ARTICLE IN PRESS
TFS-18784; No of Pages 15
Technological Forecasting & Social Change xxx (2016) xxx–xxx
Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.
Contents lists available at ScienceDirect
Technological Forecasting & Social Change

Forecasting residential solar photovoltaic deployment in California

Changgui Dong [a,b,c,*], Benjamin Sigrin [b], Gregory Brinkman [b]

[a] Renmin University of China, Beijing, China
[b] National Renewable Energy Laboratory, Golden, United States
[c] RUC National Academy of Development and Strategy, Beijing, China

ARTICLE INFO

Article history:
Received 22 March 2016
Received in revised form 15 November 2016
Accepted 26 November 2016
Available online xxxx

Keywords:
Solar PV
Forecasting
Diffusion models
Policy impact

ABSTRACT

Residential distributed photovoltaic (PV) deployment in the United States has experienced robust growth, and policy changes impacting the value of solar are likely to occur at the federal and state levels. To establish a credible baseline and evaluate impacts of potential new policies, this analysis employs multiple methods to forecast residential PV deployment in California, including a time-series forecasting model, a threshold heterogeneity diffusion model, a Bass diffusion model, and National Renewable Energy Laboratory's dSolar model. As a baseline, the residential PV market in California is modeled to peak in the early 2020s, with a peak annual installation of 1.5–2 GW across models. We then use the baseline results from the dSolar model and the threshold model to gauge the impact of the recent federal investment tax credit (ITC) extension, the newly approved California net energy metering (NEM) policy, and a hypothetical value-of-solar (VOS) compensation scheme. We find that the recent ITC extension may increase annual PV installations by 12%–18% (roughly 500 MW, MW) for the California residential sector in 2019–2020. The new NEM policy only has a negligible effect in California due to the relatively small new charges (<100 MW in 2019–2020). Furthermore, impacts of the VOS compensation scheme ($0.12 per kilowatt-hour) are larger, reducing annual PV adoption by 32% (or 900–1300 MW) in 2019–2020.
© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Renewable energy for electricity generation, led by wind and solar, is key to the global climate-change mitigation strategy (IPCC, 2014). In the United States, renewable energy now accounts for the biggest source of increase in generation capacity. In 2015, the country saw over 2.3 GW_{dc} of distributed-generation photovoltaics (DGPV)[1] connected to the grid, which corresponds to almost 200,000 installations (GTM/SEIA, 2014). In California, even after the California Solar Initiative (CSI) ended in 2014, the DGPV market remains viable (BNEF, 2015a; GTM/SEIA, 2015). The DGPV installations are determined by individual customers and installed by solar companies, which are often beyond the control of utility companies. However, such high DGPV growth has important implications on utility planning processes, especially in terms of future infrastructure needed, system cost minimization, and reliable operation of the electric system.

The rapid DGPV deployment has spurred policy debates and policy changes in many ways. Various studies have been conducted to understand the benefits and costs of DGPV to the grid (Blackburn et al., 2014; Denholm et al., 2014; RMI, 2013), how to adjust policy-support schemes (CPUC, 2015b; Rábago et al., 2012; Randazzo, 2015), and the implication on utility business models (Lehr, 2013; Richter, 2013; Satchwell et al., 2014). Recently at the federal level, the U.S. Congress (U.S. Congress, 2015) approved a five-year phase-down extension of the investment tax credit (ITC) for solar energy. In California, a February 2016 decision by the California Public Utilities Commission (CPUC) on the successor to net energy metering (NEM) maintained retail electricity rates for DGPV owners at previous levels, while introducing new tariff features—a system interconnection fee, a non-bypassable charge, and a minimal monthly bill—that anticipate higher levels of DGPV deployment.

To analyze the impact of enacted or future policy actions, it is necessary to have a valid forecast for the future technology diffusion; however, forecasting DGPV deployment is fundamentally challenging for several reasons. First, the DGPV market has historically relied on policy support (DSIRE, 2016), so future policy changes create forecasting uncertainty. Second, DGPV is a new and durable technology; so, many non-economic factors may influence adoption, such as customers' environmental attitudes, peer effects, and risk preferences (Rai et al., 2016). Finally, even forecasting future DGPV prices alone is challenging because of the global nature of the technological change and supply chain (Choi and Anadón, 2014).

The objectives of this research are to model future residential PV deployment in California leveraging a suite of techniques and then use our

* Corresponding author at: Renmin University of China, Beijing, China.
E-mail address: rosenbloog@gmail.com (C. Dong).

[1] DGPV is in comparison with utility-scale PV. In this research, we define DGPV as PV systems that are generally small (<5 MW), connect to the distribution network (rather than the transmission network), and are either behind the meter or in front of it and connected to the low-voltage distribution network.

http://dx.doi.org/10.1016/j.techfore.2016.11.021
0040-1625/© 2016 Elsevier Inc. All rights reserved.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

Nomenclature

ACS	American Community Survey
ARIMA	autoregressive integrated moving average
BNEF	Bloomberg New Energy Finance
CEC	California Energy Commission
CPUC	California Public Utilities Commission
CSI	California Solar Initiative
CSS	California Solar Statistics
DGPV	distributed-generation photovoltaics
DSIRE	Database of State Incentives for Renewables & Efficiency
dSolar	Distributed Solar Market Demand model
E3	Energy and Environmental Economics
EIA	Energy Information Administration
GTM	Greentech Media
GW_{DC}	gigawatts in direct current
IOUs	investor-owned utilities
ITC	investment tax credit
kWh	kilowatt-hour
LHS	Latin hypercube sampling
MW	megawatts
NEM	net energy metering
NEMS	National Energy Modeling System
NPV	net present value
NREL	National Renewable Energy Laboratory
OOH	owner-occupied houses
SEIA	Solar Energy Industries Association
STD	standard deviation
TPO	third-party owned
TTS	Tracking the Sun
VOS	value-of-solar

baseline results to study the impact of potential policy changes on PV deployment. Without a solid baseline forecast of PV adoption, the assessment of policy impacts could be biased from the very beginning. As California and other states in the U.S. are progressing in making policy changes related to DGPV, establishing common ground for forecasting future PV adoption should be critical during the policy-making process. That is why in this research we include multiple forecasting methods, covering both top-down and bottom-up models.

Existing forecasting of future PV adoption usually comes from three sources: research institutes, industry experts, and utility companies. However, the methods they have used vary dramatically. Research institutes and industry experts often rely on bottom-up customer adoption models, whereas utility companies sometimes simply assume an end-point DGPV deployment level or extrapolate from historical data (Mills et al., 2016). Even for the bottom-up models, they differ in specific model configuration, parameter setting, and geographic resolution. For instance, the geographic resolution could range from nationwide, to state-level, to utility area; however, few studies go to county levels. Another issue with these forecasts is that they tend to have a limited time frame, usually not exceeding 2030.

Our work extends the PV forecasting literature in several ways. First, we use and compare multiple forecasting techniques. We leverage not only bottom-up models, but also top-down models. Although all bottom-up models are based on the classical Bass diffusion model, we also introduce a different type of diffusion model, i.e., the threshold-heterogeneity model. These two types of models represent two fundamentally different views on how to generate an S-type diffusion curve. Second, in our bottom-up model, we cover the two major business models of DGPV, i.e., the customer-owned and the third-party-owned (TPO) systems, whereas most other literature does not include the most recent DGPV business model—TPO. Third, we make our forecasts for California at the county level and cover periods from now until 2050. Lastly, before making forecasts for 2050, we carefully calibrate our methods based on the extensive DGPV historical data in California.

The top-down models and bottom-up model used in this research are as follows. First, following the forecasting literature (Forte, 2015; Hyndman and Athanasopoulos, 2013) and diffusion[2] of innovation literature (Bass, 1969; Bass et al., 1994; Bemmaor, 1994), we build three top-down models that are based on theory and calibrated with real market data: a time-series forecasting model, a threshold-heterogeneity diffusion model, and a Bass diffusion model. Then, we compare these models to a more complex bottom-up techno-economic model, the dSolar model maintained by the National Renewable Energy Laboratory (NREL).[3] Furthermore, as a demonstration of uncertainties in forecasting DGPV technology diffusion, we conduct sensitivity analysis based on the dSolar model around certain key economic parameters. For simplicity, we only focus on the largest residential market segment in this research, i.e., the owner-occupied housing (OOH) market, rather than the non-OOH market or PV adoption in the commercial sector.

After first reviewing the relevant literature (Section 2), we discuss our data inputs and four methods (Section 3). We then present our baseline results for California's residential PV sector in Section 4. We use the dSolar model and the threshold-heterogeneity diffusion model to conduct three policy scenario analyses of the recent federal ITC extension, the newly approved California NEM policy, and a hypothetical value-of-solar compensation scheme (more detail in Section 4.4). Section 5 provides a conclusion.

2. Literature review

This research builds on several strands of literature: general forecasting, diffusion of innovation, PV financial attractiveness, and existing DGPV forecasting models. General forecasting is essential for planning purposes, and the methods can be very simple, such as using most recent observation as a forecast or developing a complex model such as neural networks (Forte, 2015; Hyndman and Athanasopoulos, 2013). When time-series data are available, two of the most popular univariate time-series models are the autoregressive integrated moving average (ARIMA) model and the exponential smoothing model. ARIMA focuses on the autocorrelations in the data, whereas the exponential smoothing model detects trends and seasonality in the time series (Hyndman and Athanasopoulos, 2013; Hyndman et al., 2008). Time-series forecasting uses only historical information for the variable being forecasted and assumes that the observed trend and seasonality will continue. As such, time-series models are easy to implement and rely on only one assumption (i.e., continuing trend) to work. Nevertheless, time series method can miss other external factors that affect the variable of interest, such as policy changes or changes in the sub-populations considering adoption.[4]

The literature on diffusion of innovation is vast and good review work can be seen in Sultan et al. (1990), Meade and Islam (2006), and Rogers (2003). Diffusion of innovation models started in late 1960s, with the Bass diffusion model (Bass, 1969) probably being the most commonly used model to predict technology adoption. Generalized Bass models have been proposed to incorporate other variables such

[2] Technology diffusion refers to the process of how new technologies spread throughout society over time. This paper uses the terms "diffusion," "adoption," and "deployment" interchangeably.

[3] We consider these two types of models complementary to each other (Section 3), and together they establish a more reliable baseline for future PV deployment in the sense that they represent two extremes of modeling efforts: the top-down models are generally easy to implement and require much fewer assumptions, whereas the bottom-up models are more modular and have more assumptions embedded, but are more flexible in modeling PV economics.

[4] Only future values of those external factors are relevant here, because their historical values should already be incorporated into the historical values of the variable of interest.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

as technology price and advertising expenditure (Horsky and Simon, 1983; Robinson and Lakhani, 1975); however, Bass et al. (1994) showed that the classic Bass model still fits the data well without these variables, and Bottomley and Fildes (1998) found that including price information barely improves the forecasting accuracy. Therefore, most current analysis continues to use the classic Bass diffusion model, which has been used by several PV forecasting models (CEC, 2015; EIA, 2013; Sigrin et al., 2016).[5] The Bass diffusion model is a mixed-influence model based on Meade and Islam's (2006) classification: a mix of both internal and external influence. The model posits that technology diffusion is fundamentally driven by two groups in the market: innovators and imitators, which represent the internal and external forces, respectively. The driving force behind imitators is the so-called peer effect or word-of-mouth effect, which has been consistently found in the DGPV market (Bollinger and Gillingham, 2012; Noll et al., 2014; Rai and Robinson, 2013).

The S-shaped diffusion curve can be derived even without requiring the existence of the word-of-mouth effect, as long as heterogeneity exists in people's propensity to adopt (Bemmaor, 1994). The threshold-heterogeneity diffusion model is premised on a similar idea that different people in a market have heterogeneous thresholds (i.e., reservation price) to overcome before adopting a new technology. If the thresholds for all consumers in a market follow a normal distribution, a falling technology price or linearly moving threshold (from high to low levels) will result in increasingly more people adopting the technology until the market saturates. The final result is an S-type diffusion curve (Dernburg, 1957; Duesenberry, 1949; van den Bulte and Stremersch, 2004). In this paper, we include both types of diffusion models because peer effects and varying thresholds have both been found in the DGPV markets (Bollinger and Gillingham, 2012; Rai and Robinson, 2015; Robinson and Rai, 2015). Furthermore, the threshold-heterogeneity diffusion model has not been used previously in the DGPV forecasting literature.

The literature on PV financial attractiveness consists of at least three sources: PV cost projection, PV payback or net present value (NPV), and value of solar (VOS). First, PV cost projection can be done either through an experience curve (or learning-by-doing curve) or via expert elicitation. While the PV experience curve predicts PV cost based on past cumulative installed capacity (de La Tour et al., 2013; Nemet, 2006), expert elicitation derives future PV cost based on experts' judgment (Bosetti et al., 2012). Second, PV payback or NPV have been studied extensively in the past based on various assumptions on key financial parameters (Blair et al., 2014; Gillingham et al., 2014; Rai and Sigrin, 2013). This research adopts the standard approach of calculating PV economics developed by national laboratories in the United States. Third, the VOS has become a new metric to measure the contribution of DGPV to the grid system and provide an appropriate subsidy level to PV customers (Farrell, 2014; Rábago et al., 2012; Taylor et al., 2015); therefore, this research builds on this relevant literature and evaluates the potential impact of this new compensation scheme on PV adoption.

Lastly, several DGPV forecasting models exist in the literature, including NREL's dSolar model (a successor to the SolarDS model; see Section 3.5 for more detail) and the U.S. Energy Information Administration's National Energy Modeling System (NEMS) residential model, which relies on the payback period and the number of years that the technology has been available to forecast deployment (EIA, 2013). On the industry side, Bloomberg New Energy Finance (BNEF) and Greentech Media (GTM) have developed short-term DGPV forecasting models, although their methods are not publicly available (BNEF, 2015a; GTM/SEIA, 2015). For several years, the California Energy Commission (CEC) has produced the California Energy Demand report, which includes a 10-year DGPV state forecast based on methods used in the SolarDS and NEMS models (CEC, 2015). The CPUC has recently contracted Energy and Environmental Economics (E3) to develop a NEM Successor Tariff Public Tool to forecast PV deployment until 2025 and evaluate the impact of policy changes on deployment (E3, 2015). Almost all of these forecasting models are bottom-up techno-economic models that map the economics to PV adoption (see, for example, Section 3.5); however, they tend to differ in model configuration, parameter setting, and geographic resolution. In this paper, we leverage the dSolar model as it provides both a highly geographically-resolved picture (on the county level) and a long-term view (till 2050) on future PV diffusion trends. We future compare the dSolar results with other bottom-up models when their results are available.

3. Data and method

This section first presents the various data sources used in this research and summary statistics for variables used in this research, such as socio-demographic factors. We then describe the four main models used in our analysis: a time-series forecasting model, threshold-heterogeneity diffusion model, Bass diffusion model, and the dSolar model. These methods were selected because they are commonly used in the forecasting literature of new technology adoption (Armstrong, 2001). We discuss the essence of each method, the intuition behind them, and their applicability in the case of PV adoption.

3.1. Data

We collected data, including CSI incentive application and PV system characteristic data, from California Solar Statistics (CSS) and collected interconnection data from the state's three major investor-owned utilities (IOUs). We then combined these data with PV installation records from California publicly owned utilities—which are part of the Tracking The Sun (TTS) data set collected by Lawrence Berkeley National Laboratory (Barbose and Darghouth, 2015)—and we cleaned all the data[6] to produce a complete installation time series for each California county.

For the threshold-heterogeneity diffusion model, we also needed data to calculate the NPV of historic DGPV investments so that we could characterize the profitability of PV investments over time. We combined the consumer benefit values (mostly from incentives and bill savings) calculated in Gillingham et al. (2016) with the price data in Barbose and Darghouth (2015) to determine NPV, which uses a 7% discount rate and assumes no financing is involved.[7]

Assumptions of future capital cost changes are exogenous to demand and do not consider learning effects. Specifically, in our base case, we use U.S. Department of Energy (DOE) SunShot program goals (DOE 2012) of reaching a project turn-key cost of \$1.5/W_{DC} in \$2010 for residential systems in 2020, and are constant in real terms thereafter. Three additional transformations are applied: 1) an 8.2% project adder to account for regional costs of doing business in California (EIA, 2013); 2) Costs are expressed in \$2014, using the Consumer Price Index; and 3) A scaling factor is applied to account for economies of scale, assuming a reference system size of 4 kW, where installed costs increase (decrease) by 1.67% per kW smaller (larger). We further test

[5] Although the more complex Bemmaor-type model (Bemmaor, 1994; Bemmaor and Lee, 2002) can nest the Bass model, Bemmaor and Lee (2002) also found that the latter performed better in the long-term forecasting and there were penalties for the model complexity.

[6] Major data-cleaning steps included the following in order for us to produce the most complete PV installation history at the county level. For the CSS data, we deleted project-level data entries with both missing utility names and county names, duplicate interconnection application records, and misspelled county names that we cannot reasonably fix; we then combined the customer sectors into either residential or non-residential categories, imputed the missing county names with known utility names, and imputed the system size in DC based on the AC number if missing. For the TTS data, we deleted duplicate data entries, cleaned the county and utility names, and combined the customer sectors into either residential or non-residential categories. After the data-cleaning process, our state-level installation results matched well with the industrial source by GTM.

[7] Although this assumption introduces downward bias, this bias is common for all counties; thus, it might not be substantial if only this NPV is used to explain county-level differences in PV diffusion.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

the sensitivity of our forecasting results to the technology cost assumption.

Our analysis also examines the impact of socio-demographic factors. We collected socio-demographic data such as household income and education at the census tract level from the U.S. Census American Community Survey (ACS), and we mapped all CSI PV installations to the corresponding census tracts.[8] Table 1 summarizes the socio-demographic and CSI data for variables that we end up using in our regression analysis. The CSI data include the total number of installations and average NPV, both at the tract level. We use the socio-demographic variables to understand the PV diffusion patterns across California counties together with the PV penetration and NPV variables to conduct the scenario analysis using the threshold-heterogeneity diffusion model.

3.2. Time-series forecasting model

The ARIMA model and the exponential smoothing model produce very similar results, so we only use ARIMA in this paper. ARIMA is an extension of the autoregressive moving-average (ARMA) model in time-series analysis. Without seasonality, the model ARIMA(p,d,q) is specified as follows:

$$(1-\phi(L))\left(1-L^d\right)y_t = (1+\theta(L))\varepsilon_t, \tag{1}$$

where $\phi(\cdot)$ and $\theta(\cdot)$ are polynomials of the lag operator L with order p and q respectively; d is the number of unit roots in the polynomials of the observed time series y_t; and ε_t is the white-noise error term. Because of the existence of d, ARIMA models can fit non-stationary time series in a similar manner to how exponential smoothing models can.[9] Once the ARIMA model is estimated, the time-series forecasting becomes straightforward.[10] One way to understand ARIMA models is that, for the univariate time series, the current value is a function of the last p values of itself, the last q random shocks (through the error term), and the trend captured by d.

Linking back to our county-level PV installation time-series data, if the selected ARIMA model finds some non-stationary trend in the data, then the model assumes that the trend will continue, usually in a linear fashion. However, PV installation cannot increase forever. Thus, we used time-series forecasting to demonstrate a range of model types and to show the incremental value of other more complex models in this context. Similarly, machine-learning methods such as a neural network can be insufficient for modeling changes in external forces that impact DGPV diffusion.

3.3. Threshold-heterogeneity diffusion model

The threshold-heterogeneity diffusion model states that different people in a market have heterogeneous thresholds (i.e., reservation prices) to overcome before adopting a new technology. If the market price is lower than a consumer's threshold, it becomes beneficial for this consumer to adopt. Though the installed price of PV has declined over time, California's PV subsidies have also decreased, which means that it is plausible that Californians with higher payoffs or NPV (not necessarily those with higher thresholds or higher reservation price) are the first adopters.[11] When technology cost drops, people who previously could receive lower payoffs start to see benefits from installing PV.

Assuming the payoff from adopting DGPV at a certain time point follows a bell-shaped distribution, the diffusion process can be seen as either a combination of a fixed payoff distribution and a moving threshold or a fixed threshold and a moving (but identical) distribution. In Fig. 1, Panel A shows the fixed payoff distribution with moving threshold; at year t, the threshold is at position $p1$, and at year $t+n$, the threshold moves to position $p2$. With a lower threshold, more people decide to adopt the new technology, even with a fixed payoff. Panel B shows the fixed threshold with moving payoff distribution; from year t to year $t+n$, the payoff distribution shifts gradually to the right while the threshold position remains the same. The shaded areas in both cases represent the equivalent cumulative percentages of people who have adopted the new technology. The final S-shaped diffusion curve comes from the cumulative density function of the normal distribution, assuming the movement of either the threshold (in Panel A) or the payoff distribution (in Panel B) is linear in time.[12]

We estimate the payoff distribution in a market by fitting a truncated normal or Gamma distribution to the penetration-level data and deriving the distribution parameters. Because PV penetration levels are low, generally below 10%, only a small tail of the payoff distribution can be observed with empirical data; thus, we need to estimate the parameters from this "truncated" distribution. The tail of the payoff distribution is obtained by dividing PV installation values by the market potential, another parameter we are going to estimate in the data-fitting process.

Researchers have devised many ways to fit a truncated distribution,[13] though most of the fitting methods do not work well in this case. As an alternative, we use Latin hypercube sampling (LHS) to search for the optimal parameter values based on a cost function that evaluates how well the model fits the data. Three parameters are included: the mean and standard deviation of the normal payoff distribution, and the maximal market potential (i.e., the final number of PV adopters to convert the historical number of adopters to penetration levels). LHS requires the lower and upper bounds for each parameter, as well as the number of samples to be drawn from the parameter space; with those, LHS will divide the parameter range by the number of samples and ensure there is only one sample from each row and column. This type of sampling method has been shown to be more efficient than simple Monte Carlo sampling (Carnell, 2012).

3.4. Bass diffusion model

The Bass diffusion model can be specified in either the continuous time domain or the discrete time domain. For the continuous time domain,

$$\frac{f(t)}{1-F(t)} = p + q \cdot F(t), \tag{2}$$

where $f(t)$ is the fraction of the population that adopts PV at time t, $F(t)$ is the cumulative installed fraction, p is the innovation coefficient, and q is the imitation coefficient. If $p=0$, the Bass diffusion model reduces to the logistic growth model, and if $q=0$, it reduces to an exponential distribution curve.

[8] Only the CSI data set contains street address information that can be geocoded at the tract level. So, we only use the CSI data to conduct the regression analysis described in Section 4.4.

[9] The ARIMA model in Eq. (1) can be extended to incorporate seasonality, with additional polynomials and unit roots parameterized by another set of $(P,D,Q)_m$, where m is the seasonal frequency. See Hyndman et al. (2008) for details.

[10] We use the R package "forecast" for this.

[11] At the beginning of the technology diffusion process, some people may adopt the new technology owing to other motivations (e.g., environmental benefits in the context of DGPV), and this subjective aspect of utility can be factored into the payoff calculation.

[12] In reality, there is no way to prevent both the threshold and the payoff distribution from moving. If both movements are linear in time, for simplicity purposes, only one needs be modeled. Furthermore, other distributions, such as the Gamma distribution, can simulate the S-curve, but the Gamma distribution requires the threshold (or payoff distribution) movement be exponential in time. Results from using the Gamma distribution are slightly worse than those based on the normal distribution in terms of fitting the historical data.

[13] In R, there are user-written functions to fit truncated distributions, including "fitdist," "fitTail," and "gamlssML."

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

Table 1
Summary statistics of tract-level diffusion and socio-demographic variables.

Variable name	Variable description	Min	Max	Mean	STD
count_lg	logarithm of installation counts	0	5.714	2.335	1.219
pene_level	penetration level: installation counts divided by number of OOH	0.001	0.244	0.030	0.026
npv_pw	net present value of PV in $/W	−6.570	6.300	0.665	1.020
inc_100to150k	% of OOH with income between $100,000 and $150,000	0	1.000	0.189	0.154
edu_bachorhigh	% of OOH (highest educated) with bachelor degree or higher	0	1.000	0.381	0.225
value_over1mil	% of OOH with housing values over $1 million	0	0.976	0.068	0.158
value_500kto1mil	% of OOH with housing values between $500,000 and $1 million	0	0.975	0.239	0.252
race_white	% of OOH (family heads) that are white	0	1.000	0.702	0.214
median_rooms	median number of rooms	1.4	9.000	5.907	0.826
built90s	% of OOH built from 1990 to1999	0	1.000	0.097	0.132
mortgage2orHE	% of OOH with either a second mortgage or home-equity loan	0	0.560	0.169	0.080
hh_size	mean household size	1.146	5.489	2.845	0.614
Electricity	% of OOH with electricity as a heating source	0	1.000	0.186	0.123
wood	% of OOH with wood as a heating source	0	0.837	0.021	0.061
mortgage_over40	% of OOH with mortgage over 40% of household income	0	1.000	0.470	0.186

N = 7305 tracts; OOH = owner-occupied houses.

Eq. (2) is essentially a first-order ordinary difference equation, and we can solve $f(t)$ as:

$$f(t) = \frac{(p+q)^2}{p} \frac{\exp(-(p+q)t)}{\left(1+\frac{q}{p}\exp(-(p+q)t)\right)^2} \tag{3}$$

and $F(t)$ as:

$$F(t) = \frac{1-\exp(-(p+q)t)}{1+\frac{q}{p}\exp(-(p+q)t)}. \tag{4}$$

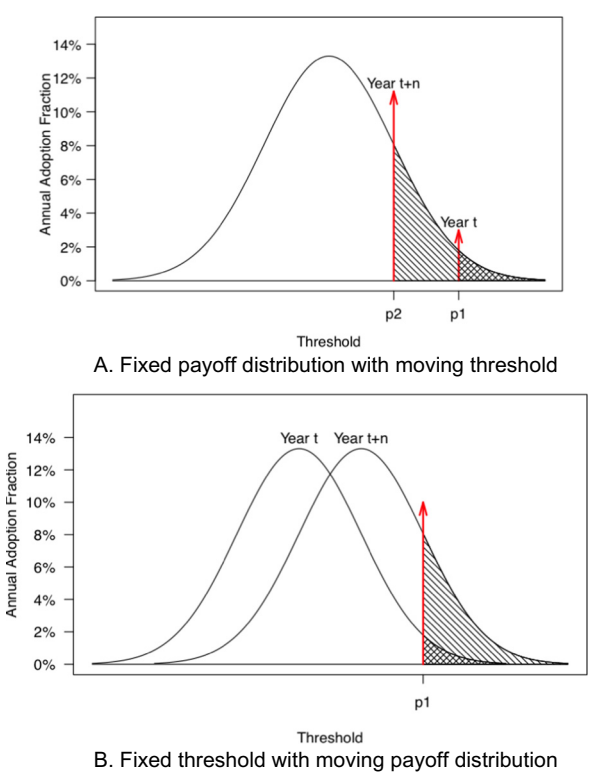

Fig. 1. Two ways to understand the threshold-heterogeneity diffusion model Shaded areas represent cumulative adoption fractions at year t and $t+n$; p1 and p2 correspond to two threshold positions.

Another interesting result is the peak time as a function of the p and q values:

$$t_{peak} = \frac{\ln q - \ln p}{p+q}. \tag{5}$$

For the discrete time domain,

$$f(t) = \left(p + q \cdot \sum_1^{t-1} f(t)\right)\left(1 - \sum_1^{t-1} f(t)\right), \tag{6}$$

where $f(t)$ has the same connotation as the continuous case, and the main predicting variable on the right-hand side is the cumulative adoption fraction until the last period. Multiplying Eq. (6) by the potential maximum market size m on both sides and collecting terms gives:

$$S(t) = p * m + (q-p) * Y_{t-1} - \left(\frac{q}{m}\right) * {Y_{t-1}}^2 \tag{7}$$

where $S(t) = m * f(t)$ is the current-year adoption number, and $Y_{t-1} = m * \sum_1^{t-1} f(t)$ is the previous-year cumulative adoption number.

We use both the discrete and continuous versions of the Bass diffusion model to fit the historical data and recover the three parameters: p, q, and m. We use ordinary least squares to fit Eq. (7) and obtain initial parameter values as inputs to fit Eq. (3) using non-linear least squares (Srinivasan and Mason, 1986).[14] When the estimation methods (ordinary least squares and non-linear least squares) do not produce statistically significant results,[15] we rely on LHS to search efficiently for the optimal parameter values again. In addition to the two coefficients p and q above, we also search for the optimal market size parameter m to obtain $f(t)$ based on the installation numbers that we have for each market.[16]

[14] As pointed out by Schmittlein and Mahajan (Schmittlein and Mahajan, 1982), a small bias is present in using ordinary least squares on discrete time series.

[15] One major reason for such insignificance is that because we have not yet seen the peaking year in PV installations, there could be multicollinearity issues between the cumulative sales and its square term (Heeler and Hustad, 1980).

[16] To increase the robustness of the LHS approach, we take several additional steps. First, we carefully choose the lower and upper bounds for each parameter. Based on Sultan et al. (1990), the average value of the coefficient is 0.03, and the average value is 0.38. Jeuland (1994) found that is usually 0.01 or less, and rarely >0.5. To cover broad enough parameter space, we set the lower bound for both coefficients at zero, the upper bound at 0.2, and the upper bound at 0.8 (we further reduce the upper bound below). With respect to the maximum market size parameter, we set the lower bound to be the cumulative DGPV installation value until Q1 2015 and the upper bound to be the number of owner-occupied housing units in the market. Second, based on the lower and upper bounds for selected parameters, we draw multiple sets of large random samples using LHS. For each draw, we set the sample size at 10,000, which improves the model fitting significantly compared to a sample size of 1000. We repeat this process 10 times using different random seeds to cover the parameter space more completely. Third, once we obtain certain parameter values from the first two steps, we keep shrinking the parameter space iteratively and search for the optimal parameters once again. After these three steps, we arrive at our best estimate of the model parameters.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

Note we use different time resolutions and cost functions to estimate the threshold-heterogeneity and Bass diffusion models. Because both approaches are trying to fit the same historical data using different distributions, differentiating them enables exploration of different aspects of the data. Thus, the threshold-heterogeneity diffusion model uses annual installation data with a cost function focusing on model fit on the annual adoption, whereas the Bass diffusion model uses quarterly data with a cost function focusing on model fit on the cumulative adoption. Although annual data are appropriate for the threshold-heterogeneity diffusion model, because the annual penetration level is essentially the density function in Fig. 1, quarterly data could provide more degrees of freedom in estimating the Bass diffusion model.

3.5. dSolar model

The Distributed Solar Market Demand model (dSolar) is a geospatially rich bottom-up market-penetration model that uses a Bass-like model to simulate DGPV deployment for residential, commercial, and industrial customers in the continental United States in two-year intervals from 2014 through 2050. As an upgrade of the SolarDS model (Denholm et al., 2009), dSolar includes better representation of customer decision making through an agent-based framework and increased use of spatially resolved data sets.

Market diffusion of DGPV is modeled through an agent-based approach that includes five overarching steps (Fig. 2). To represent accurately the population variation in the attributes influencing adoption propensity, the model first uses a statistical framework to sample agents (potential customers) for every U.S. county. Attributes assigned include agent location, solar resource (George et al., 2007), and annual electricity consumption (EIA, 2009; EIA, 2015) as well as a weighting factor corresponding to the number of similar customers that each agent represents in a county.[17] After generating agent types, dSolar establishes upper bounds on the technical potential of adoption by restricting residential owner-occupied houses to be single families only, excluding sites that are technically infeasible and requiring that agents offset a certain average percentage of their load by installing DGPV.[18]

Next, the potential economics for each adopter is determined using detailed discounted cash-flow analysis to determine the profitability (e.g., the payback period for customer-owned systems and monthly electricity bill savings for third-party-owned systems) over the system's lifetime. This approach assumes that the PV value is created by reducing the electricity bills the agent would have paid had they not adopted. Costs include initial capital costs, such as the down payment, monthly loan/lease payments, and annual operation and maintenance requirements (DOE, 2012). Revenue includes energy bill savings, applicable financial incentives, and tax-based credits, such as depreciation and interest rate deductions for nonresidential customers.

Considerable focus is given to representing regional differences in economic drivers accurately. Incentives are based on the Database of State Incentives for Renewables & Efficiency (DSIRE) database (DSIRE, 2016) and net-metering policies based on the IREC (Interstate Renewable Energy Council) Freeing the Grid data set (Barnes et al., 2013). Retail rates are based on the Utility Rate Database (OpenEI, 2015), and bill savings are calculated using the NREL System Advisor Model (Blair et al., 2014) based on an hourly time series of expected system generation and energy consumption. Using an hourly generation/consumption profile is important for capturing nuances in bill savings that arise from complex rate structures, such as tiered rates, demand charges, or time-of-use rates.

The final DGPV adoption arrives in two steps. First, we derive the maximum fractions of eligible agents who would eventually adopt DGPV for different levels of PV economics, differentiating between the customer-owned and the TPO market segments based on NREL's previous survey results (Fig. 3, Sigrin and Drury, 2014). These fractions can be considered as a cumulative density function of the underlying data while not as parameter estimates. They are then applied to the technical potential that we derived previously. Next, we use a carefully parameterized Bass diffusion curve to obtain the annual installation numbers (same and fixed for all counties in California); for each year, we observe the market penetration level from the previous year and calculate the corresponding equivalent years of market diffusion based on the parameterized Bass diffusion curve, and then we move the time scale by two more years. The result is the biannual market adoption of DGPV for a county. To ensure the validity of our results, we calibrate the dSolar results to those from GTM at the state-level in 2014 and 2016.[19]

Although dSolar uses the Bass diffusion curve to calculate PV adoption, it differs from the Bass diffusion model in how it obtains information on the maximum market size. Models like dSolar usually resort to the willingness-to-pay literature (Kastovich et al., 1982; Paidipati et al., 2008; R.W. Beck and Inc, 2009) and generate the maximum market-share curve corresponding to certain payback times, assuming an inverse exponential or inverse logistical relationship between them. Multiplying the maximum market share by technical potential numbers then gives the maximum market size. On the other hand, in our Bass diffusion model, we estimate a maximum market size (the parameter m) directly (together with the p and q values), which is not necessarily a function of payback. Even in the generalized Bass model literature, variables such as product price or payback are in addition to the parameters p, q, and m.

4. Results

This section first presents our baseline residential DGPV deployment forecasting results at the state and county levels using the four tools described above: time series, threshold, Bass and dSolar models, in terms of both the number of installations and installed capacity. We compare results from other sources where possible. Our baseline forecasts assume no future major policy changes in California as of January 22, 2016: we only included the ITC extension in our baseline results, but not the new NEM policy in California. Similarly, we did not include the impact of California's Zero Net Energy (ZNE) building goals that would require all new residential homes have solar starting in 2020, which represents roughly 410 MW installed capacity per year (E3, 2015).

The baseline results are followed first by a sensitivity analysis around three key economic parameters and then by a scenario analysis on the impact around the recent ITC extension, the new California NEM policy, and a hypothetical value-of-solar (VOS) compensation scheme. The VOS compensation scheme has been implemented in several utilities (Austin Energy as a prominent example), and also been discussed in the proposal stage of the California's new NEM policy. We model the potential impact of such policy and compare it with the new NEM policy.

4.1. State-level residential PV deployment

Almost all technology adoption forecasting models presume an S-shaped diffusion curve; adoption starts slowly, speeds up, and finally slows down again until it saturates the whole market. Fig. 4 shows cumulative-installation forecasting results based on the four methods

[17] The agent-level representation creates the potential to incorporate social demographics or other customer attributes in the adoption likelihood, which could highlight the within-county household heterogeneity, for example, between early and later adopters.

[18] Examples of technical filters include zoning restrictions and roof characteristics. We use a load-offset percentage of 95% for people with NEM based on Davidson and Margolis (2015), who use a database of system quotes received by customers that lists expected system generation and the customer's prior-year consumption. On the other hand, for people without NEM, we assume a lower load-offset percentage of 66%, which is a middle load-offset value used in E3 (2015).

[19] The sensitivity of the dSolar results to key economic variables can be found in Gagnon and Sigrin (2016).

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

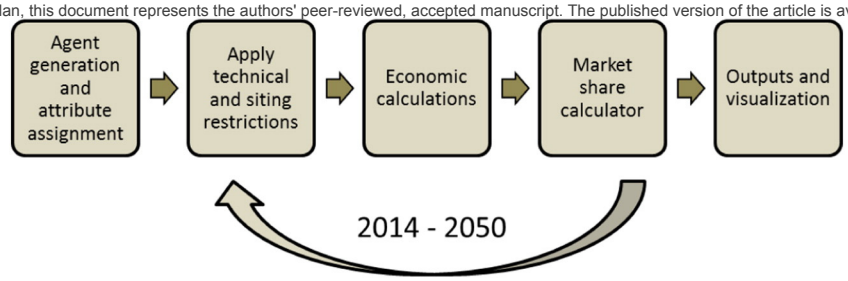

Fig. 2. dSolar model structure.

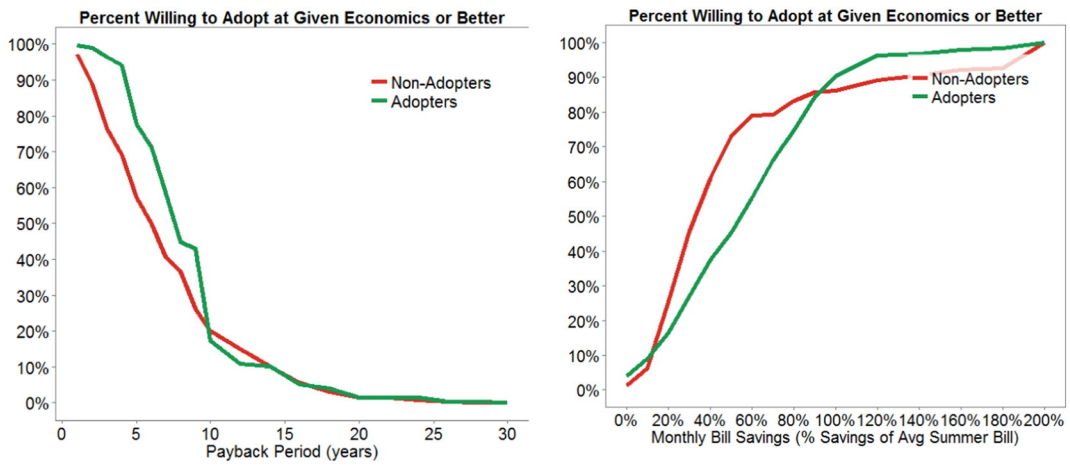

Fig. 3. Customer willingness to adopt for a given level of economic benefit or better: payback period (left for customer-owned market) and percentage of monthly bill savings (right for TPO market) In the legends, Non-adopters and Adopters refer to whether the correspondents in the survey have adopted PV.

outlined in Section 3, along with another forecast by E3 and the historical installation trend, at a time interval of two years, which is the time resolution of the dSolar model. The figure also shows the results as percentages of 2013 owner-occupied houses in California.[20]

Our four models fit the historical installation trend well but begin to diverge after 2016. The obvious outlier is the time-series forecasting method, which rises linearly at the historical rate. The time-series model is the simplest to constrain at longer time scales because it does not consider time-varying economic conditions or long-term market saturation; this suggests the incremental value of using other more complex modeling techniques.[21]

The threshold-heterogeneity and Bass diffusion models produce S-shaped curves, as an improvement on the time-series method. The maximum market size is about 6.8 million installations (98% of 2013 OOH) for the threshold-heterogeneity model and 5.3 million (76% of 2013 OOH) for the Bass diffusion model. Although these market sizes may appear large, we note that the estimate is based on assuming no changes in current NEM (i.e., PV being compensated indefinitely at the retail rate) or retail rate policies, which are already favorable to DGPV.[22] One reason for the difference in the Bass and threshold models is that the Bass diffusion density curve ($f(t)$ instead of $F(t)$) has a higher kurtosis than the normal curve; thus, after the peaking year, it forecasts much less adoption compared to the threshold-heterogeneity diffusion model.

Because the dSolar model relies in part on the Bass diffusion model, its cumulative installation curve looks similar to the Bass diffusion model's curve, though at a smaller scale. In 2050, dSolar predicts 4.8 million cumulative installations in California (70% of 2013 OOH). The reduced deployment in dSolar results, as compared to the other three models, is largely driven by two economic factors not reflected in the other models—namely, the ITC extension ending in 2022, and PV costs ceasing to decline in 2020. As a result of these two factors, PV adoption in dSolar is front-loaded to periods before 2022 and then experiences a quicker decline than others.

The E3 forecast on cumulative installation in Fig. 4 is slightly higher than the threshold-heterogeneity and Bass forecasts through 2016; it then becomes lower (at least through 2024), mostly due to differences in assumptions of the federal ITC policy in that study.[23] E3's post-2016 forecasting has not yet shown any obvious inflection point, though E3's bottom-up engineering model was also based on the classic Bass model.

An implicit assumption in the results discussed so far is that a system will automatically be rebuilt at the end of its productive life (25 years). We consider this assumption reasonable because the household has already adopted and PV prices are projected to continue decreasing. To make this assumption more explicit, we further differentiate the annual number of PV installations into first-time adoption and re-adoption below.

In Fig. 4, most diffusion curves have an inflection point around 2020 or 2022. Fig. 5 Panel A compares biannual installation results for the first-time PV adoption from three of our methods (excluding the time-series method, which would appear as a simple horizontal line) and the E3 study. All forecasts show residential PV installation peaking in

[20] In 2013, there were 6,939,104 OOH in California (ACS, 2013).

[21] Neural-network methods generally could not recognize the data-generation process as an S-shaped curve even with multiple hidden neurons and layers. However, with modifications (e.g., introducing the past cumulative installations and their square terms into the training model), a neural network could learn the S-curve.

[22] Other assumptions are similarly favorably for increased PV adoption, including assumptions of declining PV prices and increasing electricity rates. See Sigrin et al. (2016) for details.

[23] E3's work was done prior to the ITC extension approved by the U.S. Congress in December 2015.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

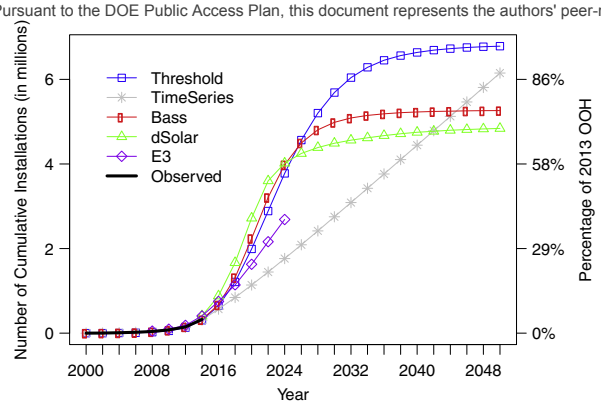

Fig. 4. Observed and forecast cumulative installations in California, 2000–2050. Labels at the right axis depict the percentages of the corresponding marked numbers of cumulative installations as in 2013.

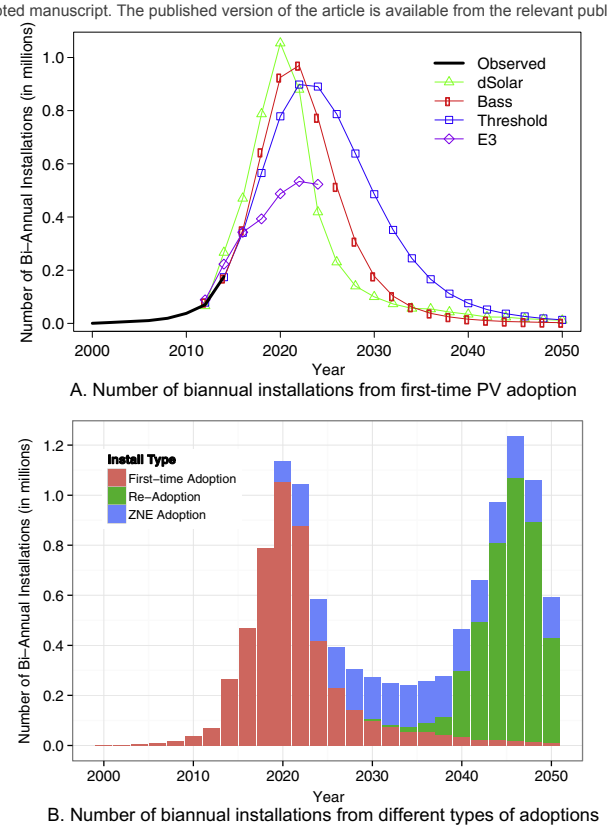

Fig. 5. Observed and forecast biannual installations in California, 2000–2050 The time series result is not shown here because it is simply a horizontal line.

2020 or 2022 at roughly half a million installations per year, suggesting that the very rapid growth among first-time adopters from 2012 through the first half of 2015 (GTM/SEIA, 2015) will likely subside in the next 5–7 years, barring a major change in the California market.[24] The dSolar forecast has a higher and earlier peak in 2020, which is consistent with the new ITC extension policy because the ITC level will drop from 28% to 11% in 2022.[25] Furthermore, as in Fig. 4, the threshold model shows more PV adoption after the peak year, whereas the Bass model lies in between. Lastly, E3 predicts a flatter biannual adoption curve until 2024 and a fatter tail afterwards (by extrapolation), whereas the other three models all predict the fastest growth from now until before 2020 or 2022 and then a decrease in the number of first-time PV adoption, with the threshold model decreasing slower than the other two Bass-type models.[26]

The above decrease in first-time PV adoption does not necessarily mean the residential PV market is going to collapse in California after the mid-2020s, because these models do not consider alternative sources of market growth. Panel B of Fig. 5 estimates the additional deployment when considering re-adoption and the Zero Net Energy new home mandate. However, even when considering these segments, there could be a market contraction following the peak of first-time adoption and when these systems are re-adopted, which is common in the multi-generation Bass diffusion models for repeated purchased technologies (Jiang and Jain, 2012; Norton and Bass, 1987). Nevertheless, three factors not explored here could fill in the gap as seen in Panel B: 1) The decrease in first-time adoption, post-peak, may be less pronounced (as shown in the threshold model); 2) Other residential market segments (not considered here), such as non-OOH, community solar, and even landlords selling solar power could grow in popularity; and 3) We assume a system lifetime of exactly 25 years, though there is likely to be a range of expected lifetimes, which would smooth the adoption curve. Finally, TPO contracts typically last for only 20 years, which means new business could come in 20-year cycles instead of 25 years for TPO companies.[27]

Fig. 6 shows biannual installation forecasts on a capacity basis from first-time PV adoption only and includes several additional forecast sources.[28] The amount of capacity simulated for the Bass and threshold-heterogeneity diffusion models derive from multiplying the installation numbers by an assumed average system size. We use the average system size from E3 for 2016–2024 and the average system size from dSolar for 2026 and beyond. Both E3 and dSolar calculate system size based on offsetting a percentage of electricity use for each customer segment[29]; average system size then accounts for all customer segments in a year at the state level. In 2014, the average system size from E3 is 5.2 kW versus 3.8 kW from dSolar, which should give dSolar a lower installed capacity forecast for that year. But since dSolar predicts a greater number of installations as seen in Panel A of Fig. 5, the impact of a small average system size is not reflected in Fig. 6.[30]

The time-series result is not shown here because it is simply a horizontal line.

After 2016, the relationship among the dSolar, Bass, and threshold modeling results for installed capacity resembles that for installations in Fig. 5 Panel A. The peaking year is still 2020 and 2022 for most models, and the peaking installed capacity is around 3–4 GW biannually based on the Bass, threshold, and dSolar results. This peaking capacity is based on several key assumptions: an average system size around 3.8 kW, a maximum market size equivalent to 70%–98% of owner-occupied houses in 2013, and no major future policy changes (e.g., on retail

[24] Again, we did not include any re-adoption that could occur at the end of system lifetime. Furthermore, major market innovations or policy changes could shift the curve peaks in Fig. 5 to the right, even without considering re-adoption.

[25] The 28% and 11% figures come from averaging the ITC level over a two-year increment, as is consistent with the dSolar model.

[26] The difference between our three methods and E3 is due in part to E3's assumption that the number of residential customers served by the three biggest California IOUs will grow 1.138% annually, resulting in 1.5 times more people in 2050 than in 2015. In contrast, dSolar assumes a 0.2% annual average population growth rate based on EIA data (EIA, 2009; EIA, 2015), and the threshold-heterogeneity and Bass diffusion models assume a fixed market size following Bass et al. (1994).

[27] This new business does not necessarily mean re-adoption, because TPO customers can still use the old PV system.

[28] *Re*-adoption will only start to happen in late-2020s and most other sources forecast PV installed capacity before that, so we only show first-time PV adoption numbers here.

[29] The percentage offset in dSolar is 95% for people with NEM and 66% for others without NEM, whereas E3 chooses the best system size from three options: offsetting 33%, 66%, or 100% of annual customer load.

[30] The relative small average system size and large number of installations reflect the underlying modeling philosophy of dSolar. Designed as an agent-based model, dSolar tends to model all agents in the market, whether they are small electricity users or big ones. As a result, we tend to see more adopters in the result with a smaller system size.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

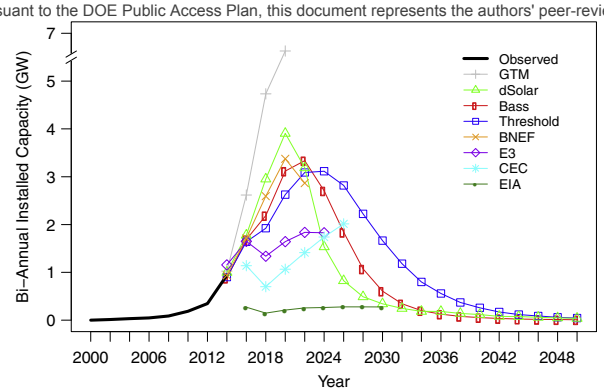

Fig. 6. Observed and forecast biannual installed capacity from first-time PV adoption in California, 2000–2050. Sources: GTM (2015); BNEF (2015b); E3 (2015); CEC (2015); and EIA (2015) EIA forecast data are from American Energy Outlook 2015 (EIA, 2015), where we assume California accounts for 88.6% (based on GTM/SEIA's capacity number in 2014) of the residential PV installed capacity in the Pacific Census Region.

rate) or market innovations. We further assume the current NEM policy in California (e.g., full retail credit for excess generation) will continue until 2050.

After incorporating the effect of ITC extension, GTM (2015) forecasts the highest biannual installed capacity that reaches 4.7 GW in 2017–2018 and 6.6 GW in 2019–2020, the latter of which is more than six times higher than recent biannual capacity additions (2013–2014) and much higher than other modeling results. Although this may be consistent with extrapolating the recent historical trend, it is much higher than estimates from other models. In contrast, BNEF (2015b) forecast resembles dSolar's result, though with a smaller peak and much shorter forecasting horizon.

CEC, E3, and EIA predict a dip in terms of installed capacity in 2018, due to the (now deprecated) assumption of ITC expiration in 2017, which was recently extended by the U.S. Congress until 2022. Specifically, CEC predicts roughly 2 GW of biannual installed capacity in 2026 with no decreasing trend, though in most of the periods before 2026, CEC's forecast is lower than other predictions (CEC, 2015). Pacific Gas and Electric (2015) pointed out previously the possibility of the CEC forecast for under-predicting DGPV adoption, because CEC has relied on payback to predict consumer response rather than using metrics such as levelized cost of energy (LCOE) or monthly bill savings to reflect increasingly popular TPO arrangements. Both E3 and dSolar have accounted for this issue.[31]

Lastly, EIA's forecast is lower than all other results. After 2020, the forecast is basically a horizontal line and the biannual installed capacity remains at roughly 270 MW within two years, which is even lower than recent installation rates in California.

4.2. County-level residential PV deployment

Of California's 58 counties, 57 have DGPV installations. This section uses the threshold-heterogeneity and Bass diffusion models to explore detailed county-level results.

Fig. 7 provides snapshots of PV penetration levels (i.e., market share) in 2014 (observed) and 2030 (from the threshold-heterogeneity diffusion model), showing a dramatic increase over time.[32] In 2014, only three counties have a penetration level higher than 8%, but in 2030, 28 counties have a penetration level higher than 80% (~20 GW installed capacity in total before any curtailment). In 2050 (not shown), only two counties have a penetration level below 80% (~25 GW installed capacity in total before any curtailment). High-penetration counties are not necessarily those with high numbers of installations. For example, in 2014, the top three counties in terms of penetration levels are Yolo County (penetration level 9.7%), Glenn County (9.2%), and Fresno County (8.8%), while the top three in terms of installations are Los Angeles County (penetration level 5%), San Diego County (7.3%), and Riverside County (6.1%). The correlation between these two variables is only 0.2, reflecting the potential complementarity of these two metrics for evaluating market development.

Fig. 8 shows the distribution of estimated p (innovation coefficient) and q (imitation coefficient) values at the county level based on the Bass diffusion model. These values determine PV diffusion speed in each county. Although most diffusion models assume certain p and q values, they are often not empirically calibrated based on observed county-specific data. For instance, E3 uses a p value of 0.02, which is the upper limit of what Jeuland (1994) found. Based on the q value assumed by E3 (0.3), a peak time (time to maximum annual adoption) of about 8.5 years can be derived using Eq. (5), which only makes sense if using 2014 as the year PV was introduced to the market.

The p and q values in Fig. 8, based on quarterly PV installation data, are roughly one quarter of those based on annual installation data (i.e., the reason they look small). Panel A suggests that most p values are <0.001 and the county-to-county difference is minor (except for two counties with relatively high p values); in other words, the proportion of innovators in the population is generally small, considering it took Los Angeles from 1998 until 2014 to reach a penetration level of 3%. In contrast, California counties differ much more with respect to proportions of imitators (the q value, Panel B). Using p and q values from Fig. 8, the median peak time across 57 counties is about 86 quarters or 21.5 years. Because many California counties saw PV installations starting in the late 1990s, it is not surprising that the peaking year in Fig. 5 is between 2020 and 2022.

Table 2 explores the relationship between q values (imitation coefficient) and socio-demographic variables. Overall, these variables seem to explain the q values well, with an adjusted R^2 as high as roughly 60%. Of the 15 socio-demographic variables that have a significant piecewise correlation with the q values, only three remain significant after controlling for other factors. Big families with larger household sizes are positively associated with high q values, which is consistent with our intuition. Households relying on a utility gas company (rather than wood) for heating tend to have high imitation behavior. Lastly, counties with a high percentage of other races tend to have low q values. Among those insignificant variables, it seems that higher education level is negatively correlated with imitation behavior, whereas lower education level is positively correlated with it, though neither are statistically significant with a 57 sample size.

4.3. Sensitivity analysis

As a demonstration of key uncertainties in forecasting DGPV technology diffusion, we conduct a sensitivity analysis around three economic parameters using the dSolar model because the other three methods do not explicitly rely on information of these economic variables to make forecasts. The three economic parameters are: technology cost, technology performance, and cost of grid integration.[33]

- Technology Cost: The installed cost of rooftop solar, in addition to the value of solar generation, is a key determinant of the economic value of the PV system. The reference cost trajectory is the "mid" cost trajectory from NREL's 2015 Annual Technology Baseline (Blair et al., 2014).

[31] E3 converts the project-level lifetime benefit-to-cost ratio to an implied payback period, which works well for TPO PV systems. Then, E3 uses a method similar to CEC's to map the payback period to market diffusion for various market segments within each investor-owned utility (IOU) (E3, 2015). In contrast, dSolar uses monthly bill savings to model TPO adoptions.

[32] Penetration level is calculated by dividing the number of cumulative installations by the estimated and constant maximum market size based, in 2030, on the threshold-heterogeneity diffusion model, which is constrained to be no greater than the number of owner-occupied houses in a county in 2013.

[33] We thank the anonymous reviewer for pointing out this direction to us.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

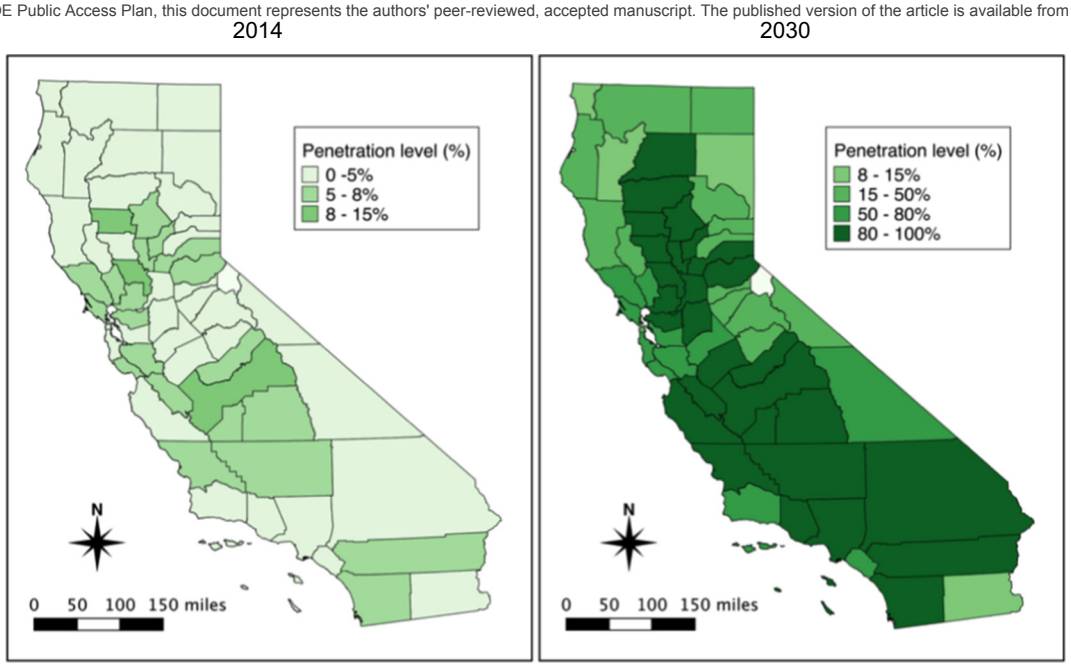

Fig. 7. California county-level DGPV diffusion results in 2014 (observed) and 2030 from the threshold-heterogeneity diffusion model.

In addition, a "high" cost sensitivity is modeled which assumes cost decreases to $1.5/W in 2040 (*High Cost*). The "low" cost assumption assumes continued cost declines to $1/W in 2030 (*Low Cost*), and the "optimistic" assumption considers additional decline to $0.5/W by 2040 (*Optimistic Cost*). Though these values represent significant technology innovation, they are in line with recent expert forecasts (Cole et al., 2016).

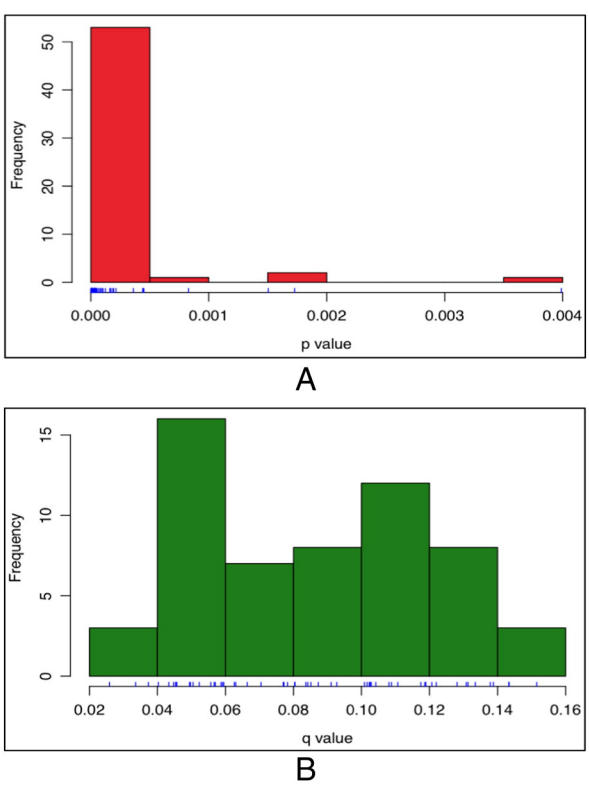

Fig. 8. Estimated county-specific *p* (Panel A) and *q* (Panel B) values from the Bass diffusion model using quarterly installation data Ticks above x-axis indicate exact values.

- Technology Performance: Several decades of PV panel reliability tests give a high degree of certainty for the expected lifetime of a PV module (Jordan and Kurtz, 2013). In addition, many panel manufacturers offer lifetime warranties certifying their performance. In this analysis, we assume that PV panels have a lifetime of 25 years and have a mean annual degradation rate of 0.5%. Nevertheless, customer uncertainty about module quality, real or perceived, could deter technology adoption. Thus, we analyze technology uncertainty assuming a 20-year lifetime and 1.5% annual degradation (*Low Performance*) and 30-year lifetime and no degradation (*High Performance*).
- Renewable Integration: Understanding the impacts of deep levels of renewable generation on operation of the electrical grid is an important issue in renewable energy modeling. The challenges are both engineering-based and policy-based. Though the true impacts are not entirely known and would vary dramatically by electrical system, the marginal value of renewable generation is likely to decrease as penetration increases (Cole et al., 2016). Impacts are likely at both the transmission level, through curtailment of surplus generation, and at the distribution level, as increasing levels of distributed

Table 2
Regression of *q* values (proportions of imitators) on socio-demographic variables.

Social demographic variables	Coef.	Std. Err.
Mean household sizes	0.062**	0.024
Median number of rooms	0.019	0.017
% of OOH with less than high school education	0.038	0.164
% of OOH with college education	0.145	0.186
% of OOH with bachelor education	−0.215	0.189
% of OOH with utility gas as a heating source	0.075**	0.031
% of OOH with wood as a heating source	0.015	0.068
% of OOH that is white	−0.098	0.072
% of OOH that is black	0.128	0.173
% of OOH that is other race (not Indian or Asian)	−0.277*	0.138
% of OOH with only one mortgage	−0.233	0.142
% of OOH with mortgage over 40% of household income	0.076	0.104
% of OOH built in 2000s	0.057	0.071
% of OOH built in 1970s	−0.020	0.097
% of OOH with income between $100,000 and $150,000	0.053	0.096

$N = 57$, Adjusted $R^2 = 0.59$.
* $p < 0.1$.
** $p < 0.05$.

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

generation could trigger feeder voltage issues (Palmintier et al., 2016). We simulate integration sensitivities only at the distribution level, as transmission-level impacts would be better reflected in retail energy prices, which are an immutable part of this analysis. A recent engineering study (Navigant, 2016) estimates that distribution-system integration costs, without any mitigation strategies, could range from \$0.1/W–\$0.5/W as PV penetration reaches 50%–100% of a feeder's thermal limits. Therefore, we test the sensitivity of our baseline results to the integration cost by imposing a low integration cost of \$250/kW (*Low Integration Cost*) and a high integration cost of \$500/kW (*High Integration Cost*), both at 2014-dollar value.

Table 3
Sensitivity analysis of dSolar to three economic parameters.

Scenario	Cumulative number of adopters in 2050	% Change from reference
Reference case	4,840,695	0.0%
High cost	2,271,746	−10.6%
Low cost	5,394,457	+11.4%
Optimistic cost	5,593,079	+15.5%
Low performance	4,744,922	−2.0%
High performance	4,869,013	+0.6%
Low integration cost	4,428,349	−8.5%
High integration cost	3,928,548	−18.8%

Adoption projections using the dSolar model were developed for each of the sensitivity categories described above. Sensitivity to inputs was analyzed relative to the Reference scenario for the cumulative number of first-time adopters in 2050 (Table 3). Results were most sensitive to costs, followed by distribution integration costs, and performance. The increase in penetration for the *High Cost* and *Optimistic Cost* scenarios is relatively modest (11.4%–15.5%), because we already observe high levels of penetration using the *Reference* values. On the other hand, the decrease in penetration for the *High Cost* case is at the same magnitude (minus 10.6%).

The modeled results are insensitive to performance assumptions, mostly because the mean payback period for most potential adopters (<7 years) is substantially less than the range of modeled lifetimes (20–30 years). Finally, there was a modest response to the distribution integration costs modeled. We note that there are still significant uncertainties about transmission-level integration challenges, such as generation curtailment, which are not modeled here.

4.4. Policy scenario analysis

This section estimates the effects of three major DGPV-related policies: the recent federal ITC extension passed by the U.S. Congress (2015), a newly approved decision from the CPUC on the NEM tariff successor (CPUC, 2015a), and a hypothetical VOS compensation scheme. The ITC extension retains the 30% tax credit until 2019 and then gradually phases it out to 0% for customer-owned PV and 10% for TPO PV in 2022. Because dSolar only resolves in even years, we assume the average ITC level for the 2020 solve year to be 28% (30% for TPO) and 11% in the 2022 solve year (24% for TPO). After 2022, the ITC level stabilizes at 0% for customer-owned PV and 10% for TPO PV.

To simulate the newly approved California NEM policy (CPUC, 2015a), we assume that PV customers will pay a one-time \$100 interconnection fee and a non-bypassable charge of 2.5 cents per kilowatt-hour (kWh) of net energy consumed from the grid starting from the 2018 dSolar model solve year. We expect this policy would have a modest impact on PV economics because the non-bypassable charge is dwarfed by high retail costs of electricity (e.g., Pacific Gas and Electric's Tier 3 rate of \$0.27/kWh and Tier 4 rate of \$0.30/kWh). Though the CPUC has indicated that it will require all new PV customers to subscribe to a time-of-use rate starting from 2018, for simplicity, we assume the time-of-use (TOU) rate is cost-neutral to PV customers.[34]

To further demonstrate the impact of NEM policy, we create another scenario where the exported PV generation is compensated at \$0.12/kWh starting in 2018. Such a policy has been discussed by E3 (2015) and is similar to Austin Energy's VOS tariff in 2015 (\$0.113/kWh).[35] Below, we use the dSolar model and the threshold-heterogeneity diffusion model to simulate the potential impact of these three policies.[36]

Fig. 9 summarizes the biannual DGPV installation results for four scenarios and the baseline result (S0) using the dSolar model. These four scenarios are as follows: ITC expires in 2017 (S1); California implements its new NEM policy starting in 2017 (S2); the combination of the first two scenarios (ITC Expiration + New NEM) (S3); and finally, the VOS compensation scheme at \$0.12/kWh replaces the existing NEM (S4).

Comparing the dSolar baseline scenario S0 to S1, the extension of the ITC is estimated to have increased the peak-installed capacity in 2018 by around 27% or roughly 800 MW and in 2020 by around 12% or roughly 500 MW. Long-term effects of the extension are mixed—after 2022, annual deployment is greater in the non-extension scenarios due to pent-up demand and generally favorable economic conditions.[37] Impacts of the ITC extension in dSolar is more modest as compared to those of two other industrial sources: GTM (2015) predicted an increase of 33% (or 1.6 GW) in 2017–2018 and a further increase of 25% (or 1.7 GW) in 2019–2020, whereas BNEF (2015b) predicted an increase of 35% (or 900 MW) in 2017–2018 and an increase of 43% (or 1.5 GW) in 2019–2020. One explanation involves the underlying method difference: whereas dSolar leverages the classic Bass diffusion model, which necessitates an inflection point in the diffusion curve, GTM did not include an inflection point in their predictions. Because BNEF also uses the Bass model, it is possible that the p and q parameters within the Bass model are different between dSolar and BNEF. Furthermore, our PV installation cost assumption could also differ from BNEF's because they have estimated a greater impact in 2019–2020.

Regarding the new California NEM policy, the impact on PV diffusion is minimal, as shown by comparing scenarios S0 and S2, or S1 and S3. In 2020, the difference between S0 and S2 is 1.1% (or 44 MW), whereas the difference between S1 and S3 is 2.8% (or 95 MW). Such modest impact is consistent with what we first expected, because the new NEM policy only imposes small new charges to future PV customers (and our assumption that customers on TOU rates continue to see the same bill savings).

However, in the scenario of the VOS compensation scheme where the hourly[38] exported PV generation is compensated at \$0.12/kWh, the decrease in PV diffusion is more substantial, mostly because the VOS credit is substantially less than the current value of net metered generation at current rate levels. In 2020, the bi-annual installed capacity in S4 is 32% lower than it is in S0, or 1.3 GW less. Such a big decrease is understandable considering the low compensation level that PV

[34] The TOU rate could increase PV economics for customers during peaking hours. But with high PV penetration that shifts the peaking hours increasingly to the evening hours (Darghouth et al., 2015), the increased PV economics could be gradually reduced. As for the \$10 minimal monthly bills, because dSolar only sizes the PV system to offset at most 95% of the annual electricity load, this requirement is non-binding.

[35] "Residential Solar Energy Rate: Value of Solar (VoS)." Austin Energy. http://austinenergy.com/wps/portal/ae/rates/residential-rates/residential-solar-energy-rate.

[36] The other two models, the Bass diffusion model and the time-series model, are not suitable for this task because their results do not necessarily rely on PV economics.

[37] The relatively higher installed capacity after the peak in S1 and S3 compared to S0 and S2 is simply because ultimately in 2050, these four scenarios would converge in terms of total number of PV adoption; with PV costs continuing to decline, PV diffusion in S1 and S3 would catch up with the other two scenarios (through faster later adoption) even without the five-year ITC extension.

[38] dSolar sizes the PV system to offset 95% of the customer's load, so accounting for the net consumption or the exported generation on an annual basis would make no difference in the VOS scenario.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

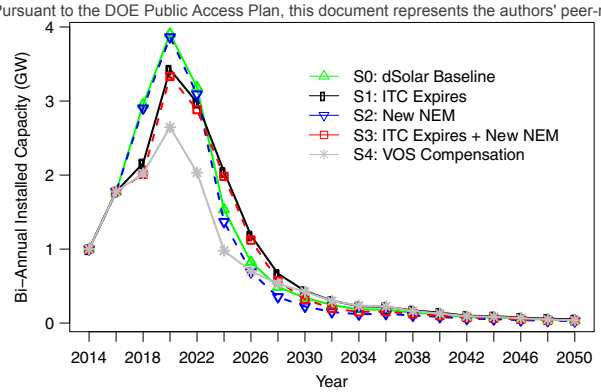

Fig. 9. Residential DGPV policy scenario analysis using dSolar: dSolar baseline (no policy change) vs. ITC expiring, new NEM and VOS policy taking effect in 2017.

customers would receive; nevertheless, the decrease could be alleviated by either a higher compensation level (i.e., closer to the retail rate) or accounting for the exported PV generation over a longer interval (i.e., monthly or quarterly).

Because the ITC expiration, new NEM policy, and VOS compensation scheme solely affect PV economics, we can also use the threshold-heterogeneity diffusion model for scenario analysis, though in an indirect way. First, we must quantify the impact of changes in PV economics (e.g., NPV) on PV diffusion in general. We extend Davidson et al. (2014) and use PV economics data from Tracking the Sun (Barbose and Darghouth, 2015) and Gillingham et al. (2016) to model PV penetration at the census tract level in California. Finally, we map the impact of the policy changes on PV NPV to changes of PV deployment in the threshold-heterogeneity diffusion model.

Table 4 compares results from the original model of Davidson et al. (2014) and from our version. We use average penetration level at the tract level from 2007 to 2014 as the dependent variable, whereas Davidson et al. (2014) use the logarithm of the total number of installations.[39] We use penetration level because it is the variable we use in the threshold-heterogeneity diffusion model, and furthermore, we believe this model of the penetration level produces more intuitive results. Specifically, we believe it is better to replace *mortgage2orHE* with *built90s*, replace *wood*[40] with *electricity*, and replace *value_over1mil* with *value_500kto1mil*, which has a positive coefficient. In addition, we include a straightforward household income variable (*inc_100to150k*) and drop the *hh_size* variable, which is correlated with *median_rooms* and *edu_bachorhigh*. All our variable coefficients are easy to interpret; census tracts with higher PV penetrations are likely to be those with better PV economics, higher education levels, higher income, greater electricity demand, higher white populations, more recently built houses, more people relying on grid electricity, and—potentially—more highly valued houses and fewer second mortgages or home equity loans. For example, reducing the NPV of PV by $1/W reduces the penetration level in a census tract by roughly 0.1%.[41]

Next, we use dSolar to determine the impact of the ITC expiration, the newly approved NEM policy, and VOS policy on NPV at the county level. Finally, we input the regression and NPV results into the threshold-heterogeneity diffusion model and compute the impact on the installed capacity in California.

Fig. 10 presents the results of the scenario analysis using the threshold heterogeneity diffusion model. The relative closeness between S0 and S2 and between S1 and S3 is still obvious, indicating the minor impact of the new NEM policy. The short-term impact of ITC until 2024 resembles what is found in Fig. 9. In 2018, the increase in installed capacity due to ITC extension is 30% (or roughly 600 MW), whereas this number is 18% (or 500 MW) in 2020. After 2022, with PV cost being low enough, the impact of ITC expiration disappears.

Regarding the scenario of VOS compensation (S4), its impact on PV installation is more substantial—a 32% decrease (or 900 MW less) in 2020. Furthermore, the impacts persist longer than they do in the dSolar scenario analysis due to the methodological difference in these two scenario analyses. Scenario results based on the threshold model are derived from comparing the PV economics in those scenarios and that in the baseline case, and then evaluating the impact of PV economics changes on PV penetration. But each scenario based on dSolar represents an individual model run, and the common Bass diffusion parameters and same PV economics in the last few years would suggest a very similar final adoption level. In other words, scenarios based on the threshold model do not need to catch up with the baseline result, whereas this is not the case in dSolar. Such a methodological difference indicates that top-down models such as the threshold model may not be well suited for long-term impact analysis, though the near-term results seem reasonable.

Table 4
Regression results of PV economics impact on diffusion, census tract level.

DV: penetration level	This report	DV: log of installs	Davidson et al. (2014)
npv_per_watt	0.00098[a]	npv_per watt	0.01703
	(0.00017)		(0.01321)
edu_bachorhigh	0.00422[a]	edu_bachorhigh	0.77748[a]
	(0.00159)		(0.09357)
inc_100to150k	0.01013[a]		
	(0.00272)		
median_rooms	0.00709[a]	median_rooms	0.73542[a]
	(0.00035)		(0.01887)
		hh_size	−0.15075[a]
			(0.02950)
race_white	0.01172[a]	race_white	0.69177[a]
	(0.00077)		(0.06652)
mortgage_over40%	−0.00207	mortgage_over40%	−0.23472[b]
	(0.00137)		(0.08030)
built90s	0.00960[a]	mortgage2orHE	0.81524[a]
	(0.00165)		(0.18310)
electricity	0.00510[a]	Wood	1.40606[a]
	(0.00175)		(0.23773)
value_500kto1mil	0.00095	value_over1mil	−0.59006[a]
	(0.00114)		(0.10110)
Sample size	7305		7305
Adjusted R square	0.259		0.378

Davidson et al. (2014) use stepwise regression to select the most important variables in the regression. This table is based on variables selected in Fig. 1 of Davidson et al. (2014), from ACS (2013).
The coefficient of our *mortgage_over40%* is significant at the 10% level. Robust standard errors are in parentheses.
[a] $p < 0.001$.
[b] $p < 0.01$.

5. Conclusions

We use three top-down models (time-series forecasting, threshold-heterogeneity diffusion, and Bass diffusion) and a bottom-up engineering model (dSolar) to forecast residential DGPV deployment in California. Top-down models require many fewer data inputs and assumptions, and they are easier to implement than bottom-up models. However, they generally lack the ability to conduct scenario analysis of major policy changes (with the threshold model as an exception), especially long-term policy impact analysis. In this report, the time-series forecasting model mainly illustrates the need to build more complex models, whereas the threshold-heterogeneity and Bass diffusion models provide relatively simple tools for checking the reliability of results from a bottom-up engineering model (dSolar). We also show how

[39] We obtain very similar results as Davidson et al. (2014) if we switch to their dependent variable.
[40] The *wood* variable in Davidson et al. (2014) has a negative coefficient, which may be correct in the sense that PV and wood are less likely to substitute each other as heating source.
[41] Note that 0.1% is at the tract level. The difference in PV economics (e.g., NPV) across census tracts in California should not be big, which means that other factors could explain the penetration-level difference in the data.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

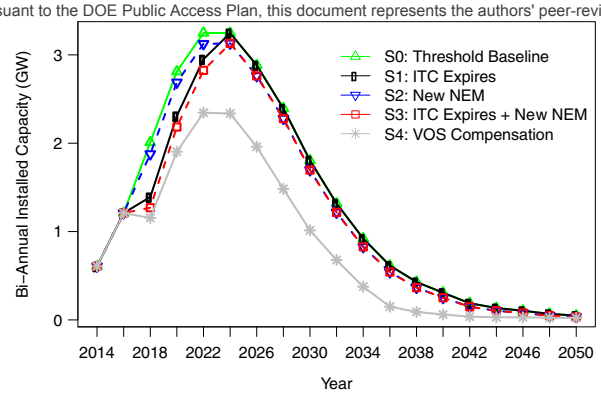

Fig. 10. Residential DGPV policy scenario analysis with the threshold-heterogeneity diffusion model: threshold model baseline (no policy change) vs. ITC expiring, new NEM and VOS policy taking effect in 2017.

the threshold-heterogeneity diffusion model can be used for policy scenario analysis. On the other hand, bottom-up models are much more detailed and powerful for conducting scenario analysis, but they require assumptions for each module of the engineering model. In this analysis, we find that the two model types complement each other and we compare our methods with industry sources.[42]

Our results from the two diffusion models and dSolar suggest, barring any major policy changes—although with federal ITC extending until 2022, four-tiered rate, and NEM policy continuing indefinitely—that the residential PV market in California could peak between 2020 and 2022 at about 3–4 GW of biannual capacity (or 1.5–2 GW annual capacity); this corresponds to about 0.9–1.1 million biannual installations (or half a million annual installations), both from the first-time adoption. Compared with other forecasting sources, our results are not as high as some industry forecasts but are greater than predictions by the EIA and CEC (at least in the short term). Importantly, our baseline forecasts assume no major policy changes in California's residential PV sector from now until 2050, though we have later used two of our methods to estimate the impact of potential policy changes in California on PV adoption.

We find significant county-level heterogeneity with respect to residential DGPV deployment in California. Currently, very few counties have a penetration level above 10% based on our estimates of maximum market size. However, barring any major changes in policy, our threshold-heterogeneity diffusion model projects that about half of California's counties will have a penetration level >80% in 2030, which could pose substantial grid-integration challenges. After imposing a $500/kW integration cost as we did in our sensitivity analysis, the 2050 penetration level could be roughly 20% lower.

Based on our Bass diffusion model results, there are small and similar proportions of innovators in each county; but the proportions of imitators are much larger and differ significantly, which is a finding similar to Guidolin and Mortarino (2010). Certain socio-demographic variables correlate highly with the degree of imitation in a county, such as mean household sizes and using the utility gas company for heating. In addition, a regression model (Table 4) that we developed using both PV economics and socio-demographic variables explains different penetration levels at the census tract level, and all regression coefficients are intuitive.

Because changes to California and national policy are likely to be implemented in 2016, we use dSolar and our threshold-heterogeneity diffusion model to analyze the impact for California of proposed changes—including the recent ITC extension, a newly approved NEM policy with additional charges to PV customers, and a hypothetical VOS compensation scheme. Our dSolar results suggest that the impact from the recent ITC extension mostly affects short-term deployment, because declining PV costs offset policy support in the long-term; this finding is also confirmed in the threshold model analysis. In terms of the magnitude of the impact, our two models have estimated a 27%–30% increase (600–800 MW reduction) in the 2017–2018 PV installation and a 12%–18% increase (roughly 500 MW reduction) in the 2019–2020 PV installation for the California residential sector. Although this impact may appear low compared to other industry forecasts by GTM and BNEF, it can be explained by differences in our underlying methods and parameterizations.

Based on our scenario analyses, we have also found that the newly approved NEM policy in California could potentially have a minor impact on future PV adoption (<100 MW) assuming TOU rate impacts are cost-neutral. This minor impact is because the imposed low non-bypassable charge would be dwarfed by the high retail rates in California as of now. Instead, additional rate restructuring is likely to have a much larger impact on long-term deployment. For example, our models show that were California to implement a VOS compensation scheme with compensation lower than the current net-metered generation value (say, $0.12/kWh starting in 2016), annual PV deployment could be reduced by 32% (900–1300 MW) compared to our baseline results.

Our assumptions have important effects on the results. First, our average PV system size is small relative to some industry estimates, and using a larger size would increase our capacity results proportionally. Second, our maximum market size is based on the number of owner-occupied houses for the two diffusion models in 2013, and it is further restricted to single-family owner-occupied houses in dSolar. If future innovation expands the residential market beyond owner-occupied houses or enables a higher percentage of owner-occupied houses to adopt DGPV, our assumption might be conservative. Third, the assumption of allowing the residential PV market to grow without any grid constraint or curtailment seems to be optimistic, even if we have imposed certain integration costs in our sensitivity analysis. Fourth, our assumption of the future TOU rate being cost-neutral to PV customers warrants future work. Finally, we assume the fitted density curves—either normal or Bass diffusion—would not move over time, which might not be true if future innovations (e.g., dramatic performance improvements or introduction of new business models) are realized. If the curves do move, the installed capacity curves (Fig. 6) will have heavier (or thinner) right tails.

Our work can be extended easily to other PV sectors and markets if high-quality data are available. Additional future work could include further combination of the various modeling methods to provide more robust forecasts, or further relaxation of some of our assumptions. In this work, we have not modeled the impact of the potential TOU rate in California but simply assumed it is cost-neutral to PV customers; additional work could address this important issue. Finally, uncertainties remain a topic to be explored as which modeling approach can predict future PV deployment most credibly. So a more detailed sensitivity analysis is needed.

Acknowledgments

The authors would like to thank Trieu Mai, Wesley Cole, Robert Margolis, and Aaron Bloom of the National Renewable Energy Laboratory (NREL), and Andrew Mills of Lawrence Berkeley National Laboratory (LBNL) for providing valuable comments, and Jarett Zuboy (consultant) and Mike Meshek (NREL) for providing technical editing help. The authors would also like to thank Jim Turnure and the Energy Information Administration for support for this study (Project # DOOE1000).

References

ACS, 2013. American Community Survey 5 year data. http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

[42] In addition, these two types of models also differ in terms of accounting PV economics. Although it is not required for top-down models to model PV economics, PV economics is the key and driving factor in bottom-up models.

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

Armstrong, J.S. (Ed.), 2001. Principles of forecastingInternational Series in Operations Research & Management Science Vol. 30. Springer US, Boston, MA . http://link.springer.com/10.1007/978-0-306-47630-3.

Barbose, Galen L., and Naïm R. Darghouth. 2015. "Tracking the Sun VIII: The Installed Price of Residential and Non-residential Photovoltaic Systems in the United States." Edited by Dev Millstein, M. Spears, Ryan H. Wiser, Michael Buckley, Rebecca Widiss, and Nick Grue, (August).

Barnes, J., Culley, T., Haynes, R., Passera, L., Wiedman, J., Jackson, R., 2013. Freeing the Grid 2013: Best Practices in State Net Metering Policies and Interconnection Procedures. Interstate Renewable Energy Council (Accessed July 7, 2014: http://freeingthegrid.org/wp-content/uploads/2013/11/FTG_2013.pdf).

Bass, F.M., 1969. A new product growth for model consumer durables. Manag. Sci. 15 (5): 215–227. http://dx.doi.org/10.1287/mnsc.15.5.215.

Bass, F.M., Krishnan, T.V., Jain, D.C., 1994. Why the bass model fits without decision variables. Mark. Sci. 13 (3):203–223. http://dx.doi.org/10.1287/mksc.13.3.203.

Beck, R.W., Inc, 2009. Distributed Renewable Energy Operating Impacts and Valuation Study. Prepared for Arizona Public Service by R.W. Beck, Inc.

Bemmaor, A.C., 1994. Modeling the diffusion of new durable goods: word-of-mouth effect versus consumer heterogeneity. In: Laurent, G., Lilien, G.L., Pras, B. (Eds.), Research Traditions in MarketingInternational Series in Quantitative Marketing vol. 5. Springer, Netherlands:pp. 201–229 (http://dx.doi.org/10.1007/978-94-011-1402-8_6).

Bemmaor, A.C., Lee, J., 2002. The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates. Mark. Sci. 21 (2), 209–220.

Blackburn, G., Magee, C., Rai, V., 2014. Solar valuation and the modern utility's expansion into distributed generation. Electr. J. 27 (1):18–32. http://dx.doi.org/10.1016/j.tej.2013.12.002.

Blair, N., Dobos, A.P., Freeman, J., Neises, T., Wagner, M., Ferguson, T., Gilman, P., Janzou, S., 2014. System Advisor Model, SAM 2014.1.14: General Description. NREL/TP-6A20-61019. Golden, CO 80401: National Renewable Energy Laboratory.

BNEF, 2015a. H1 2015 North American PV Market Outlook: Wide Open Throttle. Bloomberg Finance L.P, New York.

BNEF, 2015b. How Extending the Investment Tax Credit would Affect US Solar Build. Bloomberg Finance L.P, New York.

Bollinger, B., Gillingham, K., 2012. Peer effects in the diffusion of solar photovoltaic panels. Mark. Sci. 31 (6):900–912. http://dx.doi.org/10.1287/mksc.1120.0727.

Bosetti, V., Catenacci, M., Fiorese, G., Verdolini, E., 2012. The Future Prospect of PV and CSP Solar Technologies: An Expert Elicitation Survey. *Energy Policy*, Special Section: Fuel Poverty Comes of Age: Commemorating 21 Years of Research and Policy, 49 (October). :pp. 308–317 http://dx.doi.org/10.1016/j.enpol.2012.06.024.

Bottomley, P.A., Fildes, R., 1998. The role of prices in models of innovation diffusion. J. Forecast. 17 (7):539–555. http://dx.doi.org/10.1002/(SICI)1099-131X(199812)17:7<539::AID-FOR684>3.0.CO;2-S.

Carnell, R., 2012. Lhs: Latin Hypercube Samples. R Package Version 0.10. http://CRAN.R-project.org/package=lhs.

CEC, 2015. California Energy Demand 2016–2026, Preliminary Electricity Forecast. http://docketpublic.energy.ca.gov/PublicDocuments/15-IEPR-03/TN205141_20150623T153206_CALIFORNIA_ENERGY_DEMAND_20162026_PRELIMINARY_ELECTRICITY_FOREC.pdf.

Choi, H., Anadón, L.D., 2014. The role of the complementary sector and its relationship with network formation and government policies in emerging sectors: the case of solar photovoltaics between 2001 and 2009. Technol. Forecast. Soc. Chang. 82: 80–94. http://dx.doi.org/10.1016/j.techfore.2013.06.002 (February).

Cole, W., Lewis, H., Sigrin, B., Margolis, R., 2016. Interactions of rooftop PV deployment with the capacity expansion of the bulk power system. Appl. Energy 168:473–481. http://dx.doi.org/10.1016/j.apenergy.2016.02.004 (April).

CPUC, 2015a. Decision adopting successor to net energy metering tariff. http://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M156/K443/156443378.PDF.

CPUC, 2015b. Rulemaking 14-07-002. http://docs.cpuc.ca.gov/PublishedDocs/Efile/G000/M152/K410/152410786.PDF.

Darghouth, N.R., Wiser, R.H., Barbose, G.L., Mills, A.D., 2015. Net Metering and Market Feedback Loops: Exploring the Impact of Retail Rate Design on Distributed PV Deployment.

Davidson, C., Drury, E., Lopez, A., Elmore, R., Margolis, R., 2014. Modeling photovoltaic diffusion: an analysis of geospatial datasets. Environ. Res. Lett. 9 (7):074009. http://dx.doi.org/10.1088/1748-9326/9/7/074009.

Davidson, C., Margolis, E., 2015. Selecting Solar: Insights into Residential Photovoltaic (PV) Quote Variation. NREL/TP-6A20-64792. National Renewable Energy Laboratory, Golden, CO.

de La Tour, A., Glachant, M., Ménière, Y., 2013. Predicting the costs of photovoltaic solar modules in 2020 using experience curve models. Energy 62:341–348. http://dx.doi.org/10.1016/j.energy.2013.09.037 (December).

Denholm, P., Drury, E., Margolis, R., 2009. The solar deployment systems (SolarDS) model: documentation and sample results. NREL/TP-6A0-44073. National Renewable Energy Laboratory, Golden, CO.

Denholm, P., Margolis, R., Palmintier, B., Barrows, C., Ibanez, E., Bird, L., Zuboy, J., 2014. Methods for analyzing the benefits and costs of distributed photovoltaic generation to the U.S. electric utility system. NREL/TP-6A20-62447. National Renewable Energy Laboratory (NREL), Golden, CO.

Dernburg, T.F., 1957. Consumer response to innovation: television. In: Dernburg, T.F., Rosett, R.N., Watts, H.W. (Eds.), Studies in Household Economic Behavior. Yale University Press, New Haven, CT :pp. 3–49. http://econpapers.repec.org/paper/cwlcwldpp/36.htm.

DOE (U.S. Department of Energy), 2012. The SunShot Vision Study. DOE/GO-102012-3037 (Washington, DC).

DSIRE (Database of State Incentives for Renewables & Efficiency), 2016. Database of State Incentives for Renewables & Efficiency (Accessed January 15, 2016, www.dsireusa.org<http://www.dsireusa.org>).

Duesenberry, J.S., 1949. Income, Saving, and the Theory of Consumer Behavior. Harvard University Press.

E3, 2015. NEM Successor Tariff Public Tool. http://www.cpuc.ca.gov/NR/rdonlyres/BF21BE4D-EF73-48A6-BFDF-6F2ADEFC1BD9/0/FinalPublicTool070172015.zip.

EIA, 2009. Residential Energy Consumption Survey (RECS). http://www.eia.gov/consumption/residential/data/2009/.

EIA, 2013. Residential Demand Module of the National Energy Modeling System: Model Documentation 2013. http://www.eia.gov/forecasts/aeo/nems/documentation/residential/pdf/m067(2013).pdf.

EIA, 2015. Annual Energy Outlook 2015. http://www.eia.gov/forecasts/aeo/.

Farrell, J., 2014. Minnesota's Value of Solar: Can a Northern State's New Solar Policy Defuse Distributed Generation Battles? https://ilsr.org/wp-content/uploads/2014/04/MN-Value-of-Solar-from-ILSR.pdf

Forte, R.M., 2015. Mastering Predictive Analytics With R. Packt Publishing.

George, R., Wilcox, S., Anderberg, M., Perez, R., 2007. National Solar Radiation Database (NSRDB) - 10 km Gridded Hourly Solar Database. Proceedings of Solar 2007. Published by the American Solar Energy Society, Cleveland, OH.

Gillingham, K., Deng, H., Wiser, R.H., Darghouth, N.R., Nemet, G., Barbose, G.L., Rai, V., Dong, C.G., 2014. Deconstructing Solar Photovoltaic Pricing: The Role of Market Structure, Technology and Policy (December).

Gillingham, K., Deng, H., Wiser, R., Darghouth, N., Nemet, G., Barbose, G., Rai, V., Dong, C., 2016. Deconstructing Solar Photovoltaic Pricing: The Role of Market Structure, Technology, and Policy. Energy J. 37 (3), 231–250.

GTM, 2015. Investment Tax Credit Extension Will Increase Solar Installations 54% Through 2020. GTM Research, Boston, MA.

GTM/SEIA, 2014. U.S. Solar Market Insight Report, 2014 Year in Review. GTM Research and Solar Energy Industries Association, Boston, MA.

GTM/SEIA, 2015. U.S. Solar Market Insight Report, Q2 2015. GTM Research and Solar Energy Industries Association, Boston, MA.

Guidolin, M., Mortarino, C., 2010. Cross-country diffusion of photovoltaic systems: modelling choices and forecasts for National Adoption Patterns. Technol. Forecast. Soc. Chang. 77 (2):279–296. http://dx.doi.org/10.1016/j.techfore.2009.07.003.

Heeler, R.M., Hustad, T.P., 1980. Problems in predicting new product growth for consumer durables. Manag. Sci. 26 (10), 1007–1020.

Horsky, D., Simon, L.S., 1983. Advertising and the diffusion of new products. Mark. Sci. 2 (1):1–17. http://dx.doi.org/10.1287/mksc.2.1.1.

Hyndman, R., Athanasopoulos, G., 2013. Forecasting: Principles and Practice. OTexts.

Hyndman, R., Koehler, A., Ord, K., Snyder, R., 2008. Forecasting with Exponential Smoothing. Springer Series in Statistics. Springer Berlin Heidelberg, Berlin, Heidelberg. http://link.springer.com/10.1007/978-3-540-71918-2.

IPCC, 2014. Climate Change 2014: Mitigation of Climate Change (Summary for Policymakers). http://www.ipcc.ch/pdf/assessment-report/ar5/wg3/ipcc_wg3_ar5_summary-for-policymakers.pdf.

Jeuland, A.P., 1994. The Bass Model as a Tool to Uncover Empirical Generalizations in the Diffusion of Marketing. Working paper. University of Chicago.

Navigant Consulting, Inc., 2016. Virginia Solar Pathways Project: Study 1 - Distributed Solar Generation Integration and Best Practices Review. Dominion Virginia Power, Richmond, VA.

Jiang, Z., Jain, D.C., 2012. A generalized Norton–Bass model for multigeneration diffusion. Manag. Sci. 58 (10):1887–1897. http://dx.doi.org/10.1287/mnsc.1120.1529.

Jordan, D.C., Kurtz, S.R., 2013. Photovoltaic degradation rates—an analytical review. Prog. Photovolt. Res. Appl. 21 (1):12–29. http://dx.doi.org/10.1002/pip.1182.

Kastovich, J.C., Lawrence, R.R., Hoffmann, R.R., Pavlak, C., 1982. Advanced electric heat pump market and business analysis. ORNL/Sub/79-2471/1. Oak Ridge National Laboratory, Oak Ridge, TN.

Lehr, R., 2013. New Utility Business Models: Utility and Regulatory Models for the Modern Era. America's Power Plan. http://americaspowerplan.com/site/wp-content/uploads/2013/10/APP-UTILITIES.pdf.

Meade, N., Islam, T., 2006. Modelling and forecasting the diffusion of innovation – a 25-year review. Int. J. Forecast. 22 (3):519–545. http://dx.doi.org/10.1016/j.ijforecast.2006.01.005 (Twenty five years of forecasting).

Mills, A., Barbose, G., Seel, J., Dong, C., Mai, T., Sigrin, B., 2016. Planning for a Distributed Disruption: Innovative Practices for Incorporating Distributed Solar into Utility Planning. LBNL-1006047. Lawrence Berkeley National Laboratory, Berkeley, CA.

Nemet, G.F., 2006. Beyond the learning curve: factors influencing cost reductions in photovoltaics. Energy Policy 34 (17):3218–3232. http://dx.doi.org/10.1016/j.enpol.2005.06.020.

Noll, D., Dawes, C., Rai, V., 2014. Solar community organizations and active peer effects in the adoption of residential PV. Energy Policy 67 (April):330–343. http://dx.doi.org/10.1016/j.enpol.2013.12.050.

Norton, J.A., Bass, F.M., 1987. A diffusion theory model of adoption and substitution for successive generations of high-technology products. Manag. Sci. 33 (9):1069–1086. http://dx.doi.org/10.1287/mnsc.33.9.1069.

OpenEI, 2014. U.S. Utility Rate Database. Accessed December 15, 2014. http://en.openei.org/wiki/Utility_Rate_Database.

Paidipati, J., Frantzis, L., Sawyer, H., Kurrash, A., 2008. Rooftop photovoltaics market penetration scenarios. NREL/SR-581-42306. Navigant Consulting, Inc., Burlington, MA.

Palmintier, B., Robert, B., Barry, M., Michael, C., Kyri, B., Fei, D., Matthew, R., Matthew, L., Ashwini, B., 2016. On the Path to SunShot: Emerging Issues and Challenges in Integrating Solar with the Distribution System. National Renewable Energy Laboratory, Golden, CO (NREL/TP-5D00-65331. http://www.nrel.gov/docs/fy16osti/65331.pdf).

Rábago, K., Libby, L., Harvey, T., Norris, B., Hoff, T., 2012. Designing Austin Energy's Solar Tariff Using a Distributed PV Value Calculator. http://www.cleanpower.com/wp-content/uploads/090_DesigningAustinEnergysSolarTariff.pdf.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021

ARTICLE IN PRESS

Pursuant to the DOE Public Access Plan, this document represents the authors' peer-reviewed, accepted manuscript. The published version of the article is available from the relevant publisher.

Rai, V., Robinson, S.A., 2013. Effective information channels for reducing costs of environmentally-friendly technologies: evidence from residential PV markets. Environ. Res. Lett. 8 (1):014044. http://dx.doi.org/10.1088/1748-9326/8/1/014044.

Rai, V., Robinson, S.A., 2015. Agent-based modeling of energy technology adoption: empirical integration of social, behavioral, economic, and environmental factors. Environ. Model. Softw. 70:163–177. http://dx.doi.org/10.1016/j.envsoft.2015.04.014 (August).

Rai, V., Sigrin, B., 2013. Diffusion of environmentally-friendly energy technologies: buy versus lease differences in residential PV markets. Environ. Res. Lett. 8 (1):014022. http://dx.doi.org/10.1088/1748-9326/8/1/014022.

Rai, V., Cale Reeves, D., Margolis, R., 2016. Overcoming barriers and uncertainties in the adoption of residential solar PV. Renew. Energy 89:498–505. http://dx.doi.org/10.1016/j.renene.2015.11.080 (April).

Randazzo, R., 2015. Regulators Delay APS Solar-Fee Decision. http://www.azcentral.com/story/money/business/consumer/2015/08/19/regulators-delay-aps-solar-fee-decision/31963553/.

Richter, M., 2013. Business model innovation for sustainable energy: German utilities and renewable energy. Energ Policy 62:1226–1237. http://dx.doi.org/10.1016/j.enpol.2013.05.038 (November).

RMI, 2013. A Review of Solar PV Benefit & Cost Studies. second ed. Rocky Mountain Institute, Snowmass, CO.

Robinson, B., Lakhani, C., 1975. Dynamic price models for new-product planning. Manag. Sci. 21 (10):1113–1122. http://dx.doi.org/10.2307/2629953.

Robinson, S., Rai, V., 2015. Determinants of spatio-temporal patterns of energy technology adoption: an agent-based modeling approach. Appl. Energy 151 (August):273–284. http://dx.doi.org/10.1016/j.apenergy.2015.04.071.

Rogers, E.M., 2003. Diffusion of Innovations, 5th Edition. fifth ed. Free Press, New York.

Satchwell, A., Mills, A.D., Barbose, G.L., Wiser, R.H., Cappers, P., Darghouth, N.R., 2014. Financial Impacts of Net-metered PV on Utilities and Ratepayers: A Scoping Study of Two Prototypical U.S. Utilities (September).

Schmittlein, D.C., Mahajan, V., 1982. Maximum likelihood estimation for an innovation diffusion model of new product acceptance. Mark. Sci. 1 (1), 57–78.

Sigrin, B., Drury, E., 2014. Diffusion into New Markets: Economic Returns Required by Households to Adopt Rooftop Photovoltaics. AAAI Energy Market Prediction Symposium. Washington, November 13-15, 2014.

Sigrin, B., Gleason, M., Preus, R., Baring-Gould, I., Margolis, R., 2016. The distributed generation market demand model (dGen): documentation. NREL/TP-6A20-65231. National Renewable Energy Laboratory, Golden, CO.

Srinivasan, V., Mason, C.H., 1986. Technical note—nonlinear least squares estimation of new product diffusion models. Mark. Sci. 5 (2):169–178. http://dx.doi.org/10.1287/mksc.5.2.169.

Sultan, F., Farley, J.U., Lehmann, D.R., 1990. A meta-analysis of applications of diffusion models. J. Mark. Res. 27 (1):70–77. http://dx.doi.org/10.2307/3172552.

Taylor, M., McLaren, J., Cory, K., Davidovich, T., 2015. Value of solar: program design and implementation considerations. NREL/TP-6A20-62361. National Renewable Energy Laboratory, Golden, CO.

U.S. Congress, 2015. Text of House Amendment #1 to the Senate Amendment to H.R. 2029, Military Construction and Veterans Affairs and Related Agencies Appropriations Act, 2016. http://docs.house.gov/billsthisweek/20151214/CPRT-114-HPRT-RU00-SAHR2029-AMNT1final.pdf.

van den Bulte, C., Stremersch, S., 2004. Social contagion and income heterogeneity in new product diffusion: a meta-analytic test. Mark. Sci. 23 (4), 530–544.

Changgui Dong is an assistant professor in Public Policy at Renmin University of China, School of Public Administration and Policy. He is also a research fellow at the National Academy of Development and Strategy, RUC. He was a post-doctoral researcher at the National Renewable Energy Laboratory from 2015 to 2016. His research mainly focuses on renewable energy technology diffusion, technological change and policy evaluation. He obtained his PhD in Public Policy from the University of Texas at Austin in the United States.

Benjamin Sigrin is an energy systems modeling engineer in the Energy Forecasting and Modeling Group at the National Renewable Energy Laboratory. He is specialized in financial modeling and capacity expansion modeling of the electric sector. He obtained his M.S. in energy and earth resources and M.P.A. in public affairs from the University of Texas at Austin in 2013.

Gregory Brinkman is an energy analysis engineer in the Energy Forecasting and Modeling Group at the National Renewable Energy Laboratory. He is specialized in unit commitment and dispatch modeling of electric power systems. He obtained his PhD in mechanical engineering from the University of Colorado, Boulder CO in 2009.

Please cite this article as: Dong, C., et al., Forecasting residential solar photovoltaic deployment in California, Technol. Forecast. Soc. Change (2016), http://dx.doi.org/10.1016/j.techfore.2016.11.021