Renewable Energy 150 (2020) 570–577

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

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.

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).

https://doi.org/10.1016/j.renene.2019.12.101
0960-1481/

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:

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

Where:

C_{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)
inf: inflation rate (%)
int: lending rate (%)
$FIT\ Price$: fixed feed-in-tariff price ($/kWh)
i: year
N: lifetime of solar system (years)
M: term length of FIT Price (years)

C_{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, *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 inf assumed as 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-

Table 1
Data sources for residential solar system price, electricity prices, and subsidies. Details of data used for each region and year are available in the Electronic Annex. FIT= Feed-in-tariff.

Region	PV price	Electricity price	Subsidies		Residential adoption	Detached Households
Arizona 2011–2016	[35]	[36]	State: [33], Federal: [32]		[35]	2013–2016 [37]
California 2005–2016	[35]	[36]	State: [34], Federal: [32]		[38]	2013–2016 [37]
Germany 2005–2016	[39]	[40]	FIT: [39]	Interest rate: 2005 [30], 2006–2008 [31]	[39]	2005–2016 [41]
Japan 2005–2016	[29] + previous years	[42]	FIT: 2009–2010 [43], 2011–2016 [44]	Capital: [29] + previous years	[45] + previous years	2013, 2008 [46]
Massachusetts 2011–2016	[35]	[36]	State: [33], Federal: [32]		[35]	2013–2016 [37]

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 in local, nominal currency for a region in a given year. NPV is then moved to real 2016 values in a nation's currency using the Consumer Price Index for each nation [26]. For Germany and Japan, local currencies are then converted to US dollars using the 2016 average exchange rate [28].

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, Japan and California 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:

$$\begin{aligned} &\textit{Number of free detached houses} = \textit{Stock of detached houses} \\ &\quad - \textit{total houses with PV system already installed} \end{aligned} \tag{2}$$

Annual adoption is thus measured as

$$\textit{Adoption}\left(\frac{MW}{\textit{million houses}}\right) = \frac{\textit{Annual residential adoption}\ (MW)}{\textit{Number of free detached houses}\ (\textit{million})} \tag{3}$$

Data sources uses 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

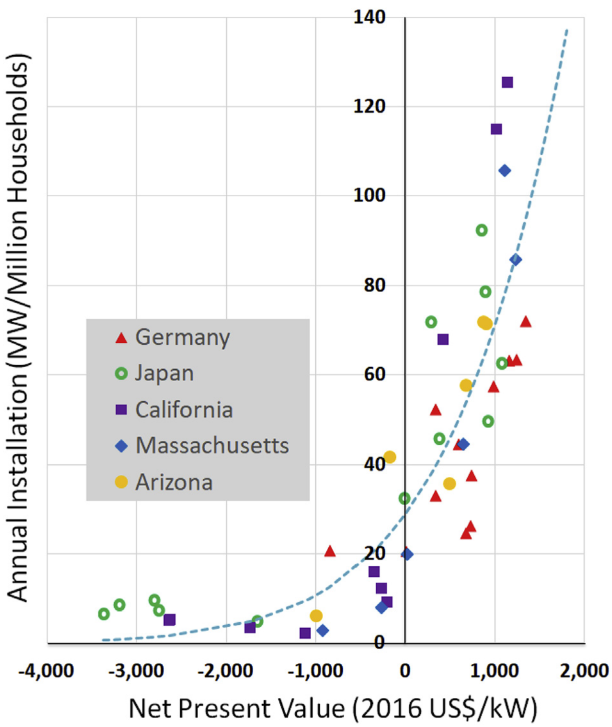

Fig. 1. Annual residential PV installations versus Net Present Value for homeowners in five regions (Germany, Japan, California, Massachusetts, Arizona). Each data point reflects an annual figure for the region, years treated are 2005–2016 for Germany, Japan and California, 2011–2016 for Massachusetts and Arizona. The dotted line is the fit from the error function model, Equation (4).

of a normal distribution,

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

$$= K\left(1+\mathrm{erf}\left(\frac{NPV-\mu}{\sigma}\right)\right) \quad (4)$$

whereerf (x) is the error function, α 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. μ is the NPV that results in maximum new customer acquisition and σ 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 customers ($\sigma \to 0$), leading to a step function for adoption as a function of NPV (i.e. If NPV is below some threshold, no one adopts. If NPV is above it, everyone adopts). Perhaps the simplest possible model would suggest that all consumers adopt if and only if their NPV >0, 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). μ values larger than 0 correspond to consumers on average expecting more than economic breakeven. Non-zero σ reflects the observed diversity in consumer preferences. Note that if σ is of reasonable size compared to μ, adoption will be positive even with negative NPV. This realistic outcome represents consumers that buy PV for non-economic 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 $\mu = \$7100$/kW and $\sigma = \$4110$/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 μ and σ 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:

$$Subsidy\ induced\ adoption\left(\frac{MW}{million\ hh}\right)$$

$$= Adoption\ (NPV\ w/\ subsidy) - Adoption\ (NPV\ no\ subsidy) \quad (5)$$

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

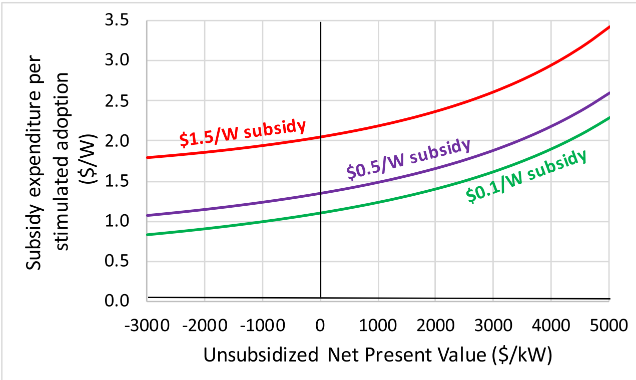

Fig. 2. Government subsidy expenditure per stimulated adoption (US$/Watt) for a flat subsidy of residential solar at three different values. Stimulating new adoption occurs at lowest public cost when the technology is new and expensive. As the technology becomes cheaper, the subsidy is increasingly paid to "free riders" who would have purchased without a subsidy.

capital cost subsidy ($/W), total expenditures on subsidies are given by:

$$Subsidy\ expenditure\left(\frac{\$}{million\ hh}\right) = Adoption\ (NPV\ w/\ subsidy)\left(\frac{MW}{million\ hh}\right) \times Subsidy\left(\frac{\$}{W}\right) \times 10^6 \tag{6}$$

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 CO_2 reduction is assumed to be the social benefit:

$$\begin{aligned} Net\ benefits\left(\$\right) &= Subsidy\ induced\ adoption\ (MW) \\ &\quad \times Lifetime\ CO_2\ reduction\left(\frac{ton\ CO_2}{MW}\right) \\ &\quad \times Social\ Cost\ of\ Carbon\left(\frac{\$}{ton\ CO_2}\right) \\ &\quad - Total\ adoption\ (W) \times Subsidy\left(\frac{\$}{W}\right) \end{aligned} \tag{7}$$

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 carbons 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 2
Input parameter values and select model outputs for benefit-cost analysis of capital cost solar subsidy in Germany, Japan and California. NPV = Net Present Value.

	Residential solar price (2016 US$/W)	Electricity Price (2016 US$/kWh)	Solar resource (kWh/kW)	Residential solar NPV (2016 $/kW)	Carbon intensity of grid (g/kWh)	Subsidy expenditure per adoption (0.5$/W subsidy)
Germany	1.65	.32	900	2400	486	1.70
Japan	3.2	.22	1150	1380	572	1.55
California	4.3	.17	1550	240	489	1.38

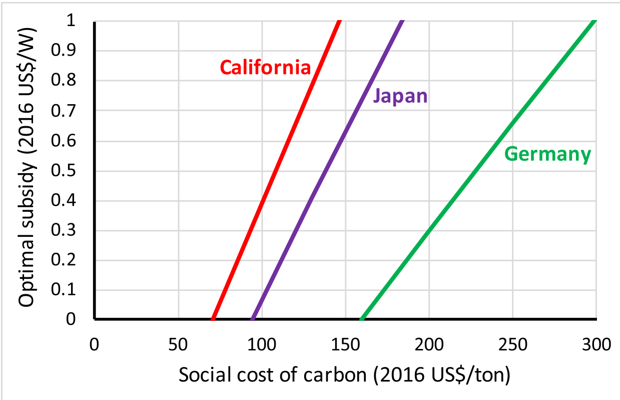

Fig. 3. "Optimal" single year subsidy for residential solar that maximizes net societal benefits (displaced CO_2 minus subsidy expenditure) in Germany, Japan and California as a function of the social cost of carbon (2016US $/ton) The optimum is determined by mapping a test subsidy level to adoption 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.

subsidy increases linearly with increasing social cost of carbon values, with slightly different slopes between the three regions.

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 describes adoption across multiple regions with only one explanatory variable. The model fits empirical data well and is attractive from a theoretical perspective because its mathematical form (error function) can be derived from a single plausible assumption: a normal distribution for additional customers as NPV increases. To compare and contrast our model with prior work, the only model of simpler mathematical form is the original Bass model, which fits adoption share directly to time series data [3]. The original Bass model however does not account for interventions (e.g. a subsidy) and needs to be calibrated separately to each region. Van Benthem and collaborators use a simple Logit model with NPV as an explanatory variable to describe solar diffusion in California [20]. Our error function model has easily interpretable regression parameters and far better empirical fit within and across regions. More complex multivariable regressions and agent-based models can explain how diffusion is influenced by other variables, e.g. education level [5]) or behavior of neighbors [18]. There are valuable applications of more complex models, but our simple (yet empirically compelling) model enables new analyses such as subsidy expenditure per induced adoption (Section 4) and optimal subsidy (Section 5). In addition, the compactness of our model facilitates its integration into larger energy systems models such as the National Energy Modeling System (NEMS) from the U.S. Energy Information Administration [21] and the World Energy Model (WEM) from the International Energy Agency [22]. These models are computationally intensive, covering large geographical areas and accounting for multiple interactions between energy technologies and the economy and can benefit from empirically-justified but computationally simple models.

Clearly, most consumers do not perform or even necessarily understand a Net Present Value calculation, as in Equation (1). How, then, could NPV be the basis for a model describing consumer choice? Solar installers do some form of calculation akin to NPV in presenting the value proposition to consumers. They may show the results in a different form (e.g. payback period or cash flow) but some form of similar measure is shown to consumers to inform their purchase decision. And while each consumer makes choices based on their particular circumstances and preferences, making individual prediction difficult, their aggregate behavior is easier to model.

While there is compelling empirical justification for the model we present, there are differences between predicted and actual diffusion. One reason for gaps between prediction and outcome arise from data limitations constraining implementation of the modeling concept. Region-aggregated variables (state, nation) were 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 is 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

Eric 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 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

[1] NREL, Solar Industry Update: Q4 2018/Q12019. NREL/PR-6A20-73992, National Renewable Energy Agency (NREL), Golden, USA, 2019a.
[2] LBNL, Tracking the Sun 11, Lawrence Berkeley National Laboratory (LBNL), Berkeley, California, 2018.
[3] M. Guidolin, C. Mortarino, Cross-country diffusion of photovoltaic systems: modelling choices and forecasts for national adoption patterns, Technol. Forecast. Soc. Chang. 77 (2) (2010) 279–296.
[4] B. Sigrin, M. Gleason, R. Preus, I. Baring-Gould, R. Margolis, Distributed Generation Market Demand Model (dGen): Documentation (No. NREL/TP-6A20-65231), National Renewable Energy Lab.(NREL), Golden, USA, 2016.
[5] T. Islam, Household level innovation diffusion model of photo-voltaic (PV) solar cells from stated preference data, Energy Policy 65 (2014) 340–350.
[6] K. van Blommestein, T.U. Daim, Y. Cho, P. Sklar, Structuring financial incentives for residential solar electric systems, Renew. Energy 115 (2018) 28–40.
[7] P. Zhai, E. Williams, Analyzing consumer acceptance of photovoltaics (PV) using fuzzy logic model, Renew. Energy 41 (2012) 350–357.
[8] C.L. Kwan, Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States, Energy Policy 47 (2012) 332–344.
[9] J. Letchford, K. Lakkaraju, Y. Vorobeychik, Individual household modeling of photovoltaic adoption, in: Energy Market Prediction: Papers from the 2014 AAAI Fall Symposium. DOE Sunshot Initiative, 2015.
[10] E. Drury, M. Miller, C.M. Macal, H. Donna, The transformation of Southern California's residential photovoltaics market through third-partyownership, Energy Policy (42) (2012) 681–690.
[11] M. Graziano, K. Gillingham, Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment, J. Econ. Geogr. 15 (4) (2015) 815–839.
[12] C.W. Hsu, Using a system dynamics model to assess the effects of capital subsidies and feed-in tariffs on solar PV installations, Appl. Energy 100 (2012) 205–217.
[13] M. Jimenez, C.J. Franco, I. Dyner, Diffusion of renewable energy technologies: the need for policy in Colombia, Energy 111 (2016) 818–829.
[14] L.L.C. Dos Santos, L.N. Canha, D.P. Bernardon, Projection of the diffusion of photovoltaic systems in residential low voltage consumers, Renew. Energy 116 (2018) 384–401.
[15] D. Noll, C. Dawes, V. Rai, Solar Community Organizations and active peer effects in the adoption of residential PV, Energy Policy 67 (2014) 330–343.
[16] C. Macal, D. Graziano, J. Ozik, Modeling Solar PV Adoption: A Social-Behavioral Agent-Based Framework. Energy Market Prediction, vols. 1–18, Association for the Advancement of Artificial Intelligence, 2014.
[17] H. Zhang, Y. Vorobeychik, J. Letchford, K. Lakkaraju, Data-driven agent-based modeling, with application to rooftop solar adoption, in: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, 2015, pp. 513–521.
[18] V. Rai, C. Reeves, R. Margolis, Overcoming barriers and uncertainties in the adoption of residential solar PV, Renew. Energy 89 (2016) 498–505.
[19] N. Mohandes, A. Sanfilippo, M. Al Fakhri, Modeling residential adoption of solar energy in the Arabian Gulf Region, Renew. Energy 131 (2019) 381–389.
[20] A. Van Benthem, K. Gillingham, J. Sweeney, Learning-by-doing and the optimal solar policy in California, Energy J. 29 (3) (2008) 131.
[21] EIA, The National Energy Modeling System: an Overview, Energy Information Administration (EIA), Washington D.C. (USA), 2018. https://www.eia.gov/outlooks/aeo/nems/overview/pdf/0581(2018).pdf.
[22] IEA, World Energy Model, International Energy Agency, Paris, France, 2018. https://www.iea.org/weo/weomodel/.
[23] NREL, PVWatts Calculator, National Renewable Energy Laboratory (NREL), Golden, USA, 2018. http://pvwatts.nrel.gov.
[24] M. Braun, K. Büdenbender, D. Magnor, A. Jossen, Photovoltaic self-consumption in Germany: using lithium-ion storage to increase self-consumed photovoltaic energy, in: 24th European Photovoltaic Solar Energy Conference (PVSEC): Hamburg, Germany, 2009.
[25] I. Kaizuka, Residential PV Market in Japan, RTS Corporation, Tokyo, Japan, 2014.
[26] World Bank, Consumer Price Index (CPI), World Bank, Washington D.C, 2019. http://data.worldbank.org/indicator/FP.CPI.TOTL.
[27] IMF, Interest Rates Selected Indicators, International Monetary Fund, Washington D.C, 2019. https://data.imf.org/regular.aspx?key=61545867.
[28] OECD, OECD Data: Exchange Rates, Organisation for Economic Co-operation and Development (OECD), Paris, 2019. https://data.oecd.org/conversion/exchange-rates.htm.
[29] IEA, National Survey Report of, PV Power Applications in Japan 2016 (Plus Previous Years), International Energy Agency (IEA), Paris, France, 2016b.
[30] IEA, KfW-Programme Producing Solar Power, International Energy Agency (IEA), Paris, France, 2019. https://www.iea.org/policiesandmeasures/pams/germany/name-23889-en.php.
[31] IEA, National Survey Report of PV Power Applications in Germany 2008 (And Previous Years), International Energy Agency (IEA), Paris, France, 2008.
[32] DOE, Residential Renewable Energy Tax Credit, U.S. Department of Energy, Washington, D.C, 2019. https://www.energy.gov/savings/residential-renewable-energy-tax-credit.
[33] DSIRE, Database of State Incentives for Renewables & Efficiency, North Carolina Clean Energy Technology Center, North Carolina (USA), 2019. http://www.dsireusa.org.
[34] LBNL, Tracking the Sun 10, Lawrence Berkeley National Laboratory (LBNL), Berkeley, California, 2017.
[35] NREL, Open PV Project Database, National Renewable Energy Agency (NREL), Golden Colorado, 2019b. https://openpv.nrel.gov/search.
[36] EIA, Electricity Data Browser, Average retail price of electricity, Energy Information Administration (EIA), Washington D.C., 2019. https://www.eia.gov/electricity/data/browser/#/topic/7?agg=0,1&geo=g&endsec=vg&linechart=ELEC.PRICE.US-ALL.A~ELEC.PRICE.US-RES.A~ELEC.PRICE.US-COM.A~ELEC.PRICE.US-IND.A&columnchart=ELEC.PRICE.US-ALL.A~ELEC.PRICE.US-RES.A~ELEC.PRICE.US- FSO 2018. (Accessed 23 December 2019).
[37] Census, American Community Survey, Physical Housing Characteristics for Occupied Housing Units, US Census Bureau, Washington D.C, 2019.
[38] CSS, Monthly Quarterly and Annual Statistics, California Solar Statistics (CSS), 2019. California, USA, https://www.californiasolarstatistics.ca.gov/reports/monthly_stats/.
[39] Fraunhofer, Photovoltaics Report, Fraunhofer Institute for Solar Energy Systems, Germany, 2019. https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf.
[40] BDEW, Electricity Price Analysis February 2017, Bundesverband der Energie-und Wasserwirtschaft (BDEW), German Association of Energy and Water Industries, Berlin, Germany, 2017.
[41] FSO, Census of Buildings and Housing, Federal Statistical Office (FSO), Wiesbaden, Germany, 2018.
[42] METI, Standardized Electricity Prices, Ministry of Economy and Trade (METI), Tokyo: Japan, 2016 (In Japanese).
[43] IEA, Purchasing Scheme for Solar PV Electricity, International Energy Agency (IEA), Paris, France, 2016a. https://www.iea.org/policiesandmeasures/pams/japan/name-24651-en.php.
[44] JETRO, Electricity and Renewable Energy, Japan External Trade Organization, (JETRO), Tokyo, Japan, 2016. https://www.jetro.go.jp/ext_images/en/invest/attract/pdf/en_2016_Energy.pdf.
[45] JPEA, PV Shipment in Japan, Japan Photovoltaic Energy Association (JPEA), Tokyo: Japan (2005-2016), 2016. http://www.jpea.gr.jp/en/statistic/index.html.
[46] MIAC, Japan Residential and Land Statistics 2013, Ministry of Internal Affairs and Communication (MIAC), Tokyo: Japan, 2013. In Japanese, http://www.stat.go.jp/data/jyutaku/2003/11.htm#p01Japan.
[47] Everett Rogers, Diffusion of Innovations, fifth ed., Simon and Schuster, 2003, ISBN 978-0-7432-5823-4.
[48] A.-N. Spiess, N. Neumeyer, An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach, BMC Pharmacol. 10 (6) (2010), https://doi.org/10.1186/1471-2210-10-6.
[49] W. Greene, The behaviour of the maximum likelihood estimator of limited

dependent variable models in the presence of fixed effects, Econom. J. 7 (2004) 98–119.
[50] R.J. Spiegel, M.R. Leadbetter, F. Chamu, Distributed grid-connected photovoltaic power system emission offset assessment: statistical test of simulated- and measured-based data, Sol. Energy 78 (6) (2005) 717–726.
[51] OECD, Climate Change Mitigation: Policies and Progress, Organisation for Economic Co-operation and Development (OECD), Paris, 2015, https://doi.org/10.1787/9789264238787-en.
[52] E.S. Rubin, I.M. Azevedo, P. Jaramillo, S. Yeh, A review of learning rates for electricity supply technologies, Energy Policy 86 (2015) 198–218.
[53] S. Herron, E. Williams, Modelling cascading diffusion of new energy technologies: case study of residential solid oxide fuel cells in the U.S. and internationally, Environ. Sci. Technol. 47 (15) (2013) 8097–8104.
[54] J. Jung, W.E. Tyner, Economic and policy analysis for solar PV systems in Indiana, Energy Policy 74 (2014) 123–133.
[55] B. Hoen, S. Adomatis, T. Jackson, J. Graff-Zivin, M. Thayer, G.T. Klise, R. Wiser, Selling into the Sun: Price Premium Analysis of a Multi-State Dataset of Solar Homes, Lawrence Berkeley National Laboratory, Berkeley, CA, 2014.
[56] W. Jager, Stimulating the diffusion of photovoltaic systems: a behavioural perspective, Energy Policy 34 (14) (2006) 1935–1943.
[57] V. Rai, B. Sigrin, Diffusion of environmentally-friendly energy technologies: buy versus lease differences in residential PV markets, Environ. Res. Lett. 8 (1) (2013), 014022.