	**-0.0667**	**-0.0651**	**-0.0650**	**-0.0651**	**-0.0611**
	(0.0527)	**(0.0532)**	**(0.0538)**	**(0.0538)**	**(0.0537)**	**(0.0559)**	**(0.0558)**	**(0.0559)**	**(0.0544)**
Nameplate Rating		-0.0763	-0.0735	-0.0736	-0.0754	-0.0750	-0.0753	-0.0750	-0.0730
		(0.0550)	(0.0542)	(0.0543)	(0.0543)	(0.0537)	(0.0537)	(0.0537)	(0.0533)
Module Efficiency			-3.468	-3.420	-3.350	-4.972	-4.941	-4.972	-5.086
			(5.369)	(5.452)	(5.483)	(7.818)	(7.813)	(7.818)	(7.702)
Third Party Ownership				0.0281	-0.00996	-0.0290	-0.0259	-0.0290	-0.0640
				(0.302)	(0.316)	(0.319)	(0.320)	(0.319)	(0.313)
Value of Solar					-0.0205	-0.0233	-0.0216	-0.0233	-0.0193
					(0.0402)	(0.0429)	(0.0429)	(0.0429)	(0.0426)
Mono-c-Si						-0.245	-0.242	-0.245	-0.179
						(0.302)	(0.301)	(0.302)	(0.304)
Multi-c-Si						-0.306	-0.303	-0.306	-0.223
						(0.412)	(0.412)	(0.412)	(0.413)
Vintage							-0.0000	0.00104	0.00888
							(0.109)	(0.110)	(0.00906)
Temperature								-0.00294	-0.0611
								(0.0095)	(0.0544)
Observations	2,878,039	2,878,039	2,878,039	2,878,039	2,878,039	2,878,039	2,878,039	2,878,039	2,878,039
R-squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Number of ID									386

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

There are three main takeaways from the results in Table 10. None of the coefficients of price are statistically significant at conventional levels, here again and evidenced by the R-squared of 0.002 in model 8, only a fraction of the variation in degradation at the hourly level is explained, and overall, the models show slightly faster degradation because of increased installation price relative to those in Table 8. From left to right in Table 10, model 2 adds nameplate rating to the model and shows faster degradation from increased price. In model 3, adding module efficiency is associated with faster degradation in that a one percent increase in module efficiency has an effect in models 3-7 of anywhere from about 3.5 to 5 percent increase in degradation. Contrary to hypothesized, this indicates that module efficiency negatively affects performance. This is a surprising result and may be explained by a lack of additional control variables. Still, the effect is statistically insignificant so conclusions based on this result must acknowledge this uncertainty. Although it changes in sign as additional covariates are added, third party ownership relative to the base category of host-owned systems has no statistically significant effect across models 4-7 and the inclusion effected the degradation response to price very little. The value of solar remains statistically insignificant across models 5-7. In those jurisdictions where the total value of all incentives and savings on utility bills increases by one dollar per Watt, systems are associated with more faster degradation by about 0.02 percent per year. Mono and Multi-c-Si module technologies show a statistically insignificant and faster degradation relative to HIT modules and a slower degradation response to price. Temperature is associated with faster degradation when price increases and the effect of temperature itself shows that an increase of one degree Celsius drives faster degradation only very slightly.

The main difference in models is the addition of over 2 million observations, which when making the correction for cluster robust standard errors, has not increased the precision of the estimates indicated by the lack of statistical significance for all covariates. Year fixed effects did not change the precision to the extent that any covariate has a statistically significant effect on degradation at conventional levels and exerted only a slightly positive influence on the magnitude of the coefficient. This means that unobserved time-varying factors in years may not be a significant source of unobserved variation because the ability to infer a statistically significant effect of installation price on degradation is not achieved. Degradation gets consistently faster with increased price and although no coefficient on installation price is statistically significant above the 90 percent level all aside from that in model 1 are significant at the 75 to 80 percent confidence levels. As in the individual systems regressions, this shows that installation price may still be an important factor in explaining degradation at the hourly level but one that does not rise above the noise even when dramatically increasing the number of observations.

Although generally in the opposite direction than hypothesized, the individual and all-systems models show no statistical significance for the effect of price on degradation. However, a lower confidence level would grant statistical significance in a stable way across models. A dollar increase in installation price per watt is associated with an increase in degradation by anywhere from about -0.05 to -0.07 percent per year. With and without fixed effects, these results have extremely limited predictive capacity because of low model fit. Any inference must be tempered by this fact. Further research to delineate the drivers of whole system degradation should include more covariates that separate system component characteristics from characteristics of installation practices, and potentially control for unobserved time-varying factors but at different temporal levels. Remembering the two systems in Oakland, CA, just blocks apart, even though price may be an important determinant, this data is inconclusive that the difference in price for each system is statistically attributable to differences they experience in terms of whole-system degradation.

Since I am interested both in the statistical and economic significance of degradation, I paint a picture of that significance in Section 3 by comparing a median degradation system to a system with zero degradation to highlight the difference in value.

Section 3: The value of degradation

The careful consumer considering a PV system should question whether paying more for a system ensures enhanced performance and return on investment. Although this relationship requires further study to demarcate the causal mechanisms, estimating the economic value of degradation based on observed rates provides a more complete portrait of the impact to consumers. I used the levelized cost of energy (LCOE) defined as the present value of all lifetime project inputs divided by the lifetime energy production, to assess this impact.

Whatever the relationship of degradation to price, the simple fact that systems vary in consistent energy production over their lifetimes begs the question of the financial impact of degradation on consumers. Studies outlining the sensitivity of LCOE to varying rates of degradation were either not done on deployed systems (Dirk C. Jordan et al. 2016) or are from utility-scale instead of residential data (Darling, You, and Velosa 2011). While the sensitivity of LCOE to degradation ranks below the real discount rate and conversion efficiency (Darling, You, and Velosa 2011) variation in degradation rates introduces risk into the cash flows needed to meet the financial obligations required if a project requires financing and return on investment in the case of outright purchase. Beyond financial risk, the perception of PV as a reliable technology is critical for positive public opinion. The viability of solar PV as an emergent technology necessitates estimating the value of degradation.

Methods

Investment in a grid-connected PV system with degradation represents the lost electricity an owner cannot send to the grid; a capital cost with declining cash-flows. I calculated the LCOE of four different degradation scenarios by using the 10^{th}, 50^{th}, 90^{th} percentiles, and zero degradation, considering the 10^{th} percentile high degradation, the 50^{th} medium, and the 90^{th} low degradation. The units of degradation are percent loss per year making high degradation the greatest loss at -2.4%/year, median at -0.9%/year, and low to be 0.00%/year, after rounding. These rates come from the individual systems rates found in the right-hand column of Table 2 in Section 1.

The dataset lacks system-level information for discount rates, useful life, and inverter replacement, and because I investigate the variation in the effect of degradation on LCOE I assumed a standard capital cost or system price. I retrieved capital cost for and consider a benchmark residential system to align with the mean cost in 2014 from Feldman et al. (2014) of $3.74/Watt. Although the minimum discount rate can be derived through after-tax income on 30-year treasury bills, or using Department of Energy rate of 3 percent for energy projects (Rushing, Kneifel, and Lippiatt 2013), these may be thought of as minimum rates. Evidence suggests that residential PV system owners have significantly higher discount rates depending on whether a system owner engages in lease (21 +/- 14%) or a purchase (7 +/- 5%) (Rai and Sigrin 2013). As have others for financial PV estimates (Feldman et al. 2014), I used a 6.2 percent discount rate in this analysis.

Many modeling efforts of cost-competitiveness of solar PV consider the useful lifetime of grid-tied solar PV to be near 30 years (Reichelstein and Yorston 2013; Drury, Denholm, and Margolis 2011; Darling, You, and Velosa 2011; Tidball et al. 2010). Instead empirically-based system lifetimes, much of the thinking behind useful lifetime relates to the warranty period offered by either the module manufacturer or the installer. Photovoltaic-specific information services aiming to drive uncertainty out of the PV purchase process such as Energy Sage, use product and performance warranty information for modules instead of whole systems, where many top module manufacturers offer 10-year warranties with performance guarantees around 80 percent performance output after 25 years of use (Energy Sage, 2017). Assuming a 25-year system life may underestimate the productive life but represents a decision point for system owners. Therefore, I used 25 years as the system life by which I model the change in LCOE due to degradation. Table 11 lists all these assumptions.

Table 11. Assumptions for value of degradation

Assumption	Value	Source
Capital cost	$3.74/W*	Feldman et al. 2014
Inverter replacement	$0.57/W*	Kneifel and Webb 2016 from Liu et al. 2014; Goodrich, James, and Woodhouse 2012
Discount rate	6.2%	Feldman et al. 2014
Useful life	25 years	Author’s assumption

* These were translated to $/kW for use in Equation 4 below.

The assumptions in Table 11 input into Equation 5 result in LCOEs with variation resulting from high, median, and low degradation rates. Equation 5 takes mean values of yearly insolation, system rating, efficiency, area, and PR from the dataset. Here I assumed no annual operations and maintenance costs nor residual value after 25 years of useful life. The numerator in Equation 5 includes the capital cost and an inverter replacement cost multiplied by the nameplate rating of the system to represent lifetime cash outflows. Typical inverter warranty periods are 10 years. I therefore include a warranty replacement cost discounted to present value in dollars per Watt at 10 years from installation. Because I assumed no annual costs for operations and maintenance nor other residual value post-project life, these cash flows are not discounted. They are the full lifetime cost of the system. The denominator shows the energy generated in the first year of operation multiplied by the degradation factor R_d, which decreases annual production as n increases from year 1 to 25. The initial production ($kWh_{initial}$) is the result of multiplying the number of days in the year by the mean POA irradiance per square meter per year across all systems sites. For more details, see Appendix C, LCOE Calculations.

$$LCOE = \left(\frac{\left(CapCost_{system} + \ Inv_{replace}\right) * Rat}{\sum_{n=1}^{N}(kWh_{initial} * Eff * A * PR * \ (1 + R_d)^n)}\right) \qquad \textbf{(4)}$$

Where:

$CapCost_{system}$ = Cost of system before incentive and tax credits (\$/kW)
$Inv_{replace}$ = Inverter replacement at 10 years of system life (\$/kW)
Rat = Mean system nameplate rating (kW)
PR = Mean system performance ratio over the study period
$kWh_{initial}$ = Mean energy produced in the first full year of operation (AC-kWh)
Eff = Mean efficiency across systems in this dataset
A = Mean area of systems in this dataset
R_d = Annual linear degradation rate estimated in Section 1 at p10, p50, and p90 (%/year)
n = Useful life of the system (years)

I created the differences between the 10th, 50th, and 90th percentiles to zero degradation as percentage changes in LCOE. The lost value from degradation (V_{lost}) is shown here as the net present value of the difference between lifetime energy production with and without degradation over a 25-year useful life using discount rate (D_r).

$$V_{lost} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{Rev_{noDeg\_n} - Rev_{Deg\_n}}{(1+D)^n}\right) \quad \textbf{(5)}$$

Where:

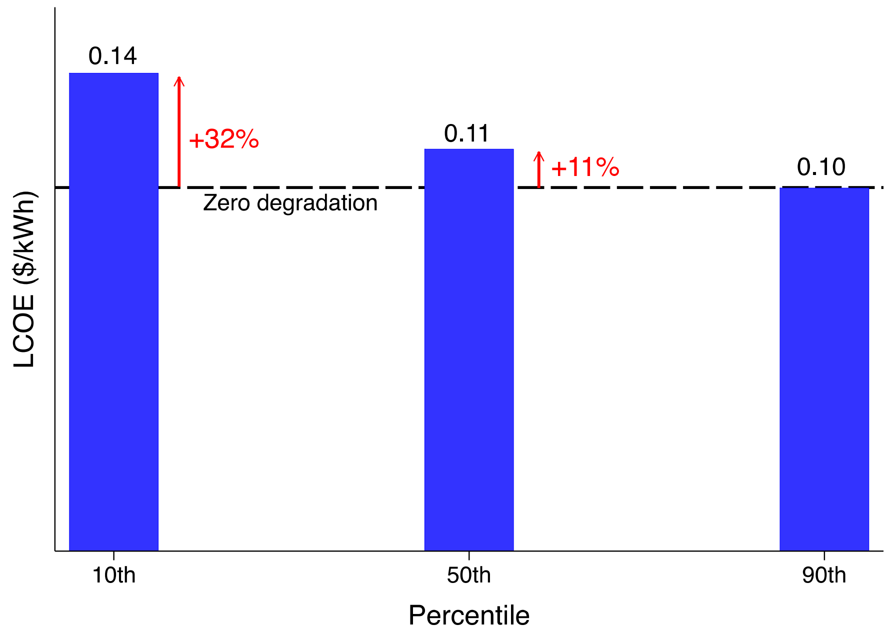
V_{lost} = Lost lifetime value from degradation ($)
Rev_{noDeg_n} = Revenue in year *n* with zero degradation ($/year)
Rev_{Deg_n} = Revenue in year *n* with 10th, 50th, or 90th percentile degradation ($/year)
D = Discount rate (6.2%)

The calculation for annual revenue is the single year energy production in the denominator of Equation 4 multiplied by the mean retail electricity rate in the dataset. This calculation is shown in Appendix C, LCOE Calculations.

Results

The economic significance of degradation is far from trivial. Figure 6 shows that there is a range of $0.10/kWh (low) to $0.14/kWh (high) around the median of $0.11/kWh caused by degradation. A system with zero degradation has an LCOE of $0.10/kWh. The changes in LCOE resulting from changes in degradation percentile are shown in Figure 6. Here, a high degradation system has an impact of increasing the cost per kWh of solar 32% over the horizontal dotted line that represents zero degradation. A low degradation has almost no effect because it is so close to the zero-degradation line. The median case is 11% higher than the dotted line meaning that median degradation adds 11% to the cost of a kWh from solar PV in the residential systems in this dataset. Note that LCOE numbers are rounded so simple arithmetic produces different results.

Figure 6. Lifetime levelized cost and percent changes from degradation



Regarding annual dollars lost to degradation, these numbers translate to a decrease in net present value (NPV) of all the cash flows over a 25-year system life of $1,740 for the high degradation system, $700 for a median system, and a $0 decrease for a low degradation system. Dividing these number by 25 years results in the mean annual dollar loss of $70 for high, $28 for median degradation, and $0 for a low degradation system.

Median degradation observed in this dataset adds 11% to the LCOE of residential solar. High degradation adds almost triple that, about 32%, and low degradation—because of estimates of positive degradation—adds almost no cost at all. Lost solar for a mean system in this dataset centers around $700 but can lose anywhere from $1,740 to $0 over the lifetime of the system. In the hypothetical example of two systems in Oakland, CA, just blocks apart, a similar investment could return $1,740 less over a 25-year system life, which signifies a 32% increase in the LCOE for that customer. Importantly, the degradation rates including the median case are estimates derived from the analysis detailed in Section 1 and represent real-world financial impacts of degradation.

Discussion and Conclusion

This work assesses the relationship between whole system degradation in mature systems and the price customers pay for a residential solar system in California under the California Solar Initiative. I hypothesized that higher cost has some advantage in performance seen through degradation. I found that median linear degradation rate estimates across all systems in the study sample are close to 1% annually, with a small confidence interval and similar when the fleet median is estimated across the medians of individual systems. However, median degradation rates for individual systems have widely varying uncertainties, which when included in a regression framework allows for very little explanatory power given the covariates. An attempt at a more precise relationship through a regression framework that includes all systems and all their estimates similarly reveals that a one dollar per Watt increase in system price is associated with faster degradation by about 0.05 to 0.07 per year but the effect is not statistically significant at conventional levels. Thus, even though higher price appears to result in faster degradation and the relationship is stable across models, they explain only a tiny fraction of the variation in degradation. Ability to explain variation is important but the persistence of statistical significance at high but not conventional levels points to installation price as a potentially important indicator of degradation, especially in future work.

The economic significance from degradation is also significant. Although there is a large range of 32% increase in LCOE for high degradation to 0% for low degradation, a median degradation system adds 11% to the LCOE. The consequences of an underperforming investment are not trivial. The federally legislated investment tax credit that is currently set to start decreasing from 30 percent in 2019 to 10 percent in 2022. Without this incentive, anywhere where the retail electricity rate is below the levelized cost of solar, investment in a residential system is non-economic because the cost to produce energy is higher than its compensation. For example, the current average retail rate across the U.S. is $0.1053/kWh. Estimates from this analysis suggest that without incentives, investment with even median degradation where price approaches $0.11/kWh, would not likely be considered by residential customers. Because this cost is largely hidden from consumer's economic decision-making is obscured. This could affect not only value to consumers but incentive structures, and discourse between proponents and investors, and requires verification and ultimately disclosure to stakeholders.

As previously stated, analysis results beyond showing that degradation is both variable and impacts the lifetime value of solar PV are limited by the ability of these regression models to explain the variation in degradation with statistical significance. Although beyond the scope of this report, further research should seek variables that help separate system components from installation practices such as cost of labor, materials, customer acquisition, etc. and ways of limiting the uncertainty in individual degradation rates such as decreasing the width of the degradation rate distributions by performing YOY differencing from one year to all other years.

Additional and more in-depth areas for future research should include unobserved temporal factors such as installer learning, changes in system shading over time, experimenting with lifetime kWh production or performance ratio instead of degradation as a proxy for system quality, investigating the causes of the worst cases of degradation, controlling for geographic variation in price at the utility-level, and as noted in Section 1, a full analysis of soiling.

Determining the relationship between degradation and price may be more tractable when studied across time to control for installer learning and installer market competition. As installers move down a learning curve one would expect that reputation is at least somewhat dependent upon degradation (system quality) and price, especially within the same area such as a utility service area. An index of the median degradation rate and median price within a utility service area could compare degradation between high and low-cost systems within the same area. Therefore, a reputation or learning variable could capture some variation from installer learning that impacts both price and degradation. Furthermore, because installer learning is likely correlated with price and degradation the current models are likely biased from this omission.

Changing shade conditions over time at individual system sites, which is currently not controlled for in this study warrants future study as well. Future investigation should separate performance data seasonally because the sun is lower in the sky over winter months in the northern hemisphere and system design may fail to account for difference in shade and partial shade structures such as trees that only shade an array when the sun angle is low in winter months. A more specific sampling procedure could also include site surveys that inquire directly into customer or third-party owner habits that deal with shade over time. Quantifying changing shade conditions is important because it is unobservable and heavily impacts system performance and financial return.

Different measures of system quality may more easily reveal a relationship to price. Using degradation to establish a relationship between price and system performance over time has inherent difficulties. This study consistently shows low model fit to degradation either because of data quality issues in the degradation rate calculation or lack of additional covariates. However, this could also be the result of looking for small effects in degradation when that measure includes large variation. Another option could be to simply use total lifetime system output in kWh or the performance ratios themselves. The potential advantages are in the simplicity of calculation and less overall variation relative to effect size.

Information about the performance of residential PV systems is important both for transparency in price and credibility of a PV technology. Therefore, an inquiry into the causes of the worst cases of degradation warrant future investigation on its own. This could also inform educational efforts that decrease risk by making system owners proactive both during the purchase process and system lifetime. The information could also inform incentive structures that weigh the costs and benefits of compensating performance versus simply installation.

Finally, because there is known variation between utility service areas two approaches warrant inquiry. First, interacting utility with price would capture the effect of price variation across utilities. Second, full analysis of soiling using more granular rain data would be useful in isolating the effect of degradation from soiling that may appear across years and utilities, which is currently not captured in the models in this study.

Still, this analysis combines degradation and price at the level of the entire installation over a portion of its lifetime and offers a starting point for understanding the potential trade-off between system quality and price. This relationship is important because consumer perception of the benefits they will receive from purchasing a potentially important low-emissions technology at “low-cost” need not indicate low quality. Given the variation in degradation and the implications for lifetime cost, the results shown here are encouraged by the existence of performance guarantees, real-time customer performance information, and product warranties. The performance differences hypothetical PV owners in Oakland experience have yet to be explained by any larger pattern of installation price. Still, neighbors and colleagues will be asking how satisfied they are with what they got. Therefore, future research into mechanisms promoting performance is required to increase transparency to potential customers.

References

Anderson, Mike, Zoe Defreitas, and Ernest Hasselbrink, Jr. 2013. “A System Degradation Study of 445 Systems Using Year-over-Year Performance Index.” *Sandia PV Performance Modeling Workshop Santa Clara, CA. May 1-2, 2013*, 95134.

Barbose, G., Darghouth, N. R., Weaver, S., Feldman, D., Margolis, R., & Wiser, R. (2015). Tracking US photovoltaic system prices 1998–2012: a rapidly changing market. *Progress in Photovoltaics: Research and Applications*, *23*(6), 692-704.

Barsun, Stephan. 2010. “Sample design_Rev6.” (Sampling plan for performance moniroting of CSI 15-minute interval data) Itron, Inc. Oakland, CA. (Received 2-4-2016)

Bollinger, Bryan, and Kenneth Gillingham. 2012. “Peer Effects in the Diffusion of Solar Photovoltaic Panels.” *Marketing Science* 31 (6): 900–912. doi:10.1287/mksc.1120.0727.

Bollinger, Bryan, and Kenneth Gillingham. 2014. “Learning-by-Doing in Solar Photovoltaic Installations.” *Yale University Working Paper*, 46. http://ssrn.com/abstract=2342406.

State of California, California-Nevada Climate Applications Program, Center for Western Weather and Water Extremes, Southwest Climate Science Center. 2014. “California Rain.” https://scripps.ucsd.edu/programs/cnap/wp-content/uploads/sites/109/2017/02/CA_Precip_climatology_Final.pdf. (Accessed 4-7-2017). http://www.water.ca.gov/floodmgmt/hafoo/csc/docs/CA_Rain_2pager.pdf.

Cambell, Matt. 2008. “The Drivers of the Levelized Cost of Electricity for Utility-Scale Photovoltaics.” *Sunpower*. sunpowercorp.com.

Cameron, A. Colin, and Douglas L. Miller. "A practitioner’s guide to cluster-robust inference." Journal of Human Resources 50, no. 2 (2015): 317-372.

Darling, Seth B, Fengqi You, and Alfonso Velosa. 2011. “Environmental Science Assumptions and the Levelized Cost of Energy for Photovoltaics.” *Society* 4 (9). doi:10.1039/c0ee00698j.

Deceglie, Michael G., Matthew Muller, Zoe Defreitas, and Sarah Kurtz. "A scalable method for extracting soiling rates from PV production data." In *Photovoltaic Specialists Conference (PVSC), 2016 IEEE 43rd*, pp. 2061-2065. IEEE, 2016.

Dierauf, Timothy, Aaron Growitz, Sarah Kurtz, Jose Luis Becerra Cruz, Evan Riley, and Clifford Hansen. *Weather-corrected performance ratio*. No. NREL/TP-5200-57991. National Renewable Energy Laboratory (NREL), Golden, CO., 2013.

Drury, Easan, Paul Denholm, and Robert Margolis. 2011. “The Impact of Different Economic Performance Metrics on the Perceived Value of Solar Photovoltaics.” *Technical Report (NREL/TP-6A20-52197)*, no. October: 1–44.

Duffie, J. a., William a. Beckman, and W. M. Worek. 2013. *Solar Engineering of Thermal Processes, 4nd Ed. Journal of Solar Energy Engineering*. 4th ed. Hoboken, NJ: John Wiley & Sons, Inc. doi:10.1115/1.2930068.

Energy Sage. "What to Know About a Solar Panel Warranty" http://news.energysage.com/shopping-solar-panels-pay-attention-to-solar-panels-warranty/. (accessed June 5, 2017)

Feldman, David, and Travis Lowder. *Banking on solar: an analysis of banking opportunities in the US distributed photovoltaic market*. No. NREL/TP-6A20-62605. National Renewable Energy Laboratory (NREL), Golden, CO., 2014. doi:10.2172/1164885.

Fthenakis, Vasilis, James E. Mason, and Ken Zweibel. 2009. “The Technical, Geographical, and Economic Feasibility for Solar Energy to Supply the Energy Needs of the US." *Energy Policy* 37 (2): 387–99. doi:10.1016/j.enpol.2008.08.011.

Gillingham, Kenneth, Hao Deng, Ryan H Wiser, Naïm R Darghouth, Gregory Nemet, Galen L Barbose, Varun Rai, and C G Dong. 2014. "Deconstructing Solar Photovoltaic Pricing: The Role of Market Structure." *Technology and Policy* 37 (3): 231–50.

Goodrich, Alan, Ted James, and Michael Woodhouse. *Residential, commercial, and utility-scale photovoltaic (PV) system prices in the United States: current drivers and cost-reduction opportunities*. No. NREL/TP-6A20-53347. National Renewable Energy Laboratory (NREL), Golden, CO., 2012. doi:10.2172/1036048.

Hansen, Clifford W. 2015. "Final Technical Report Advanced Solar Resource Modeling and Analysis." *Sandia National Laboratories* SAND2015-1. pv.sandia.gov.

Hay, John E. 1979. "Calculation of Monthly Mean Solar Radiation for Horizontal and Inclined Surfaces." *Solar Energy* 23 (4): 301–7. doi:10.1016/0038-092X(79)90123-3.

Itron, Inc. 2012. "CPUC California Solar Initiative 2010 Impact Evaluation: Addendum to Final Report." (California Public Utilities Commission archived, Davis, CA. 2012) 1-376.

Jones-Albertus, Rebecca, David Feldman, Ran Fu, Kelsey Horowitz, and Michael Woodhouse. 2015. "Technology Advances Needed for Photovoltaics to Achieve Widespread Grid Price Parity." *Prog. Photovolt: Res. Appl.* 15 (February 2013): 659–76. doi:10.1002/pip.

Jordan, D. C., and S. R. Kurtz. 2013. "Photovoltaic Degradation Rates - An Analytical Review." *Progress in Photovoltaics: Research and Applications* 21 (1): 12–29. doi:10.1002/pip.1182.

Jordan, Dirk C., Sarah R. Kurtz, Kaitlyn VanSant, and Jeff Newmiller. 2016. "Compendium of Photovoltaic Degradation Rates." *Prog. Photovolt: Res. Appl.* Vol. 24. doi:10.1002/pip.

Jordan, Dirk C, and Sarah R Kurtz. 2015. "Field Performance of 1 . 7 GW of Photovoltaic Systems." *IEEE Journal of Photovoltaics* 5 (1): 243–49. doi:10.1109/JPHOTOV.2014.2361667.

Kann, Shayle, MJ Shiao, Shyam Mehta, Cory Honeyman, Nicole Litvak, Jade Jones, Justin Baca, Will Lent, and Shawn Rumery. 2013. "U.S. Solar Market Insight Report: 2013 Year in Review Executive Summary." http://www.seia.org/sites/default/files/resources/5jBprenCY92013ye.pdf?key=5937919.

Kimber, A., L. Mitchell, S. Nogradi, and H. Wenger. "The effect of soiling on large grid-connected photovoltaic systems in California and the southwest region of the United States." In *Photovoltaic Energy Conversion, Conference Record of the 2006 IEEE 4th World Conference on*, vol. 2, pp. 2391-2395. IEEE, 2006.

King, D L, W E Boyson, and J A Kratochvil. 2004. "Photovoltaic Array Performance Model." *Sandia Report No. 2004-3535* 8 (December): 1–19. doi:10.2172/919131.

Kneifel, Joshua, David Webb, and Eric O'Rear. "Energy and Economic Implications of Solar Photovoltaic Performance Degradation." *Special Publication (NIST SP)-1203* (2016).

Kymakis, Emmanuel, Sofoklis Kalykakis, and Thales M. Papazoglou. "Performance analysis of a grid connected photovoltaic park on the island of Crete." *Energy Conversion and Management* 50, no. 3 (2009): 433-438. doi:10.1016/j.enconman.2008.12.009.

Lave, Matthew, William Hayes, Andrew Pohl, and Clifford W. Hansen. "Evaluation of global horizontal irradiance to plane-of-array irradiance models at locations across the United States." *IEEE journal of Photovoltaics* 5, no. 2 (2015): 597-606.

Liu, Xiaohui, G. O. Eric, Wallace E. Tyner, and Joseph F. Pekny. "Purchasing vs. leasing: A benefit-cost analysis of residential solar PV panel use in California." *Renewable Energy* 66 (2014): 770-774. doi:10.1016/j.renene.2014.01.026.

Lonij, Vincent P., Adria E. Brooks, Kevin Koch, and Alexander D. Cronin. 2012. "Analysis of 80 Rooftop PV Systems in the Tucson, AZ Area." *Conference Record of the IEEE Photovoltaic Specialists Conference*, 549–53. doi:10.1109/PVSC.2012.6317674.

Maghami, Mohammad Reza, Hashim Hizam, Chandima Gomes, Mohd Amran Radzi, Mohammad Ismael Rezadad, and Shahrooz Hajighorbani. "Power loss due to soiling on solar panel: A review." *Renewable and Sustainable Energy Reviews* 59 (2016): 1307-1316. doi:10.1016/j.rser.2016.01.044.

Marion, B, J Adelstein, K Boyle, H Hayden, B Hammond, T Fletcher, D Narang, A Kimber, L Mitchell, and S Richter. 2005. "Performance Parameters for Grid-Connected PV Systems." In *Photovoltaic Specialists Conference, 2005. Conference Record of the Thirty-First IEEE*, 31:1601–6. doi:10.1109/PVSC.2005.1488451.

Mejia, Felipe A., and Jan Kleissl. "Soiling losses for solar photovoltaic systems in California." *Solar Energy* 95 (2013): 357-363.

Nemet, Gregory, Eric O'Shaughnessy, Ryan H. Wiser, Naim R. Darghouth, GalenL. Barbose, Ken Gillingham, and Varun Rai. 2017. "Sounces of Price Dispersion in U.S. Residential Solar Installations." Berkeley. https://emp.lbl.gov/publications/sources-price-dispersion-us.

Noll, Daniel, Colleen Dawes, and Varun Rai. "Solar community organizations and active peer effects in the adoption of residential PV." *Energy Policy* 67 (2014): 330-343. doi:10.1016/j.enpol.2013.12.050.

Perez, Richard, Pierre Ineichen, Robert Seals, Joseph Michalsky, and Ronald Stewart. 1990. "Modeling Daylight Availability and Irradiance Components from Direct and Global Irradiance." *Solar Energy* 44 (5): 271–89. doi:10.1016/0038-092X(90)90055-H.

Phinikarides, Alexander, Nitsa Kindyni, George Makrides, and George E. Georghiou. "Review of photovoltaic degradation rate methodologies." *Renewable and Sustainable Energy Reviews* 40 (2014): 143-152. doi:10.1016/j.rser.2014.07.155.

Rai, Varun, and Ariane Beck. 2015. "Public Perceptions and Information Gaps in Solar Energy in Texas." *Environmental Research Letters* 10 (7). IOP Publishing: 74011. doi:10.1088/1748-9326/10/7/074011.

Rai, Varun, and Benjamin Sigrin. 2013. "Diffusion of Environmentally-Friendly Energy Technologies: Buy versus Lease Differences in Residential PV Markets." *Environmental Research Letters* 8 (1): 1–8. doi:10.1088/1748-9326/8/1/014022.

Reichelstein, Stefan, and Michael Yorston. "The prospects for cost competitive solar PV power." *Energy Policy* 55 (2013): 117-127. doi:10.1016/j.enpol.2012.11.003.

Rode, Johannes, and Alexander Weber. 2011. "Knowledge Does Not Fall Far from the Tree-A Case Study on the Diffusion of Solar Cells in Germany." *ERSA Conference Papers*.

Rushing, Amy S., Joshua D. Kneifel, and Barbara C. Lippiatt. 2013. "Energy Price Indices and Discount Factors for Life-Cycle Cost Analysis - 2013." doi:http://dx.doi.org/10.6028/NIST.IR.85-3273-28.

Sengupta, Manajit, A. Habte, S. Kurtz, A. Dobos, S. Wilbert, E. Lorenz, T. Stoffel, D. Renne, C. Gueymard, D. Myers, S. WIlcox, P. Blanc, R. Perez. 2015. "Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications." *National Renewable Energy Laboratory*. doi:10.1016/j.solener.2003.12.003.

Sharma, Vikrant, and S. S. Chandel. "Performance and degradation analysis for long term reliability of solar photovoltaic systems: a review." *Renewable and Sustainable Energy*

Reviews 27 (2013): 753-767. doi:10.1016/j.rser.2013.07.046.

Swanson, Richard. 2006. “A Vision for Crystalline Silicon Photovoltaics.” *Prog. Photovolt: Res. Appl.* 14 (January 2006): 443–53. doi:10.1002/pip.

Tidball, Rick, Joel Bluestein, Nick Rodriguez, and Stu Knoke. *Cost and performance assumptions for modeling electricity generation technologies*. No. NREL/SR--6A20-48595. ICF International, Fairfax, VA (United States), 2010. doi:10.2172/993653.

U.S. Energy Information Administration. 2016. *International Energy Outlook 2016.* Vol. 0484(2016)doi:www.eia.gov/forecasts/ieo/pdf/0484(2016).

U.S. Energy Information Administration. What is U.S. electricity generation by energy source? https://www.eia.gov/tools/faqs/faq.php?id=427&t=3. Accessed August 28, 2017.

U.S. Energy Information Administration. Solar, natural gas, wind make up most 2016 generation additions. https://www.eia.gov/todayinenergy/detail.php?id=25172. Accessed August 28, 2017.

Woodhouse, Michael, Rebecca Jones-Albertus, David Feldman, Ran Fu, Kelsey Horowitz, Donald Chung, Dirk Jordan, and Sarah Kurtz. *On the Path to SunShot. The Role of Advancements in Solar Photovoltaic Efficiency, Reliability, and Costs*. No. NREL/TP--6A20-65872. National Renewable Energy Lab.(NREL), Golden, CO (United States), 2016. http://www.osti.gov/scitech/biblio/1253983.

Wooldridge, Jeffrey. 2011. *Introductory Econometrics*. *A Modern Approach*. Cengage Learning. 2012.

Appendix A: Degradation Calculations

Examples of Discarded Systems

Figure A - 1 is an example of data verified as unusable for this analysis due to insufficient data from too few observations. Systems with similar trends seen in daily data were excluded because a degradation rate calculation using the YOY approach requires at least 2 years of data by which to difference PRs. Figure A - 2 and Figure A - 3 show data shifts. The shift in Figure A - 2 cannot be reconciled because the factor to correct the shift is unknown whereas the shift in Figure A - 3 represents a reversed current transducer and can be corrected using the absolute value of the magnitude. Several systems exhibited this condition and were corrected. Note that plots were made with daily points instead of hourly points.

Figure A - 1. Too few observations

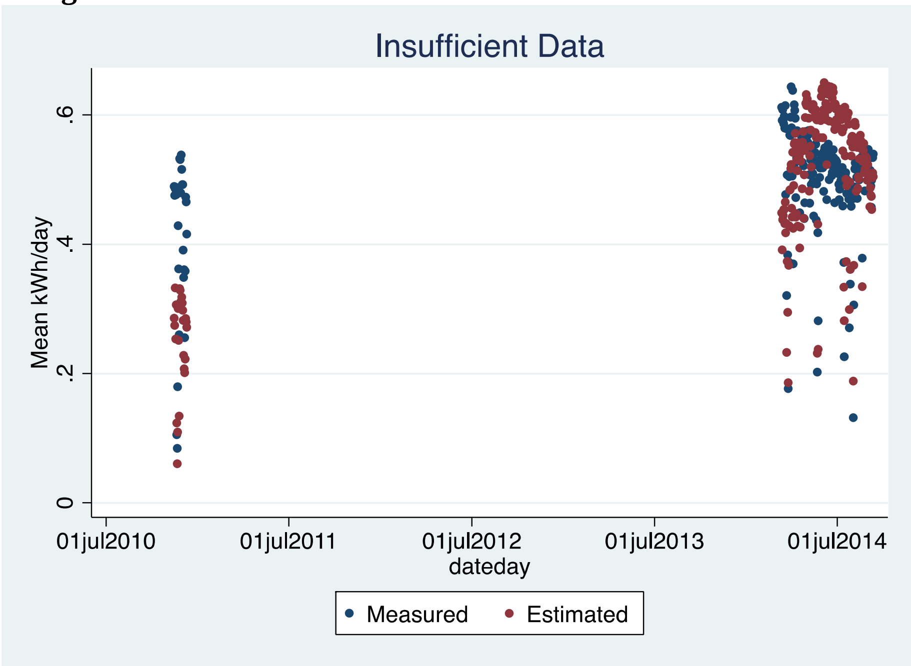


Figure A - 2. Unreconciled shift

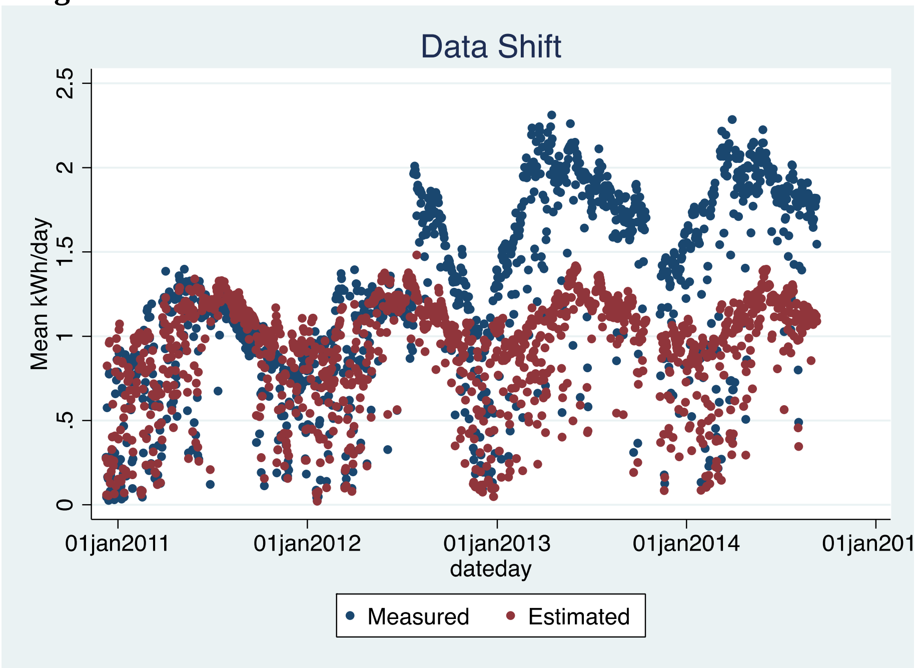
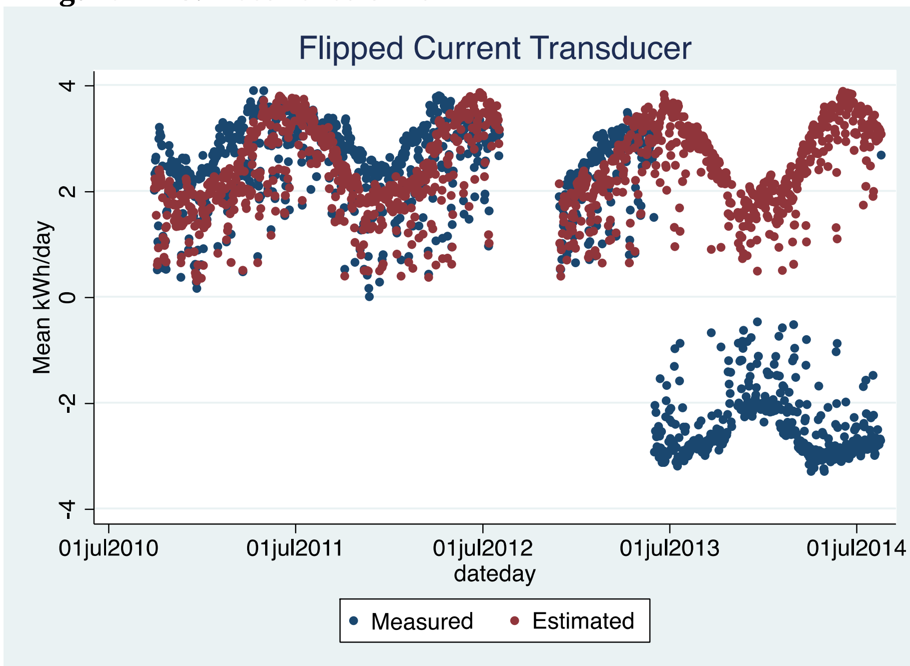


Figure A - 3. Reconciled shift



Plane of Array Radiation

Equation A3 – A6 show the calculations made to arrive at POA irradiance using the Perez model (1990). Equations A3 – A5 were taken from PV Performance Modeling Collaborative available at: https://pvpmc.sandia.gov/.

$$\boldsymbol{E_{POA} = E_B + E_G + E_D} \qquad \textbf{(A3)}$$

Plane of array beam component

$$\boldsymbol{E_B = DNI * \cos(\theta)} \qquad \textbf{(A4)}$$

DNI = Direct normal irradiance
θ = Angle of incidence

Plane of array ground-reflected component

$$\boldsymbol{E_G = GHI * Albedo * \frac{\left(1 - cos\left(\theta_{T,Surface}\right)\right)}{2}} \qquad \textbf{(A5)}$$

Where:

E_G = Ground-reflected irradiance
GHI = Global horizontal irradiance
Albedo = Reflectivity of the surrounding ground surface.
$\theta_{T,Surface}$ = Surface angle of the surrounding ground

Plane of array sky-diffuse component

$$\boldsymbol{E_D = DHI\left[(1-F_1)\left(\frac{1+cos(\theta_T)}{2}\right)+F_1\left(\frac{a}{b}\right)+F_2 sin(\theta_T)\right]} \qquad \textbf{(A6)}$$

Where:

F_1= Function that describe circumsolar and horizon brightness
F_2 = Function that describe circumsolar and horizon brightness
a = max(0, cos(θ))
b = max(cos(85º)), cos(θz))
θ = Angle of incidence
θz = Solar zenith angle
θ_T = Array tilt from horizontal

This is a basic description of the Perez model (Perez et al. (1990). For a more detailed description see Duffie, Beckman, and Worek (2013).

Standard Test Conditions

Standard Test Conditions (STC) are a set of laboratory conditions under which PV modules are tested to provide a nominal rating and is how system nameplate ratings are defined. The conditions are standardized to the following

- Ambient temperature of 25°C
- Irradiance of 1000 W/m^2
- American Society for Testing and Materials (ASTM) G173-03 Standard Spectrum Irradiance
- Air mass index of 1.5

Convective Heat Transfer Equations and Coefficients

Equation A1 describes a generalized formula for empirically derived module back surface-temperature based on King, Boyson, and Kratochvil (2004b) and used in Dierauf et al. (2013) to correct for weather influence on module and array DC output.

$$T_m = G_{POA} * \left\{e^{(a+b*WS)}\right\} + T_a \qquad \textbf{(A1)}$$

Where:

T_m = module back-surface temperature [°C]
G_{POA} = POA irradiance from calibrated reference cells [W/m2]
T_a = ambient temperature [°C]
WS = the measured wind speed corrected to a measurement height of 10 meters [m/s]
a = empirical constant reflecting the increase of module temperature with sunlight
b = empirical constant reflecting the effect of wind speed on the module temperature [s/m]
e = Euler's constant and the base for the natural logarithm.

While wind speed and module temperature increase with incident sunlight, it is their influence on module temperature cell temperature that influences the module temperature coefficient of power. The effect of module cell temperature is derived using Equation A2 found in Dierauf et al. (2013).

$$\boldsymbol{T_{cell} = T_m + \left(\frac{G_{POA}}{G_{STC}}\right) * \Delta Tcnd} \qquad \textbf{(A2)}$$

Where:

T_{cell} = predicted operating cell temperature [°C] 13
T_m = predicted module surface temperature as determined by Equation (3) [°C]
G_{POA} = POA irradiance, as described above [W/m2]
G_{STC} = reference irradiation for the correlation; constant at 1,000 [W/m2]
$\Delta Tcnd$ = conduction temperature drop as presented in Table 2 below**.**

I used row-two coefficients. The data do not provide information regarding mounting type so coefficients specific to each system cannot be known.

Table A - 1. Conduction coefficients

Module Type	Mount	a	b	ΔTcnd(°C)
Glass/cell/glass	Open rack	-3.47	-0.0594	3
Glass/cell/glass	Close-roof mount	-2.98	-0.0471	1
Glass/cell/polymer sheet	Open rack	-3.56	-0.0750	3
Glass/cell/polymer sheet	Insulated rack	-2.81	-0.0455	0
Polymer/thin-film/steel	Open rack	-3.58	-0.1130	3

Source: Table 2, page 12 in King, Boyson, and Kratochvil (2004b).

Computing the irradiance-weighted average cell temperature includes the cell temperature (T_{cell}) and POA irradiance as shown in Equation A3. This is a location-specific yearly-average cell temperature that corrects the PR from measured data to generate a predicated value.

$$\boldsymbol{T_{cell\_typ\_avg}} = \frac{\sum\left(\boldsymbol{G_{POA\_typ\_j}} * \boldsymbol{T_{cell\_typ\_j}}\right)}{\sum\left(\boldsymbol{G_{POA\_typ\_j}}\right)} \qquad \textbf{(A3)}$$

Where:

$T_{cell\_typ\_avg}$ =average irradiance-weighted cell temperature from one the current year of weather data
$T_{cell\_typ\_j}$ = calculated cell operating temperature for each hour (ºC)
$G_{POA\_typ\_j}$ = POA irradiance for each hour determined from one the current year
j = hour of the year
This was taken directly from Dierauf et al. (2013), page 13.

Positive Degradation

Further illustrating the point that no physical impossibility precludes degradation rates from being negative, Figure A - 4 and Figure A - 5 show that degradation rates can be positive when all available years are included. The median degradation rate in Figure A - 4 is 0.1 %/year. However, notice the gap in data and then increase in hourly degradation rates following the gap in 2014. When I discarded all the data in 2014 as shown in Figure A - 4, the resulting calculation of the median degradation rate for this system becomes slightly negative at -0.008 %/year, in Figure A - 5. Therefore, rather than a manifestation solely of uncertainty, a period of low followed by a period of higher measurements influence median degradation rates. Underlying these rates are PR values that show a large positive change from 2013 to 2014. Therefore, low performance ratios in one year affect the degradation rate when differencing over years.

Figure A - 4. Time-series plot of degradation exhibiting negative degradation

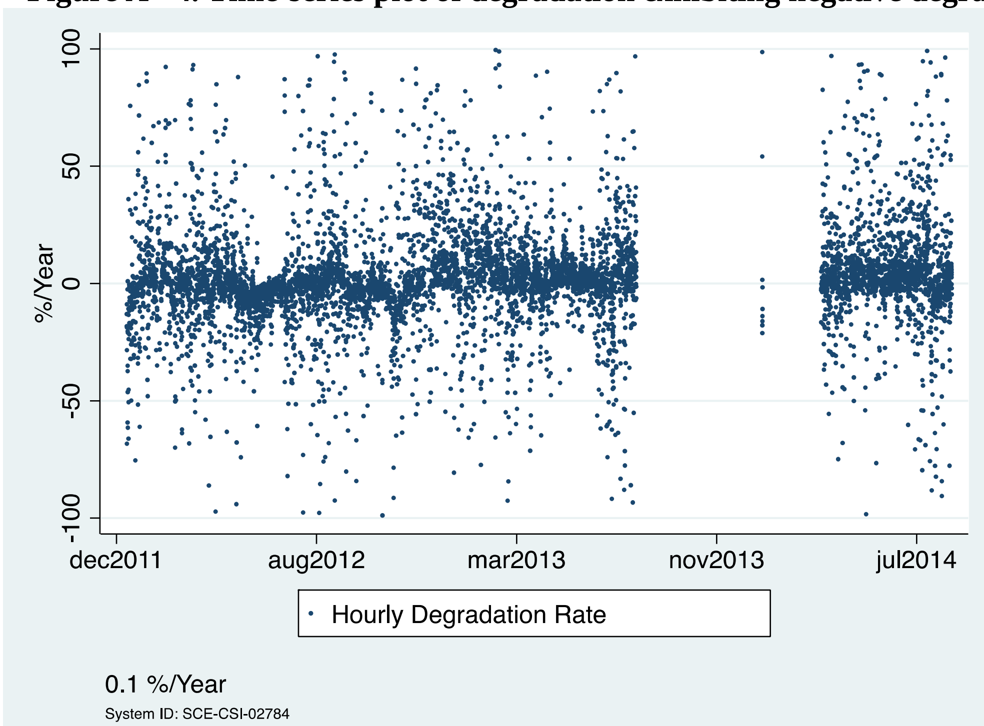
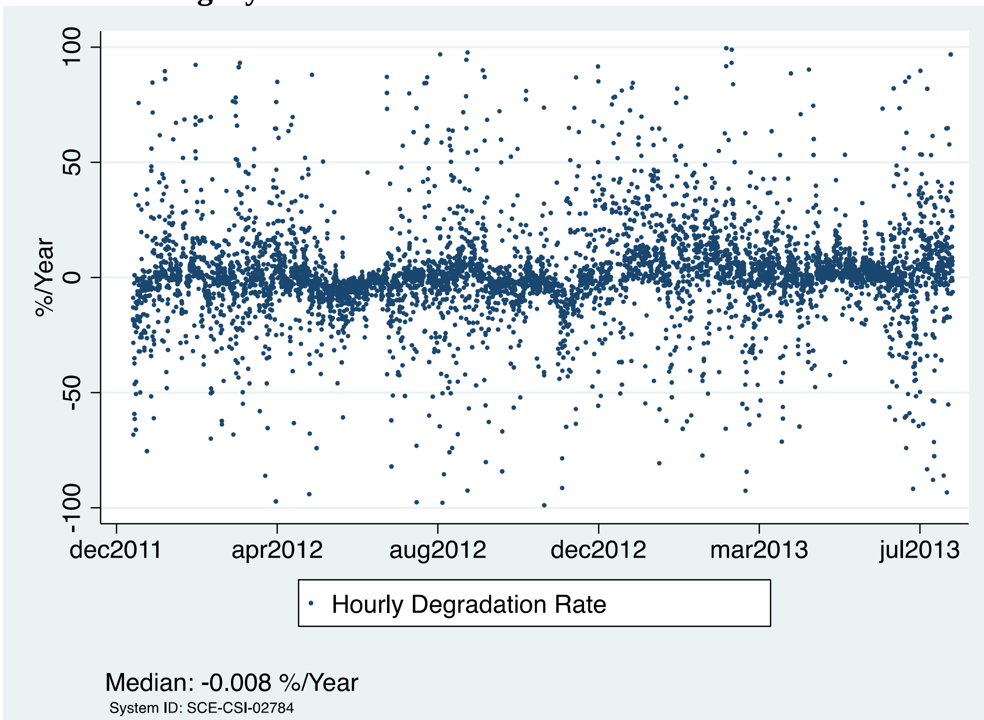


Figure A - 5. Time-series plot of degradation exhibiting positive degradation after discarding a year



Appendix B: Degradation and Price

Gauss Markov assumptions for OLS

To eliminate the potential for biased coefficients when using ordinary least squares (OLS) for multiple linear regression five assumptions must be defended. These are commonly known as the Gauss-Markov assumptions. In short, these justifications must demonstrate that the expected value of the errors resulting from the regression is equal to zero, these errors are uncorrelated with any of the independent variables, and they have equal variance. Below, I provide a brief justification for the variables used in Equations 2 and 3 in Section 2 meeting these criteria. These criteria are taken from Wooldridge (2011) and the interpretation are my own.

1. **Linear in parameters**
   The regression coefficients have a ceteris paribus population effect on the variables excluding the situation where the coefficients are non-linear. Even though there is the potential for the line of best fit to include a squared term, the squared term is a result of the variable and not the coefficient or parameter. In the case of the degradation regression coefficients, evidence of non-linearity is beyond the scope of this study but does exist. Thus, the parameters are linear in any case.

2. **Random sampling**
   The sampling of the PV systems used in this analysis is random and has been chosen to be a representative sample of the population of the systems in the CSI rebate program. The sites chosen were selected using a randomized Monte Carlo Simulation to arrive at sample requirements based on year of installation, program administrator, location and mounting type across commercial and residential sites to ensure

statistically significant results can be gathered for different types of systems. This information was gathered from Barsun (2010).

3. **No perfect collinearity**
Sample variation is required in each independent variable with no independent variables being an exact linear combination of another. The variation included in each independent variable is listed in Table 7 and Table 9.

4. **Zero conditional mean**
Unobserved factors must be unrelated to independent variables and on average the expected value of the error term given all variables is equal to zero. Thus, given the data, there are no additional covariates hypothesized to impact degradation and are therefore introducing bias into the coefficient on the price variable in Equation2 and 3 in Section 2. In other words, there are no variables that contain unaccounted-for variation in the models presented above. Further research should be conducted to prove the validity of this assumption.

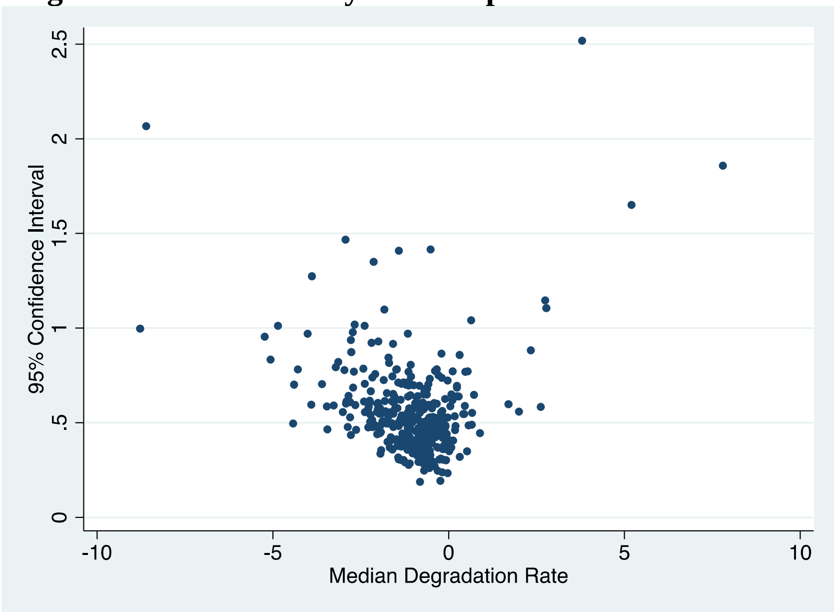
5. **Homoskedasticity**
The error term must be the same variance across independent variables. This can be empirically tested and statistically treated by estimating standard errors robust to heteroskedasticity, which I have done above. In Equation 2. This assumption affects the ability to make a sound inference based on the magnitude of the estimated standard errors. By introducing robust standard errors, the estimates of statistical significance I made in Section 2 can be considered conservative while correcting for any heteroskedasticity. Equation 3 uses cluster robust standard errors to account for the fact that regression model errors are independent and uncorrelated across systems but within systems this correlation exists. Although this correction has no bearing on

the magnitude of the coefficients, not employing this correction is likely to result in standard errors that overstate the precisions with which a statistically significant effect can be discerned (Cameron and Miller 2015). Again, this is a method of making inference from the model in Equation 3 more conservative.

Uncertainty in Degradation Rate Medians

As noted in Section 2, the relationship between positive degradation and uncertainty is plotted in Figure B - 1. The median degradation rates for each system are on the vertical axis and the bootstrapped standard error of the median on the horizontal axis. A pattern exists in which the most extreme values of the median degradation, both in the positive and negative direction, have the largest uncertainties associated with them. Although beyond the scope of this work, this indicates that controlling for the uncertainty in the estimate of the median degradation rate is important.

Figure B - 1. Uncertainty in the dependent variable



Appendix C: Value of Degradation Calculation

LCOE Calculations

I modified the levelized cost of energy in Equation 5 above, from Equations C1 and C2 used in previous research (Cambell 2008; Darling, You, and Velosa 2011). Reviewing Equation 3.5 shows that I excluded the factor that appears to discount kWh as Equation C2 shows. Solutions have been rounded to the nearest cent.

$$LCOE = \frac{Lifecycle\ cost}{Lifetime\ energy\ production} \tag{C1}$$

$$LCOE = \frac{Project\ cost + \sum_{n=1}^{N} \frac{AO}{(1+D_r)^n} - \frac{RV}{(1+D_r)^n}}{\sum_{n=1}^{N} \frac{Initial\ kWh * (1-R_d)^n}{(1+D_r)^n}} \tag{C2}$$

$$LCOE_{p10} = \frac{\left(\frac{\$3.74}{Watt} + \frac{\$0.57}{Watt}\right) * \frac{1000\ Watt}{kW} * 4.8kW}{\sum_{n=1}^{N} \left(\frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR}\right) * \left(1 + \left(-\frac{0.024}{year}\right)\right)^n} = \frac{\$\mathbf{0.14}}{\boldsymbol{kWhh}}$$

$$LCOE_{p50} = \frac{\left(\frac{\$3.74}{Watt} + \frac{\$0.57}{Watt}\right) * \frac{1000\ Watt}{kW} * 4.8kW}{\sum_{n=1}^{N} \left(\frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR}\right) * \left(1 + \left(-\frac{0.009}{year}\right)\right)^n} = \frac{\$\mathbf{0.11}}{\boldsymbol{kWh}}$$

$$LCOE_{p90} = \frac{\left(\frac{\$3.74}{Watt} + \frac{\$0.57}{Watt}\right) * \frac{1000\ Watt}{kW} * 4.8kW}{\sum_{n=1}^{N} \left(\frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR}\right) * \left(1 + \left(\frac{0.000}{year}\right)\right)^n} = \frac{\$\mathbf{0.10}}{\boldsymbol{kWh}}$$

$$LCOE_{NoDeg} = \frac{\left(\frac{\$3.74}{Watt} + \frac{\$0.57}{Watt}\right) * \frac{1000\ Watt}{kW} * \ 4.8kW\ * 0.86_{PR}}{\sum_{n=1}^{N}\left(\frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR}\right)} = \frac{\$\mathbf{0.10}}{\boldsymbol{kWh}}$$

Additional lifetime cost from zero to high degradation (10th percentile) and zero to low (90th percentile) is shown in the calculations below. Percentages have been rounded to the nearest whole percent.

$$\Delta\ LCOE_{p90\_NoDeg} = \frac{LCOE_{p90} - LCOE_{NoDeg}}{LCOE_{p50NoDeg}} = \mathbf{0.0}\%$$

$$\Delta\ LCOE_{p50\_NoDeg} = \frac{LCOE_{p50} - LCOE_{NoDeg}}{LCOE_{NoDeg}} = \mathbf{11}\%$$

$$\Delta\ LCOE_{p10\_NoDeg} = \frac{LCOE_{p10} - LCOE_{NoDeg}}{LCOE_{NoDeg}} = \mathbf{32}\%$$

Lost value

The following calculations show the revenue after the first year of operation with a mean system in the dataset. The only difference is a result of the degradation rates. The only difference between these calculations and those above is the retail electricity rate $R_{e,}$ which across all 386 systems in the dataset is $0.15/kWh (with rounding). These are the intermediate step not shown in Equation 5, Section 3. Due to rounding, solutions do not equate but were modeled without rounding until the final solution, which is shown here.

$$Rev_{NoDeg} = \frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR} * \frac{\$0.15}{kWh} R_e = \$\mathbf{1,225}$$

$$Rev_{p90} = \frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR} * \frac{\$0.15}{kWh} R_e * (\mathbf{1 - 0.000}) = \$\mathbf{1,225}$$

$$Rev_{p50} = \frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR} * \frac{\$0.15}{kWh} R_e * (\mathbf{1 - 0.009}) = \$\mathbf{1,214}$$

$$Rev_{p10} = \frac{1746\ kWh}{m^2 * year} * 0.15 * 32.5m^2 * 0.86_{PR} * \frac{\$0.15}{kWh} R_e * (\mathbf{1 - 0.024}) = \$\mathbf{1,195}$$

Note that the changes between the 90th percentile degradation and zero degradation are equivalent above and slightly different below because 90th percentile degradation from Section 1 is very slightly positive. The following calculations show the NPV of the difference in the lifetime cash flows from the 10th, 50th, and 90th percentiles to zero degradation.

$$Value\ Difference_{p10\_NoDeg} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{Rev_{p10} - Rev_{NoDeg}}{(1+D_r)^n}\right) = -\$\mathbf{1740}$$

$$Value\ Difference_{p50\_NoDeg} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{Rev_{p50} - Rev_{NoDeg}}{(1+D_r)^n}\right) = -\$\mathbf{700}$$

$$Value\ Difference_{p90\_NoDeg} = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{Rev_{p90} - Rev_{NoDeg}}{(1+D_r)^n}\right) = \$\mathbf{0}$$