Empirical development of parsimonious model for international diffusion of residential solar

Eric Williams a,*, Rexon Carvalho b, Eric Hittinger c, Matthew Ronnenberg d

a Golisano Institute for Sustainability, Rochester Institute of Technology, USA
b North American Power, Energy Aspects, USA
c Department of Public Policy, Rochester Institute of Technology, USA
d Program of Color Science, Rochester Institute of Technology, USA

A R T I C L E   I N F O

Article history:
Received 27 March 2019
Received in revised form 23 November 2019
Accepted 23 December 2019
Available online 27 December 2019

Keywords:
Solar power
Technology diffusion
Subsidy
Empirical model

A B S T R A C T

We develop a new parsimonious model of residential solar diffusion that, with only two regression parameters and one independent variable, reasonably explains empirical observations. Additional solar customers resulting from an increase in Net Present Value (NPV) are modeled as a normal distribution. This leads to adoption as a function of NPV being the integral of the Gaussian, producing the error function, which demonstrates S-curve behavior commonly seen in technology diffusion. Empirical analysis for five regions (three U.S. states: Arizona, California, and Massachusetts; and two countries: Germany and Japan) from 2005 to 2016 shows a consistent relationship between annual adoption per million households and NPV. Non-linear regression indicates good agreement between data and the error function model, the adoption rate peaking at an NPV of $7100/kW with standard deviation of $4110/kW. Consumer purchases of rooftop solar across multiple regions are explained with a single variable, making this model simpler than traditional diffusion approaches. A novel implication of the model is that the subsidy cost to stimulate additional solar adoption increases as the technology becomes cheaper. This is because the same subsidy is paid to all consumers, including those who would have purchased solar without subsidy.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

With the expectation that solar photovoltaic (PV) power will play a key role in mitigating climate change and improving energy security, governments around the world continue to invest in policies to promote adoption. Policies include tax credits, rebates, feed-in tariffs, net metering mandates and Renewable Portfolio Standards (RPS). These efforts have contributed to dramatic price reductions in PV systems. The global average module price has dropped substantially in the last several decades: from $29/Watt (2018US$) in 1981 to $0.42/Watt (2018US$) in 2018 [1]. Installed system costs vary by region, but have also fallen substantially, e.g. the average installation cost of residential solar in the U.S. went from $11.50/Watt (2017US$) in 2000 to $3.7/Watt in 2017 (2017US$) [2].

While progress has been encouraging, the installed costs of PV systems must fall further to enable widespread adoption without undue increase in energy prices. State and federal policymakers need straightforward answers to the question of what rates of solar adoption can be expected with a given set of policy incentives. There is a substantial history of modeling to understand relationships between PV adoption, policy choices and consumer behavior. Broadly speaking, the models aim to predict an aspect of solar diffusion such as adoption rate or purchase decision of a consumer as a function of explanatory variables, including time, economic costs/benefits, and demographics (including environmental attitudes, and location of consumers). A variety of modeling frameworks have been used, including Bass diffusion, discrete choice, fuzzy logic, agent-based and generic multivariable functions. Bass and related diffusion models describe penetration rates as following a S-curve relationship over time e.g. Refs. [3,4]. Discrete choice models typically construct a probability of purchase from a mathematical representation of utility [5,6]. Fuzzy logic simulates the solar purchase decision as a set of rules to combine fuzzy set representations of decision

* Corresponding author. 111 Lomb Memorial Drive, Rochester, NY, 14623, USA.
E-mail addresses: exwgis@rit.edu (E. Williams), rexon.carvalho@energyaspects.com (R. Carvalho), eshgpt@rit.edu (E. Hittinger), mxr8103@rit.edu (M. Ronnenberg).

0960-1481/© 2020 Elsevier Ltd. All rights reserved.
variables [7]. Multivariable regression models have been used to analyze the impacts of local environmental, social, economic and political variables on the adoption of residential solar PV [8,9], to identify demographic characteristics that enable the adoption of residential solar PV through third-party ownership [10], and spatial adoption patterns [11]. System dynamics models have been developed for residential solar diffusion and represent the problem as a network of algebraic relationships for interactions between influencing factors [12–14]. Recent years have seen the development of agent-based models of household residential PV adoption, which often aim to simulate spatially-resolved PV diffusion with explanatory variables such as socio-economic demographics, behavioral motivations, and/or technical advancements, e.g. Refs. [15–19].

The above trends in solar diffusion modeling reflect how developments in Information and Communications Technologies (ICTs) are enhancing and enabling data collection and manipulation, and the complexity of descriptive models. While the power of ICTs should be utilized to continue to understand the nuances of consumer adoption, it is also important to search for parsimonious models that can explain general trends and relationships. We propose that a good parsimonious model of a system satisfies the following criteria: 1. It explains the most empirical data with fewest free parameters. 2. It is reducible to logical assumptions about fundamental system interactions and 3. All parameters should have straightforward interpretations in terms of the system’s dynamics.

Here we pursue the hypothesis that a parsimonious model with few explanatory variables can reasonably describe residential PV adoption across multiple regions with different policy regimes and consumer cultures. We start with a model with one explanatory variable, Net Present Value (NPV) as experienced by residents in a particular region, to explain the rate of residential PV adoption. While economics is obviously an important factor in any consumer decision, note that describing solar panel purchases purely as a function of net economic value is in contrast with recent literature emphasizing additional factors such as type of purchase (lease vs. loan), adoption by neighbors, and environmental attitudes.

Annual data is collected for five regions to build the model: three U.S. states (Arizona, California, Massachusetts) and two countries (Germany, Japan). The choice of these regions was driven by data availability and because they are several of the most important regions in the world for residential PV adoption. Each region is treated as a geographic aggregate, using regional average data for all variables except solar insolation, which is taken from a central location. To enable comparison of adoption in different regions, we transform annual residential solar installations to an intrinsic variable by normalizing to the number of available detached homes in the region. Net Present Value is estimated based on installed system cost, electricity price, subsidies (feed-in tariff or otherwise), real interest rates and region-aggregated output of a solar system.

This model contributes to the theory and practice of modeling solar diffusion by proposing a simple but effective method for predicting the adoption of rooftop solar photovoltaics. It is the first model to reasonably reproduce solar diffusion in three countries with only one explanatory variable (Net Present Value) and two regression constants common to all regions. This is in contrast with previous models, which have more explanatory variables (often many more) and/or apply only to a single region [5–9,15–20]. Empirical validation for many prior models is complicated by a large number of explanatory variables predicting a limited data set, and may require even more variables to increase the resolution of adoption in a particular region. Furthermore, there is value in a baseline prediction that only depends on Net Present Value. First, this model supports planning of subsidy policy through direct prediction of adoption as function of support, an analysis that can be completed quickly and easily without advanced training. We apply the model in sections 5 and 6 to derive new results on relationships between subsidy level, technology price and local economic conditions. Second, the model is simple enough to be easily incorporated into national and international energy system models such as the National Energy Modeling System from the U.S. Energy Information Administration [21] and the World Energy Model (WEM) from the International Energy Agency [22] or applied with little modification to any research or analysis seeking to predict residential solar adoption.

2. Data and methods

2.1. Net Present Value

The explanatory variable used is average Net Present Value (NPV) as experienced by a resident in a given region in a given year. NPV combines a number of economic and policy variables such as system cost, electricity price and subsidies into one value. A general equation for the NPV of a subsidized residential PV system is:

\[
\text{NPV} (\$) = ( -C_{\text{total}} + S) + \sum_{i=1}^{N} \frac{TE \times SC \times RP \times (1 + \text{inf})^i}{(1 + \text{int})^i} + \sum_{i=1}^{M} \frac{TE \times (1 - SC) \times \text{FIT Price}}{(1 + \text{int})^i} 
\]

(1)

Where:

- \(C_{\text{total}}\): capital cost of the PV system (\$)
- \(S\): capital cost subsidy (\$)
- \(TE\): total electricity produced by the PV produced in one year (kWh)
- \(SC\): self-consumption share (%)
- \(RP\): retail price of electricity ($/kWh)
- \(\text{inf}\): inflation rate (%)
- \(\text{int}\): lending rate (%)
- \(\text{FIT Price}\): fixed feed-in-tariff price ($/kWh)
- \(i\): year
- \(N\): lifetime of solar system (years)
- \(M\): term length of FIT Price (years)

\(C_{\text{total}}\) is the investment cost of the PV system, \(S\) is an initial capital cost subsidy, e.g. the 30% federal tax rebate in the U.S. \(TE\) is the total energy produced in one year by the solar panel system, determined using the PVWATTS model from the National Renewable Energy Laboratory [23]. All installed solar in a region is assumed to have the output of a system placed in a latitudinally central city. Feed-in-tariff (FIT) subsidies, such as in Germany and Japan, pay consumers a fixed price for electricity supplied to the grid, i.e. energy after self-consumption in the home. The share of self-consumed electricity, \(SC\), has been estimated to be 33% in Germany [24] and 45% in Japan [25]. The FIT electricity price, \(\text{FIT Price}\), is typically fixed for a given time period, 20 years in Germany and 10 years in Japan. \(M\) refers to the term of the FIT policy. \(N\) is the total lifetime of a solar system, taken as 20 years. The retail price of electricity is denoted by \(RP\), which is assumed to increase every year with inflation rate \(\text{inf}\) assumed to be a continuation of a historical average [26]. For a system in a net metering regime, such as many U.S. states, all electricity generated garners the retail price - in the formula this corresponds to \(SC = 100\%\). In FIT regions, income is divided into a part from self-
consumption that gets the retail price (which increases with inflation) and a part from FIT income (which does not increase with inflation). The interest rate, \(int\), is the average annual lending rate in a region that year [27]. Note that this is a purely financial measure of discount rate and that solar-specific issues such as perception of risk and benefits are implicitly accounted for in the model developed below (Equation (4)).

This multi-year international analysis necessitates choices regarding the treatment of currencies, inflation and interest rates. Our method is as follows: NPV is first calculated mid-year. Given the availability of quarterly data for Japan, we split 2009 into two half years and calculate adoption and NPV for each half year period. There are capital subsidies in Japan in addition to FIT support, and in both cases except for Japan in 2009, when a new FIT policy dramatically spurred adoption when it was introduced mid-year. Given the availability of quarterly data for Japan, we split 2009 into two half years and calculate adoption and NPV for each half year period.

Data sources used to calculate NPV in the different countries are listed in Table 1. Constrained by availability of data, treatment years are 2005–2016 for Germany and Japan and 2011–2016 for Arizona and Massachusetts. Annual averages are used in all cases except for Japan in 2009, when a new FIT policy dramatically spurred adoption when it was introduced mid-year. Given the availability of quarterly data for Japan, we split 2009 into two half years and calculate adoption and NPV for each half year period. There are capital subsidies in Japan in addition to FIT support, and we use typical numbers reported by the International Energy Agency (IEA) [29]. The German government has guaranteed favorable interest rates for loans to purchase solar panels, and information from the IEA suggests these rates were 3.6%–4.15% in 2005 [30] and 4.5% for 2006–2008 [31]. Since 2006, the U.S. offered a federal tax incentive of 30% of PV system capital cost, initially capped at $2000 but later removed in 2009 [32]. In addition, California, Arizona and Massachusetts have state-level subsidies [33,34].

### 2.2. Adoption rates

There are several important details involved in developing an appropriate measure of adoption rate. First, in order to compare different regions, PV adoption must be normalized to derive an intrinsic quantity that is independent of the population of the region. We do this by dividing residential adoption by the number of detached houses in the region. This is a proxy for the number of potential residential adopters, neglecting inappropriately sited houses, community and multi-family installations.

Second, the number of houses available for new solar installations falls with increasing penetration of the technology. We thus define the “number of free detached houses” in a given year as:

\[
\text{Number of free detached houses} = \text{Stock of detached houses} - \text{total houses with PV system already installed}
\]

Annual adoption is thus measured as

\[
\text{Adoption} \left(\frac{\text{MW}}{\text{million houses}}\right) = \frac{\text{Annual residential adoption (MW)}}{\text{Number of free detached houses (million)}}
\]

Data sources used to quantify adoption in each region for each year are detailed in Table 1. Note that the number of detached houses was not available for every year, so we extrapolated/interpolated missing years using a linear fit.

### 2.3. Model for annual adoption as function of Net Present Value

The model to be tested starts from a simple hypothesis: the number of additional customers purchasing a PV system given an increase in NPV follows a normal distribution. In qualitative terms, this follows the standard “diffusion of innovation” logic but now linked with NPV, where new adoption is lower at the start and end of the distribution and fastest in the middle [47]. In this model, few customers will purchase residential solar when it has strongly negative NPV (the left side of a normal distribution), but more and more will become interested in adopting as the NPV improves, up to some peak level. After this point, improvements in NPV continue to drive more adoption, but at a declining rate. On the right side of the curve, very high NPV levels do not drive much new adoption because most customers were already willing to adopt at some lower NPV. For a given NPV, the number of consumers who will purchase is the integral

<table>
<thead>
<tr>
<th>Region</th>
<th>PV price</th>
<th>Electricity price</th>
<th>Subsidies</th>
<th>Residential adoption</th>
<th>Detached Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011–2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>California</td>
<td>[35]</td>
<td>[36]</td>
<td>State: [33], Federal: [32]</td>
<td>[38]</td>
<td>2013–2016 [37]</td>
</tr>
<tr>
<td>2005–2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>[39]</td>
<td>[40]</td>
<td>FIT: [39]</td>
<td>[39]</td>
<td>2005–2016 [41]</td>
</tr>
<tr>
<td>2005–2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005–2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011–2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that the number of detached houses was not available for every year, so we extrapolated/interpolated missing years using a linear fit.

\[
\text{number of detached houses} = \text{Stock of detached houses} - \text{total houses with PV system already installed}
\]
of a normal distribution,

\[ \text{Annual adoption} \left( \frac{\text{MW}}{\text{million houses}} \right) (\text{NPV}) = \alpha \int_{-\infty}^{\infty} dx e^{-\left( \frac{(x-e)^2}{2\sigma^2} \right)} \]

\[ = K \left( 1 + \text{erf} \left( \frac{\text{NPV} - \mu}{\sigma} \right) \right) \]

(4)

where \( \text{erf}(x) \) is the error function, \( \alpha \) is an arbitrary constant, later combined with integration constants to yield \( K \). \( K \) is one-half maximum annual adoption, thus can be fixed at 2000 MW/million households (hh), corresponding to every single homeowner in a region buying a 4 kW system that year. \( \mu \) is the NPV that results in maximum new customer acquisition and \( \sigma \) is the spread in this value, both to be determined empirically.

While non-economic factors relating to solar adoption are not modeled explicitly, this model implicitly includes those factors through the empirically-derived distribution. A perfectly uniform preference for PV would result in a delta-function in additional adoption, which is the result achieved when both \( \mu = 0 \) and \( \sigma \to 0 \). However, the empirical data shows that this is a poor fit to empirical data (Fig. 1). \( \sigma \) values larger than 0 correspond to consumers on average expecting more than economic breakeven. Non-zero \( \sigma \) reflects the observed diversity in consumer preferences. Note that if \( \sigma \) is of reasonable size compared to \( \mu \), adoption will be positive even with negative NPV. This realistic outcome represents consumers that buy PV for noneconomic reasons such as concern for the environment.

3. Results: Adoption model

Fig. 1 shows 47 data points for observed annual residential PV adoption versus Net Present Value (NPV) in five regions for different years. Notably, despite some noise, all regions appear to fall on a common curve without location-specific tuning of any sort. Apart from which model used to describe it, the empirical trend shown in the data in Fig. 1 is new and important: it shows that PV adoption in different regions and years fall on a similar curve. Given the many variables that could influence PV adoption rates, we find the explanatory power of NPV over multiple regions surprising and useful.

Numerical non-linear least squares regression yields \( \mu = \$7100/\text{kW} \) and \( \sigma = \$4110/\text{kW} \), with a total square error (TSE) of 13,770. The root-mean standard-deviation normalizes square error for the size of the data set (47 data points) to indicate average prediction error. The resulting root-mean standard-deviation is 17 MW/million households. Equation (4) using these values is plotted on Fig. 1, indicating a good empirical fit. Assessing the fit is complicated by non-linearity, leading to difficulties in using methods based on the assumptions of random effects, e.g., r-squared [48] or fixed effects [49]. We steer clear of perilous methodological waters here, noting that linear and exponential fits yield TSE of 22,186 and 13,893 respectively. In addition to having lower TSE, the error function model is preferred for three reasons. First, it is based on a logical fundamental assumption (normal distribution for additional solar customers). Second, the function has a maximum value bounded by the theoretical limit of 100% adoption. Third, the error function model has regression parameters \( \mu \) and \( \sigma \) that are directly interpretable in terms of solar purchasing preferences.

4. Model application: Subsidy payments per stimulated adoption

One particularly useful application for this model is prediction of PV adoption induced by policies that affect the financial benefits of rooftop solar, relevant to policy makers designing any form of consumer-facing adoption policy or seeking to meet adoption targets. Subsidies (or other subsidy-equivalent policies such as feed-in tariffs) are intended to induce adoption, but their actual effect lies between two extremes. At one end of the spectrum, if government subsidizes a product that consumers would have bought anyway, the subsidy spends public money with no benefit. At the other extreme, all adoption can be attributed to the subsidy, implicitly assuming that no consumer would have bought an unsubsidized product. Reality lies between these two unrealistic assumptions, and the model in Equation (4) enables quantification of induced adoption as follows:

\[ \text{Subsidy induced adoption} \left( \frac{\text{MW}}{\text{million hh}} \right) = \text{Adoption (NPV w/ subsidy)} - \text{Adoption (NPV no subsidy)} \]

(5)

The public cost of a subsidy, at least as they are typically implemented, depends on total adoption, i.e., the subsidy is paid to both consumers who would have purchased without the subsidy and those who wouldn’t have. For example, with a flat
capital cost subsidy ($/W), total expenditures on subsidies are given by:

$$\text{Subsidy expenditure} = \frac{\text{Adoption} \times \text{NPV} \times \text{subsidy}}{\text{million} \text{ hh}}$$

The ratio of equations (6) and (5) yields “subsidy expenditure per stimulated adoption”, an indicator of the economic effectiveness of a subsidy. This is plotted in Fig. 2 for three values of flat subsidy. The central result is that the public expenditures to induce a certain amount of adoption are lowest when the technology is new and increase as the technology becomes more economically attractive. This is because the subsidy is paid to all consumers, including free riders (Equation (6)), and the cost to the government grows faster than the stimulated adoption from the subsidy (Equation (5)).

5. Model application: Optimal solar subsidy

Equations (4)–(6) can also be used to estimate the “optimal” subsidy level that maximizes net social benefits. The public benefits of a subsidy, e.g. emissions reductions, scale according to the subsidy-induced adoption. The costs are the expenditures in Equation (6), which scale with total adoption. Considering a flat capital cost subsidy for solar ($/W), this leads to the following expression for the net benefits of a subsidy when CO2 reduction is assumed to be the social benefit:

$$\text{Net benefits} = \text{Subsidy induced adoption} \times \text{Lifetime CO2 reduction} \times \text{Social Cost of Carbon} - \text{Total adoption} \times \text{Subsidy}$$

Given a set of economic and physical conditions that determine NPV, i.e. system and electricity prices, solar resources and the social cost of carbon, Equation (7) quantifies the net benefits as a function of the subsidy, allowing determination of the optimal level.

We undertake a limited analysis of this type for a single year (2016) of up-front capital cost subsidy in Germany, Japan and California. Germany and Japan currently rely on FIT policies, so this analysis considers these regions enacting retail net metering and an up-front capital cost subsidy ($/W) similar to the U.S. This is not to suggest that the up-front subsidy is preferable. Rather, our idea is to explore how a similar policy framework would play out in regions with different economic and resource contexts. Furthermore, direct capital cost subsidy and feed-in tariff can be converted into one another using financial measures such as NPV, but we show them in the same format for comparison. We consider a decision to determine the optimal subsidy level in 2016 using solar system and electricity prices in that year and local solar resources. We follow the common practice of assuming that solar displaces the regional average carbon emissions per kWh [50], with data for regional average carbon emission intensities (for 2013) from Ref. [51]. Data and select model outputs for Germany, Japan and California are shown in Table 2.

The results of the analysis are shown in Fig. 3. It was found that a relatively high threshold of social cost of carbon was needed for the subsidy to realize net benefits, $71, $94 and $159/ton in California, Japan and Germany respectively. This is partly because this calculation neglects other important social benefits of rooftop solar, criteria pollution displacement for example, and partly because all three regions are expected to see substantial solar adoption without any subsidy whatsoever. Because of this, the social cost of carbon must be high to balance out subsidy payments to free riders. German solar requires a higher social cost of carbon to justify a subsidy, due to the lower solar resource in that region. The optimal

<table>
<thead>
<tr>
<th>Region</th>
<th>Residential solar price (2016 US$/W)</th>
<th>Electricity Price (2016 US$/kWh)</th>
<th>Solar resource (kWh/kW)</th>
<th>Residential solar NPV (2016 $/kW)</th>
<th>Carbon intensity of grid (g/kWh)</th>
<th>Subsidy expenditure per adoption (0.5$/W subsidy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>1.65</td>
<td>.32</td>
<td>900</td>
<td>2400</td>
<td>486</td>
<td>1.70</td>
</tr>
<tr>
<td>Japan</td>
<td>3.2</td>
<td>.22</td>
<td>1150</td>
<td>1380</td>
<td>572</td>
<td>1.55</td>
</tr>
<tr>
<td>California</td>
<td>4.3</td>
<td>.17</td>
<td>1550</td>
<td>240</td>
<td>489</td>
<td>1.38</td>
</tr>
</tbody>
</table>
carbon is needed to justify solar subsidies in Japan and Germany, only the direct and other emissions benefits are considered. This optimum is determined by mapping a test subsidy level to adoption by residents and estimating carbon displacement benefits. Note that while this analysis indicates a high social cost of carbon is needed to justify solar subsidies in Japan and Germany, only the direct benefits of adoption on carbon emissions are included here, and technological progress and other emissions benefits are neglected.

It is important to emphasize that this limited analysis only accounts for the direct benefits of a subsidy: the lowered carbon emissions from solar panels adopted in a given year. Subsidies for solar and other evolving technologies also induce indirect benefits by stimulating demand, which in turn contributes to lowering future costs of the technology through experience and technological progress. A host of experience curve analyses show that, at least retrospectively, lower solar prices and increased adoption have gone in hand-in-hand following a regular pattern [52]. Market development is indeed a large part of the motivation behind many technology subsidies and it is possible, at least in principle, for a subsidy to stimulate cost reductions leading to competitiveness of a technology [53]. While it is useful to understand the direct benefits of a subsidy, ultimately a larger analysis is needed that also includes indirect effects on technological progress and market development. We do not undertake this here, but note that significant indirect benefits are needed to justify solar subsidies for lower values of social cost of carbon.

6. Discussion

The contribution of this work is development of a new model of residential PV diffusion that maximizes net societal benefits (displaced CO₂ minus subsidy expenditure) in Germany, Japan and California as a function of the social cost of carbon (2016 US $/ton). The optimum is determined by mapping a test subsidy level to adoption by residents and estimating carbon displacement benefits. Note that while this analysis indicates a high social cost of carbon is needed to justify solar subsidies in Japan and Germany, only the direct benefits of adoption on carbon emissions are included here, and technological progress and other emissions benefits are neglected.

The model provides a reasonably robust explanation of diffusion of residential PV, as it uses a measure of potential adopters, but clearly not all houses are suitable for solar. The estimation of Net Present Value is used despite intra-regional differences in electricity prices, subsidy levels and solar insolation. For example, electricity production of solar panels in California is estimated using a site roughly in the middle of the state (1550 kWh/kW annually), but varies from 1340–1720 kWh/kW in different parts of the state. Model fit could be improved by collecting data to implement the model for smaller regions (e.g. by county in California). Also, the number of detached houses used a measure of potential adopters, but clearly not all houses are suitable for solar. The estimation of Net Present Value does not include factors such as house tenure, roof angle and direction, resale value and maintenance costs [54,55].

The second origin of differences between model and data is use of a single explanatory variable (Net Present Value) to explain a complicated outcome (residential PV diffusion). There are factors contributing to PV adoption not explicitly accounted for, including perception of environmental issues [56], degree of adoption by neighbors [11], financial risk (e.g. loan versus lease) [57], and concerns over maintenance [7]. There are many possible extensions of this model that would account for more factors, as well as potential simplifications of prior complex diffusion models. Model complexity usually comes at the price of higher data requirements and more free parameters. Additional free parameters complicate understanding even when explanatory power is actually gained. Tradeoffs between adding variables and explanatory power are reasonably well understood for multivariable linear regression, less so for non-linear regression, agent-based, fuzzy logic and system dynamics models.

The potential to improve our model should not distract the reader from the central point: an aggregated and relatively simple model provides a reasonably robust explanation of diffusion of residential PV in a variety of regions with different consumers and policy approaches. This simpler description enables quantification of how much a given subsidy contributes to adoption, which is needed for sound policy design.
Author contributions section
Erik Williams: Conceptualization, Methodology, Supervision, Data curation, Writing- Original draft preparation, Writing-Reviewing and Editing. 
Rexon Carvalho: Data curation, Investigation. 
Matthew Ronnenberg: Data curation, Investigation, Writing-Original draft preparation. 
Eric Hittinger: Visualization, Writing-Reviewing and Editing.

Declaration of competing interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements
This work was partially supported by the Civil Infrastructure Systems program of the U.S. National Science Foundation (grant# 1436469).

Appendix A. Supplementary data
The file “Data and model details.xls” is a spreadsheet showing all the data used and calculations for the diffusion model. Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2019.12.101.

References
[34] LBNL, Tracking the Sun 10, Lawrence Berkeley National Laboratory (LBNL), Berkeley, California, 2017.
[42] METI, Standardized Electricity Prices, Ministry of Economy and Trade (METI), Tokyo: Japan, 2016 [In Japanese].
[49] W. Greene, The behaviour of the maximum likelihood estimator of limited

576 E. Williams et al. / Renewable Energy 150 (2020) 570–577