Applied Energy 151 (2015) 273–284

Contents lists available at ScienceDirect

Applied Energy

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

Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach

Scott A. Robinson [a,b], Varun Rai [a,c,*]

[a] LBJ School of Public Affairs, The University of Texas at Austin, United States
[b] Jackson School of Geosciences, The University of Texas at Austin, United States
[c] Department of Mechanical Engineering, The University of Texas at Austin, United States

HIGHLIGHTS

- We present an agent-based model of residential solar photovoltaic (PV) adoption.
- Model integrates social, behavioral, and economic elements of agent decision-making.
- Real-world, large-scale integrated dataset used for model validation and testing.
- We study the importance of using disaggregated empirical data on model performance.
- Social and attitudinal components are critical for spatial and demographic accuracy.

ARTICLE INFO

Article history:
Received 20 September 2014
Received in revised form 16 April 2015
Accepted 17 April 2015
Available online 15 May 2015

Keywords:
Agent-based modeling
Solar photovoltaic (PV)
Complex systems
Technology adoption
Bounded rationality
Residential energy

ABSTRACT

Energy technology adoption is a complex process, involving social, behavioral, and economic factors that impact individual decision-making. This paper uses an empirical, geographic information system (GIS)-integrated agent-based model of residential solar photovoltaic (PV) adoption to explore the importance of using empirical household-level data and of incorporating economic as well as social and behavioral factors on model outcomes. Our goal is to identify features of the model that are most critical to successful prediction of the temporal, spatial, and demographic patterns that characterize the technology adoption process for solar PV. Agent variables, topology, and environment are derived from detailed and comprehensive real-world data between 2004 and 2013 in Austin (Texas, USA). Four variations of the model are developed, each with a different level of complexity and empirical characterization. We find that while an explicit focus only on the financial aspects of the solar PV adoption decision performs well in predicting the rate and scale of adoption, accounting for agent-level attitude and social interactions are critical for predicting spatial and demographic patterns of adoption with high accuracy.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Demand-side behavior has important implications for local and global emissions reductions [1–4] and for the future of the electric grid [5–9]. In particular, a robust understanding of the rate and pattern of consumer adoption of durable energy technologies is critical for forecasting energy demand and emissions, as well as for infrastructure planning and development [6,10–12]. Modeling of energy technology adoption is particularly challenging, since the nominal economics (price) of the technology is but just one determinant of consumers' likelihood to adopt [13–15]. Other behavioral and social phenomenon such as decision heuristics, anchoring, path-dependence (past experiences), risk aversion, trust-based information networks, and social norms are also quite important in understanding decision-makers with bounded rationality in general [16–18], and energy-related consumer decision-making in particular [2,14,15,19–22]. As such, the development of analytical techniques that are able to appropriately represent and model the bounded rationality of economic agents, including the relevant social and spatial factors, is important for better understanding of the technology adoption process and the resultant emergent phenomena [23,24].

Agent-based modeling (ABM) has emerged as a methodology that provides a suitable framework for explicitly modeling decision-makers with bounded rationality, their social interactions, and the (physical and economic) environments surrounding them

* Corresponding author at: 2315 Red River St., LBJ School of Public Affairs, The University of Texas at Austin, TX 78712, United States.

E-mail address: raivarun@utexas.edu (V. Rai).

http://dx.doi.org/10.1016/j.apenergy.2015.04.071
0306-2619/© 2015 Elsevier Ltd. All rights reserved.

[23,25–28]. Energy and environment related consumer technology adoption has been a particular area of growth in the development and applications of ABM [29–44]. Because the underlying components of the system and how they interact with each other are modeled explicitly in ABM, the processes that lead to observable emergent phenomena (such as the rate and pattern of adoption) can be altered through simulation experiments, creating virtual laboratories [24,45,46]. The depth, generalizability, and flexibility of ABM make it applicable to a wide range of problems such as the modeling of traffic patterns, growth of civilizations, land use change, group dynamics, molecular self-assembly, electricity markets, and stock markets (see [24] for a more comprehensive review). However, there are important challenges for ABM in consumer energy technology adoption – and human-technical systems in general – especially regarding the integration of theoretical elements and empirical patterns in the model structure, initialization, and validation efforts. While the potential of ABM in enabling detailed bottom-up modeling of technology adoption is quite promising, ABM has been criticized on two important fronts [47–50]: (i) agent decision rules in ABM are often over simplified or even *ad hoc*, rendering connections to the broader theoretical context difficult, and (ii) models often have inadequate empirical emphasis on initialization and validation against real-world data [51]. These factors have drawn increasing attention in the literature toward the importance of methodological rigor and the use of adequately-resolved empirical data within ABM [49,52–54]. It is now recognized in the literature that the extent to which the strengths of ABM techniques offer an advantage over conventional modeling techniques in policy analysis and system design is not a given, rather it depends critically upon careful theoretical and empirical underpinning of agent-based models [47,54].

This paper uses a theoretically and empirically grounded agent-based model of residential solar photovoltaic (PV) adoption (henceforth, the "solar ABM") to analyze the importance of using localized (disaggregated) empirical data and of including social and attitudinal components in the adoption model in addition to purely economic factors. Specifically, we develop four different variations of the solar ABM – "Base-case", "Simple Environment," "Random Fitted," and "Economic Only" – to study residential solar PV adoption through a thorough integration of economic valuation, attitudinal evolution, and social interactions. Each of the four models has a different level of model complexity and empirical characterization. Using a rich and comprehensive dataset between 2004 and 2013, the models simulate the adoption of residential solar PV in the city of Austin (Texas, USA), which has a population of approximately 900,000. Using the four variations of the solar ABM we systematically examine the effect of progressively increasing the empirical basis and the complexity of agent-based models on model outcomes through external validation. We emphasize that these are four different *models*, not just different scenarios – as we discuss later, they vary in the basic model formulation in important ways. We focus on aspects of model fitting and validation, with the goal of identifying features of the solar ABM that are most critical for accurately describing the solar PV adoption process. We analyze the cost (in terms of predictive power) of decreasing the empirical foundation and complexity of the model.

We chose solar PV as the empirical test-bed in our study for two reasons: (i) the growing importance and impact of solar PV in the electric industry globally [55,56], and (ii) the relatively complex decision-making process associated with solar PV adoption, which offers a unique opportunity to study and quantify how economic, attitudinal, and social factors impact individual behavior and lead to emergent phenomena [57,58]. The two main contributions of this paper are: (i) identification of variables and processes key to the successful modeling of residential solar PV adoption using an ABM approach, and (ii) detailed comparisons of different model variations in order to quantify the value of increasing model complexity and of using empirical distributions in terms of increased accuracy of the model for predicting the rate *and* pattern of solar PV adoption.

2. Material and methods

In this section we provide a conceptual overview of the model components and how they fit together. All model components were integrated in the R programming language and additional supporting methods were written in Python. We also briefly describe the integrated dataset and the validation procedures that are used for the analysis in this study. A more comprehensive discussion of the underlying data and methodology is covered elsewhere [59]. All simulations were run on the 10PF Stampede Supercomputer at the Texas Advanced Computing Center (TACC), utilizing 16 tasks per server node (each with two 350GF Intel Xeon E5-2680 processors and one 1070GF Intel Xeon Phi SE10P Coprocessor) on 100 nodes per batch (1 batch = 100 simulations). Depending on the exact specification, each batch took between 20 and 35 min to execute.

2.1. Data

We use a granular household-level dataset including: (i) *for solar adopters* ($N = 2738$): time-series utility solar program data (rebate; price; system technical details; timing of adoption) and survey data (attitude; motivators; perception; information seeking; financial aspects)[1] and (ii) *for all households* ($N = 173,466$): geo-location, home value, and environmental variables (roof size; lot size; tree cover; elevation; slope; shading; insolation). The utility solar program data for Austin ranges from 2004 to mid-2013. Additional datasets including solar-adopter surveys, appraisal district data, and light detection and ranging (LiDAR) data were overlaid upon the solar program data to build the comprehensive and granular integrated dataset. The solar program data for *each* installation were matched to geocoded addresses, allowing for the analysis of solar adopter distribution over space and time. Each geocoded address was matched to a single family residential parcel from the Travis County Appraisal District, and a home footprint from the City of Austin. These polygons were overlaid with household-level home value and environmental layers in a geographic information system (GIS) to define the agent attributes. Finally, solar adopter survey data were joined to the solar program data. These data streams were combined to create agent and environment classes that were rigorously grounded in real-world data and closely reflected the actual population they were intended to represent, allowing for an empirical ABM methodology [59].

2.2. Model design

The solar ABM used in this paper is a *household-level agent-based model* able to generate the empirically observed temporal and spatial patterns of the adoption of residential solar [59]. There is only one agent type – a household. The number of agents in the solar ABM is 173,466: all the actual single-family residential households in Austin, Texas as of mid-2013. The study period is from 2004 to June 2013, during which the solar adoption level in Austin grew from only $N = 20$ to 2738. Data from 2004 through 2007 is used for initialization, while 2008–mid-2013 is the simulation period.

[1] For the solar adopter survey data, $N = 616$ (22.5% response rate).

Two main elements in our model determine the decision of agents to adopt or not adopt solar: an attitudinal component ("attitude") and a control component ("control"). Both of these components are dynamic in the model. To adopt, an agent i must have sufficiently high levels of both attitude and control (we discuss the calculation of these attributes in Section 2.3). The sufficiency of an agent's attitude (sia_i) means it has to be higher than an attitudinal threshold (sia^{thresh}). Thus, the attitudinal rule is: $sia_i > sia^{thresh}$. Note that sia^{thresh} is a global parameter, in the sense that it is common to all agents. The sufficiency of an agent's control is determined by an agent-level threshold attribute (pbc_i), which corresponds to the minimum acceptable financial return agent i is willing to accept. Thus, the control rule is: $PP_{it} < pbc_i$, where PP_{it} is the payback period of the solar investment for agent i in time period t. We use payback period in the control criterion because the majority (87%) of the respondents in a survey of solar adopters in the study area reported using payback period as the financial decision criterion [60]. Adoption occurs when an agent meets both the attitudinal and the control criteria. The expression of the behavioral model in terms of *attitude* and *control* was motivated by the Theory of Planned Behavior (TPB) – a widely applied behavioral model in psychology [61–64] used frequently in ABM [36,39,43,59,65–68]. We note that while *pbc* in TPB is a more general concept than just the perception of financial attractiveness, given the strong influence of PV economics on PV adoption our focus on just the economic component of behavioral control is justified.

In each simulated quarter, a household's attitude *sia* can be altered by interactions with other agents. This attitudinal evolution is driven by the social component. The Relative Agreement (RA) algorithm [69–71] describes this evolutionary process. In the RA algorithm, large differences in the attitude of two agents will make influence exchange less likely, while agents with low uncertainty U (around their attitude) are more likely to influence others. A social network model is used to determine which agents interact with each other (see Section 2.3.3 for more details).

On the other hand, the dynamic nature of the control component of the decision model is driven by the changing economics of solar [57,60]. Agents compare their (static) threshold control attribute *pbc* (i.e., perceived economic ability) with the payback period on the solar system. As shown in Eq. (1), payback (in years) was calculated as a function of system cost and output (which were modeled at the household level) and of rebate levels and electricity prices corresponding to the actual values during a given time period t:

$$PP_{it} = (p_t - R_t - (p_t - R_t) \times ITC_t)/(G_i \times e_t), \tag{1}$$

where e is the value of the electricity produced by the solar system (in \$/kWh), p is the per unit price of the solar system (in \$/kW), R is the utility rebate (in \$/kW), *ITC* is the federal investment tax credit (0.3, i.e. 30%), and G is the annual system electricity generation (in kWh/kW/year, calculated based on insolation; see Section 2.3.2). Thus, an agent's economic evaluation changes through exogenous price and subsidy changes, which mirror the empirically observed levels. Note that except the system price p_t all other variables in Eq. (1) are known exactly for each agent. Using actual price data on all solar systems sold in the study area and time period, we model p_t using non-parametric local polynomial regression (LOESS) [59].

2.3. Initialization

Initializing the ABM meant assigning agent state values at t_0 (the end of 2007). Any empirically grounded ABM with dynamic attributes will need to address this requirement in order to proceed with model fitting and validation [72]. However, this important aspect is dealt with varying degrees of rigor in the literature, if not grossly neglected. In the solar ABM here, each agent is initialized with pbc, sia, and uncertainty (U) around the *sia*. Household-level insolation – a static attribute – is also calculated at initialization. In order to empirically situate the model temporally and best represent the ground truth as in Q4 2007, we carefully leverage the layered, longitudinal dataset we have developed (Section 2.1). We describe the initialization process below, additional details of which are covered elsewhere [59].

2.3.1. Attitude and control attributes

For the attitudinal component, we use a three-step process for calculations of initial attitude $sia_i(0)$ (*sia* at t_0) for all agents. Survey data were used to create an index of financial, environmental, and social factors influencing agent attitude. The index components were weighted by revealed individual preferences, also derived from the survey. The full weighted index was regressed on publicly available socio-economic demographic and home-feature variables. Finally, these estimates were used in a kriging spatial autocorrelation model to take advantage of any spatial patterns in the data for calculating $sia_i(0)$ for the non-adopters. Thus, we address the initialization requirement through an empirically-grounded statistical modeling approach using multiple data-streams, while avoiding the use of *ad hoc* random distributions (for additional details see [59]). Initial uncertainty U_i for agent i is set at $-1 \times |sia_i|$, with the assumption that agents with more extreme attitudes are likely to have lower uncertainty about their attitude [73,74]. We emphasize that the entire initialization process is grounded in real empirical data on the relevant attributes of the underlying population. As such, there is no subjectivity associated with the initialization process. However, clearly this is a data-intensive process, especially for the attitudinal portion. While reasonable proxies (such as home value) may be found for agents' financial capability, doing so for attitudinal metrics is not straightforward. Instead of proxies, we use survey data on the relevant population for calculating the attitudinal indices. Although collecting survey data is more expensive and time consuming, it significantly enhances the accuracy of the indices by empirically grounding agent attributes. We believe that finding suitable proxies for calculating attitudinal indices is a fruitful study area that could be explored in future research.

The control attribute *pbc*, assumed to be static in the model, is calculated using household-level insolation, tree cover, and financial resources (proxied by home value) – these are factors that may be expected to impact agents' perception of their ability to adopt solar:

$$pbc_i = \alpha_0 + \alpha_1 (I_i + W_i^* - Tr_i^*), \tag{2}$$

where I is the household-level insolation (discussed in Section 2.3.2), W^* is the home value, and Tr^* is the tree-cover ratio (shaded fraction of roof).[2] The advantage of this approach to calculate *pbc* is that these variables are available for the entire population of the agents ($N = 173,466$), allowing for a systematic, empirically-grounded initialization. The alternative would be to use survey data to estimate and initialize *pbc*. But, as we point out in [59], given the large number of agents in our model, that approach is both infeasible and/or prone to high measurement error. That is because the survey approach to calculating *pbc* would require a survey of *all* households in the study area at the *beginning* of the study period (i.e., early 2004) or asking all agents now about the perceptions they held back in 2004. The former is simply infeasible, while the latter is also prone to high measurement error.[3]

[2] The $*$ on W and Tr indicate that these variables were normalized to account for differences in scale compared to I.

[3] Another possibility is to measure current perceptions and make the additional assumption that those same perceptions were held throughout the study period.

2.3.2. Household-level insolation

Variations in slope, aspect, elevation, and shadowing all effect the amount of direct and indirect sunlight that reaches a given area over a period of time (insolation, I). This heterogeneity directly impacts the generation potential of the system, and thus the economic value of the system for a household (Eq. (1)). In hilly areas, these differences can be dramatic, making regional average values less useful. Insolation was modeled for the location of each household by summing the direct (I_{dir}) and indirect I_{dif} sunlight received at each location on a raster grid [75] over a one year period:

$$I_i = I_{dir} + I_{dif}. \tag{3}$$

Direct radiation was modeled by creating a two dimensional representation of the sun's relative path over time for each location, which was divided into sectors with zenith θ and azimuth α angles at their centroid. The total direct radiation for a location is the sum of the direct radiation from each sector $Dir_{\theta,\alpha}$, which is a function of the solar constant S, average atmospheric transmissivity β along the angle θ, the relative optical path length $m(\theta)$ influenced by elevation [75,76], the time duration $SunDur$ for the sector, the gap fraction $SunGap$ for the sector, and the angle of incidence between the sector centroid and the surface, $AngleIn$:

$$\begin{aligned} I_{dir} &= \sum Dir_{\theta,\alpha} \\ &= S \times \beta^{m(\theta)} \times SunDur_{\theta,\alpha} \times SunGap_{\theta,\alpha} \times \cos(AngleIn_{\theta,\alpha}). \end{aligned} \tag{4}$$

Indirect radiation was modeled by creating a two dimensional representation of the entire sky for each location, also divided into sectors. The indirect radiation is a function of the global normal radiation R_g, the ratio of global normal radiation flux diffused P_{diff}, the time interval Dur, the proportion of visible sky $SkyGap$ for the sector, and the relative amount of diffuse radiation $Weight$ for the sector:

$$\begin{aligned} I_{dif} &= \sum Dif_{\theta,\alpha} \\ &= R_g \times P_{dif} \times Dur \times SkyGap_{\theta,\alpha} \times Weight_{\theta,\alpha} \\ &\quad \times \cos(AngleIn_{\theta,\alpha}). \end{aligned} \tag{5}$$

The above calculations were performed using ESRI's solar radiation toolset [77]. Following [78], the average total insolation for each household's roof was used to calculate system generation G_i used in Eq. (1).[4]

2.3.3. Social network model

Unless noted otherwise, we use a small-world network where λ^r connections were random and $1 - \lambda^r$ connections were locals [79]. Locals were defined geographically by a radius parameter ($r = 2000$ ft) around each household. Economic similarity (calculated as the squared difference in home values) was used to further refine the locals set, whereby only 5% of the geographic locals set was retained in the final locals set used in the simulations. This resulted in a median of 27 neighbors per agent in the network. As a final step $\lambda^r (= 0.1)$ fraction of these connections were randomly rewired with agents anywhere in the study area. Thus, in our social network model the majority of an agent's connections are geographic and economic neighbors. That is, a household is more likely to be connected with (and hence interact with) other households that are nearby and have similar wealth characteristics. Fig. 1 shows a portion of the resulting network. As noted in Section 2.2, the social network model is central to the evolution of agents' attitude sia_i and uncertainty about the attitude U_i. Every time step (quarter), each agent interacts with ϕ other agents randomly selected from her social network. These interactions form the basis of updating an agent's attitude according to the Relative Agreement (RA) algorithm [80], which is commonly used for modeling agent interactions in opinion dynamics models [68,71,81–84]. In the RA algorithm the level of attitude modification for an agent i following an interaction with agent j depends upon the overlap in their attitudes. In general, the RA algorithm does not permit two agents with radically different opinions to influence each other significantly. Thus agents are largely influenced by other agents with relatively similar attitudes. This property allows for clusters of interconnected agents to reinforce each other's opinions, eventually converging to local stable values [69]. The speed of this convergence is controlled by the coefficient μ, which acts as a weighting term on each agent-agent interaction. Higher levels of μ result in speedier convergence.

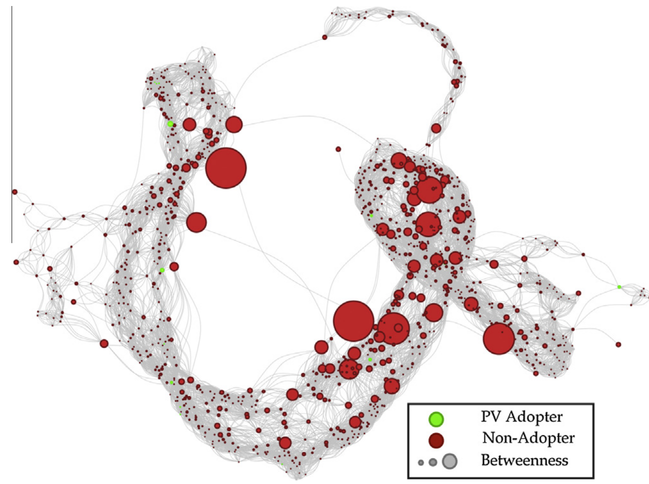

Fig. 1. Representative portion of the small world network used in the solar ABM. The shown network is only for the $N = 1054$ agents in a single zip code in the study area (the full network has $N = 173,466$ nodes). Nodes represent individual household agents, and edges represent agents connected in the network. While the underlying network uses actual geographical location of the agents to define the locals set (see text), for a convenient visualization, length of the edges do not represent actual geographical distance. Green dots are solar adopters ($N = 15$) in the zip code used for the visualization, red dots are non-adopters, and size of the dots reflects the betweenness centrality of each node. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Model variations

Using the general framework described above in Section 2.2, we develop four model variations (Table 1): "Base-case", "Simple Environment," "Random Fitted," and "Economic Only." These four model variations were chosen on the basis of empirical and theoretical relevance and represent significantly differentiated modeling approaches. The Base-case is defined first, and the other models are derived from it by stripping down specific components and variables, thus quantifying the effects of increased simplicity or lack of empirical data in terms of fitting and validation.

2.4.1. Base-case

The Base-case model is the full model that uses all model components (attitudinal, economic, and social networks) as described in Section 2.2. In addition, the Base-case model uses empirical distributions from the integrated dataset (Section 2.1) for initializing all agent states and for the empirical description of all environmental factors at the household-level (Section 2.3).[5]

[4] As will be discussed below, in the 'Simple Environment' model aggregate city-level insolation value is used instead of a household-level value.

[5] It should be noted that the Base-case model is equivalent to the "Integrated Model" in Rai and Robinson [59].

Table 1
Model features that have major differences between the four models analyzed in this study. In all models, PV economics is based on comparison of simple payback to *pbc* (see Section 2.2). Additional details on each of the models is provided in the text. **BC**: Base case; **SE**: Simple Environment; **EO**: Economic Only; **RF**: Random Fitted.

Model features				
Model	Initialization	Social network	Attitudes	Insolation resolution
BC	Empirical *pbc*, *sia*, and *U* distributions at t_0	Geographically and economically based small-world networks. Mostly, agents are connected if in the same neighborhood and have similar economic status	Attitudes toward solar PV are modeled explicitly and in parallel with economic factors through the *sia* term	Individual household level (Section 2.3.2)
SE	Empirical *pbc*, *sia*, and *U* distributions at t_0. *pbc* generated according to Eq. 6	Geographically and economically based small-world networks. Mostly, agents are connected if in the same neighborhood and have similar economic status	Attitudes toward solar PV are modeled explicitly and in parallel with economic factors through the *sia* term	City average value used, leading to the same value of *G* for all agents (Eq. 1): 1361 kW h/kW/year
EO	Empirical *pbc* distribution at t_0	No social network model. Agents do not gain or share information through contact with others, but rather have "perfect" information/attitude about solar PV	Attitudes toward solar PV play no role in the model	Individual household level (Section 2.3.2)
RF	Random *pbc*, *sia*, and *U* distributions at t_0	Erdős-Rényi random network, such that connections between any two agents occur with the same (binomial) probability	Attitudes toward solar PV are modeled explicitly and in parallel with economic factors through the *sia* term	No insolation value was used, as the payback period and *pbc* were generated through draws from random distributions

2.4.2. Simple Environment

The Simple Environment model was created in order to study how a simpler economic module closer to those used in practice by utility planners affects fit and validation of the solar ABM. In this model, empirical data were used to initialize the *pbc*, *sia*, and *U* distributions at t_0, thus creating a fully empirically initialized model. However, both in the literature discussing drivers of solar adoption as well as in practice (for example in electric utility estimates) the level of solar insolation is often left at the regional or city level of aggregation. Modelers often make such aggregation choices either to reduce modeling complexity or due to lack of data or both. Accordingly, to study the impact of using aggregate insolation data rather than using the actual insolation levels for each household as in the Base case, in the Simple Environment model the simple average insolation value for the city of Austin was used. As a result, for all agents the value of *G* used in Eq. (1) in the Simple Environment model was 1361 kWh/kW/year – the electricity generated per kW of a solar PV system per year using city-level average insolation value. This clearly impacts calculations of the payback period (Eq. (1)). It also impacts the calculation of the control variable *pbc*. Because insolation does not vary by household in the Simple Environment model, as shown in Eq. (6) only home-value W^* and the tree cover ratio Tr^* were used in calculating the *pbc* index in the Simple Environment model:

$$pbc_i = \alpha_0 + \alpha_1\left(W_i^* - Tr_i^*\right) \tag{6}$$

In contrast, the 'Base-case' and 'Economic Only' models also use household-level insolation for *pbc* calculations (see Eq. (2)).

2.4.3. Random fitted

One of our major emphases in this paper is the importance of using empirical data to inform the model during initialization. However, these distributions may not be available, even while the theoretical underpinning and reasonable parameter estimates may be accessible through existing literature or aggregate data. It is worth considering how successful the model can be under these circumstances. In the Random Fitted model, empirical distributions are *not* used to initialize the agents' state; rather these distributions are taken to be unknown, and thus are approximated using random distributions with variable parameters (i.e. shape and scale, or mean and variance). We tested uniform, normal, poisson, beta, and gamma distributions for the *pbc*, *sia*, payback period (*PP*), and agent uncertainty (*U*). The end result is that agent state variables are realizations of the specified random distributions: for example, pbc_i, rather than being determined by Eq. (6) (or its Base-case variant shown in Footnote 3), is determined by a random draw from a Poisson distribution with an expected value of 3. Finally, the social network model is also random – we use an Erdős–Rényi random graph, wherein any two agents are equally likely to be connected in the social network.

2.4.4. Economic Only

Recall that the Base-case solar ABM uses fully integrated economic, attitudinal, and social modules to model the solar technology adoption process. In order to test the importance of social interactions and household attitudes, in the Economic Only model we ran and fit the economic module of the solar ABM in isolation, by entirely leaving out the social component of the Base-case. The decision to adopt in this model was thus based only on the financial (control) criterion: each agent's comparison of their perceived control *pbc* against the payback *PP* on a solar system at a given time.

2.5. Fitting

Parameter values were chosen though a fitting process to minimize the objective function of the model, namely the root mean squared error (RMSE) in the cumulative number of installations. As shown in Eq. (7), this was assessed by comparing the predicted cumulative adoption levels $\hat{a}_t$ to the cumulative empirical level a_t over $n = 22$ quarters (from Q1 2008 to Q2 2013):

$$RMSE = \sqrt{\sum_{t=1}^{n}\left(\frac{(\hat{a}_t - a_t)^2}{n}\right)}. \tag{7}$$

Both fitting and validation (Section 2.6) metrics for each model were calculated for a "batch" run. Each batch consisted of 100 runs of the same model with the same parameters, used in order to account for randomness in the model (for example, randomness associated with the selection of which agents interact each time period). While there are multiple outcome variables for each model, only the cumulative number of solar system installations is used for fitting. As discussed next, this allows for rigorous validation of the model on unfitted criteria.

2.6. Validation

In our systematic comparison of models with varying empirical basis and complexity (Section 2.4), we validate the models across three model outcomes: temporal, spatial, and demographic. While the core function of the validation metrics is to measure each model's ability to accurately represent solar technology adoption, adopter demographics and location are critical to policy evaluation and planning, making their successful prediction of even greater interest. Below we describe the temporal, spatial, and demographic validation process.

1. *Temporal Validation*. The number of new solar adopters in a given time-period provides policy-makers and planners with a simple indicator of price elasticity, for example after lowering rebate levels. Temporal validation is done on the instantaneous rate of quarterly adoption – the marginal number of adoptions predicted for each quarter. Validated models will display low RMSE according to Eq. (7) (where the predicted and empirical values for quarterly adoption are used for $\hat{a}_t$ and a_t, respectively). Short term volatility, time periods with rapid or slow swings in adoption due to price and rebate changes make this metric adequately independent of the fitting criteria (cumulative adoption rate) discussed in Section 2.5.
2. *Spatial Validation*. How solar adopters are distributed over space can in part determine the value of the electricity produced by the solar systems to the electric utility as well as the need for potential infrastructure improvements [56]. In a spatially validated model, neighborhoods of high and low adoption will match those observed in the empirical data. In order to evaluate this similarity, we created Gaussian density rasters (grids) at 100 ft resolution using ESRI's KernelDensity() function in ArcPy for each simulation as well as for the empirical adopter coordinates. We used three metrics to assess the spatial similarity: simple arithmetic error where the predicted density is subtracted from the observed density cell-by-cell; fuzzy numerical similarity (κ^*) [85,86], and the wavelet correlation coefficient (r^w) [87,88]. Both κ^* and r^w allow for a degree of local smoothing, which better captures the similarity between maps where the structure is equivalent but individual point values may not match up exactly [59].
   The fuzzy numerical statistic κ^* between the empirical raster A and the simulated raster B was calculated according to Eq. (8) [85,86], where N is the total number of cells in the raster maps. κ_i is the maximum distance-weighted similarity for cell i in raster A with raster B. The maximum is calculated over all the n cells in a 20 cell radius from cell i. For the smoothing function ($2^{-d/5}$), a simple exponential decay function was used, with a half-life of five cells. d is the number of cells in the raster separating cells i and j. The function f used in calculating κ_i is a measure of similarity between two values a and b on the empirical map and the simulated map, respectively.

$$\begin{aligned}
\kappa^*(A,B) &= \frac{1}{N}\sum_{i=1}^{N}\kappa_i(A,B), \\
\kappa_i(A,B) &= max_j^n(f(A_i,B_j)\times 2^{-d/5}), \\
f(a,b) &= 1-\frac{|a-b|}{max(|a|,|b|)}.
\end{aligned} \tag{8}$$

3. *Demographic Validation*. Solar adopters typically have several distinguishing demographic characteristics, including higher than average wealth. As of 2013, the average home value for solar adopters in Austin, TX ($475,326) was over 77% higher than that of non-adopters on average ($267,966). Over time, this gap has been gradually decreasing in the study area. This is of considerable policy interest due to the offered rebates that are allocated from public funds, raising questions of equity [89,90]. To validate the model's ability to generate this disparity as well as the trend over time, we calculate the RMSE in median adopter home-value between the simulation and the empirical data. In order to further validate the model demographically we calculate the κ^* and r^w statistics for raster maps of adopter home-value over space.[6] This was intended to capture any local variation not accounted for in the global comparison using just the RMSE.

3. Results

In this section we present the results of fitting and validation for the four models presented in Section 2.4: "Base-case", "Simple Environment," "Random Fitted," and "Economic Only." Recall that model fitting is the basis for parameter value selection, and is done only on the RMSE in the cumulative installation levels over time (Section 2.5).

3.1. Base-case

The Base-case model was fit using six parameters (Table 2): the number of other agents a household would interact with over each quarter (ϕ), the weight on each interaction (μ) within the RA algorithm, the proportion of random connections in each household's network (λ^r), the intercept (α_0) and slope (α_1) used to scale the initial *pbc* distribution (see Section 2.4), and the global attitude threshold sia^{thresh}. Base-case model fit (RMSE in cumulative adoptions) was found to be 117.81, the best of any model attempted (Fig. 2). This result is noteworthy given that the Base-case has the same number of parameters (six) as the other models (apart from the Economic Only case, which has only two parameters).

Validation of the Base-case strongly supports the representation of the solar adoption process as a function of economic, attitudinal, and social factors as described in Section 2.2. Overall, there was little or no evidence of over-fitting, as demonstrated by the validation metrics in Table 3. The marginal rate of adoption predicted by the model matched well with that seen in the empirical data, and was the second lowest of any model attempted. Most of the remainder error can be attributed to Q4 2011, in which strong price signals experienced in the solar market resulted in the model over-predicting adoption. The relatively large error generated in this quarter resulted in the Base-case model being out-performed in terms of temporal validation by the Random Fitted model. This is discussed further in Section 3.3. The Base-case model predicts the spatial pattern of adopters observed in Austin more accurately than the simpler (fewer components) and less empirically grounded model variations discussed later. Average simple error in spatial validation was 0.46. This is also shown in Fig. 3. The Base-case predicts most of the local solar-dense neighborhoods, if not always with the exact magnitude. The average fuzzy similarity (κ^*) between the observed density and the predicted density was 0.43 and the wavelet correlation (r^w) was 0.86. The spatial demographic predictions and simple error are shown in Fig. 4. The demographic accuracy of the model was also quite high, both over time and space. Temporally, the RMSE in median adopter home values was the lowest of any model (110,580.2), and the overall temporal pattern matched (gradual decline). The spatial demographic accuracy was the highest of all the models tested

[6] Rasters were generated using Inverse Distance Weighted (IDW) interpolation [91], as density would not be an appropriate measure for home-value. For the IDW calculation, we used a variable radius neighborhoods of 15 points with a power parameter of 2.

Table 2
Parameters active in the different models. In the Random Fitted model: sia_μ and sia_σ are parameters for a Normal distribution; pbc_λ and PP_λ are parameters for a Poisson distribution; and U_α and U_β are parameters for a Beta distribution. Entries marked with an asterisk (∗) indicate that the parameter was held at the Base-case value.

Model parameters												
Model	ϕ	μ	λ^r	α_0	α_1	sia^{thresh}	sia_μ	sia_σ	pbc_λ	PP_λ	U_α	U_β
Base-case	4	0.38	0.1	−60.61	2.46	0.6						
Simple environment	5	0.55	0.1	−14.74	0.32	0.5						
Economic only				−32.21	1.31							
Random fitted	∗	∗	∗	∗	∗	∗	0.20	0.30	3	13	5	5

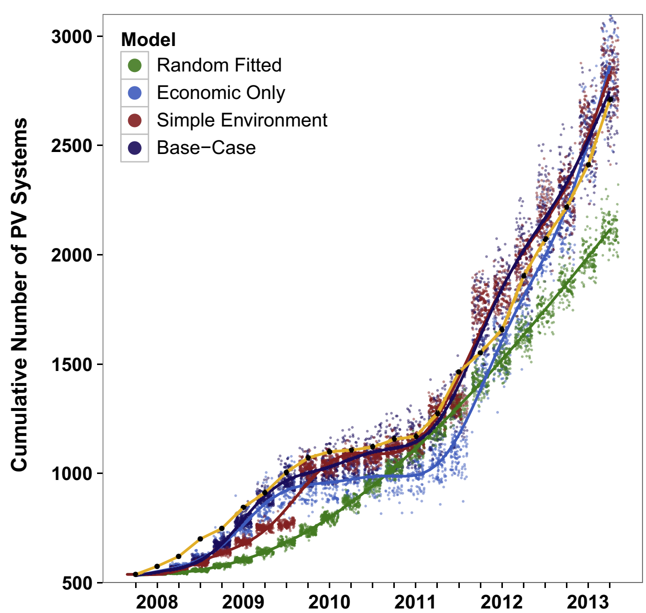

Fig. 2. Base-Case, Simple Environment, Random Fitted, and Economic Only model predictions of cumulative solar installations between Q1 2008 and Q4 2013, after fitting. The yellow line and black points shows the empirical adoption curve used for fitting. Semi-transparent colored points show individual simulation run outcomes (100 runs per model) while solid lines show the expected value of the 100-run batch as a whole for the corresponding model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Table 3), in terms of simple arithmetic error (-37,967.7), fuzzy κ^* (0.81), and r^w (0.81).

3.2. Simple Environment

The Simple Environment model fit the empirical cumulative installation data well (Fig. 2). Cumulative RMSE for this model was 126.6, worse than the Base-case but better than both the Random Fitted and the Economic Only models, discussed below. Temporal validation (Table 3) showed that the Simple Environment model was less successful at predicting the marginal diffusion rate than either of the two models below due to overestimating adoption volatility, particularly around 2009–2010, as shown in Fig. 2.

Spatial similarity statistics were well below the Base-case (Table 3), but very close to those reported by the Economic Only model, although the structure of those errors was much different, as shown in Fig. 3. The Simple Environment model over-predicted a large area in west-central Austin, which is characterized by relatively large, expensive homes built on steep, wooded slopes. The steep slopes and surrounding tree cover drive low-insolation values for the area, and thus make the economics of solar less attractive. But this heterogeneity is left out in the Simple Environment model, leading to the economics-related over-prediction. The importance of available sunlight in driving adoption patterns is further accentuated by the demographic validation. The RMSE in quarterly median adopter home value was the largest of any model, driven by the high home values predicted by the model (Fig. 4). That is because all other factors being equal, the economic component of the model (Section 2.2) favors wealthier households for adoption.

3.3. Random Fitted

Through the iterative process of selecting a distribution and parameters for *sia*, payback (*PP*), *pbc*, and *U*, the Random Fitted model was fit to the empirical data by minimizing the RMSE in the cumulative number of installed systems (Section 2.5). These distributions were parameterized iteratively by adjusting shape and scale parameters of the agent state distributions, while holding the Base-case parameters at their known, constant values (Table 2). This gives the advantage of starting the Random Fitted model with very reasonable parameter values, such as one might glean from the literature or aggregate data. As in the other models, spatial structure and other emergent properties were not used in fitting. In the optimized model, *sia* was initialized to a Normal (0.2, 0.3) distribution, payback was described by a Poisson (13) distribution, *pbc* was initialized to a Poisson (3) distribution, uncertainty was initialized to a Beta (5, 5) distribution, and λ^r was initialized to 1 (Erdős–Rényi random graph). Cumulative RMSE in the number of installations for the random fitted model using the above distributions was found to be 257.49, the largest among the four models considered. The cumulative adoption curve is quite smooth (Fig. 2), as was expected based on the use of smooth parametric distributions used in initialization. This model showed the worst

Table 3
Summary of validation results for the four models. Descriptions of each of the metrics can be found in Section 2.6. Temporal, spatial ("Sp"), and demographic ("Dem") validation metrics are independent of the fitting criteria.

Validation metrics								
Model	RMSE marginal	Sp: simple	Sp: fuzzy κ	Sp: r^w	Dem: RMSE	Dem: simple	Dem: fuzzy κ	Dem: r^w
Base-case	76.93	0.46	0.43	0.86	110,580.2	−37,967.7	0.81	0.81
Simple environment	106.91	0.99	0.42	0.56	162,436.7	109,296.8	0.73	0.74
Random fitted	65.95	−1.09	0.34	0.46	113,289.5	−88,463.62	0.77	0.76
Economic only	88.60	1.37	0.41	0.55	157,047.6	147,748.8	0.74	0.74

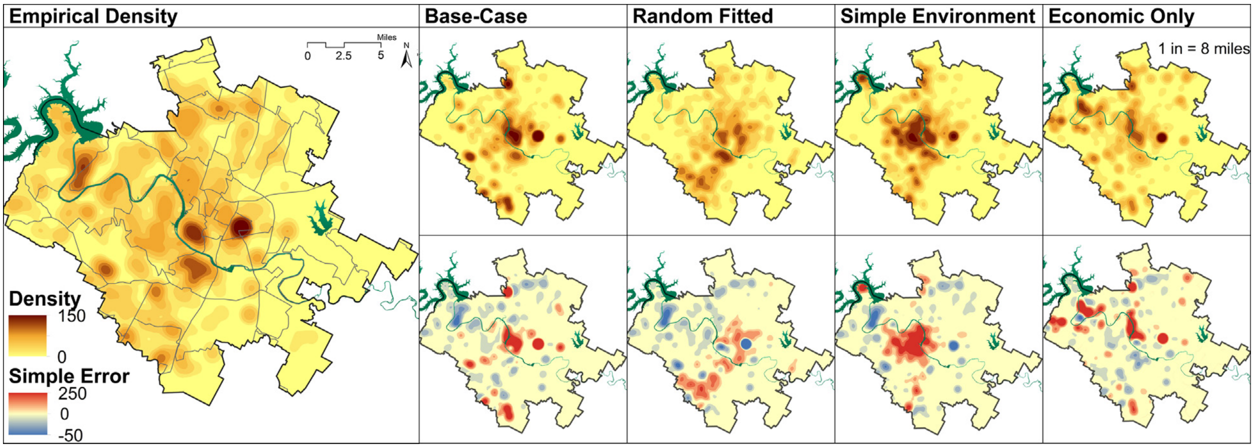

Fig. 3. Spatial validation for the Base-case, Simple Environment, Random Fitted, and Economic Only models. The empirical system density (systems per square mile or 1.61 km^2) is shown in the far left panel. The top (yellow–brown) panels show simulated system density, while the lower (divergent blue–red) panels show simple error (Empirical–Simulated) in system density. Raster resolution is 100 ft (30.5 m). Additional spatial validation metrics, fuzzy numerical κ, and r^w are reported in Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

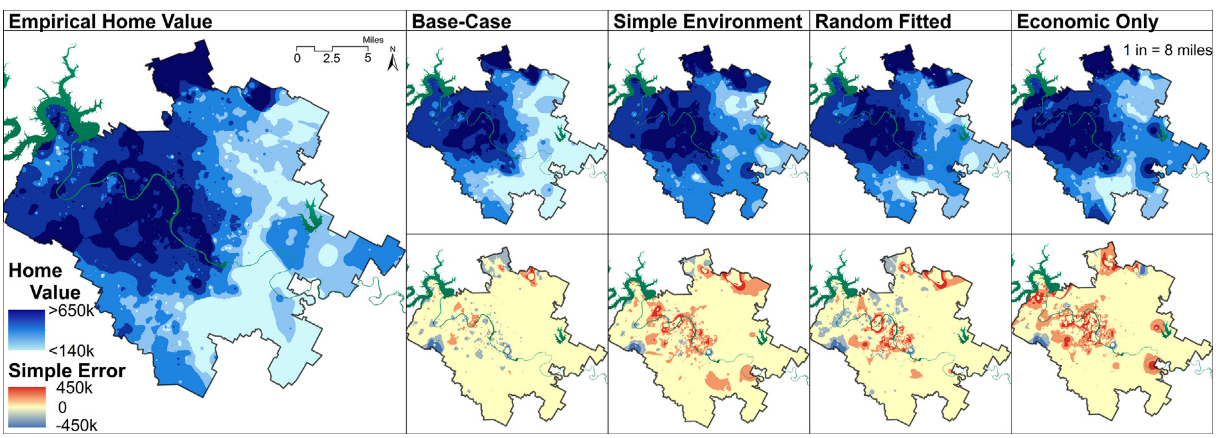

Fig. 4. Spatial demographic validation for the Base-Case, Simple Environment, Random Fitted, and Economic Only models. The large panel on the left shows the empirical adopter home-value distribution over space. The top (blue-scale) panels show simulated adopter home values. Adopter home values were interpolated using the Inverse Distance Weighted technique. The lower (divergent blue–red) panels show simple error (Empirical–Simulated). Raster resolution is 100 ft (30.5 m). Additional spatial validation metrics, RMSE, fuzzy numerical κ, and r^w are reported in Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

fit to the cumulative adoption data, and as shown in Fig. 2, it did not adequately replicate the structural features visible in the empirical adoption curve. However, temporal validation showed that the RMSE in predicting quarterly (i.e., marginal) adoption by this model was more consistent with the empirical data than any other model (see Table 3, marginal RMSE = 65.95).[7]

The spatial validation measures were generally poor for this model, suggesting that empirically deriving the *pbc* variable using insolation, the tree-cover ratio, and home value at the household-level (as is done in the Base-case) in an important determinant of adopter location. Demographic validation, however, was more favorable. The RMSE in quarterly median adopter home value was lower than all other models except for the Base-case. Over space, in this model adopter home values were lower than empirically observed. Better validation on the demographic criterion despite a poor spatial performance suggests that the Random Fitted model replicated the aggregate demographic pattern more accurately for the wrong reasons. In the empirical data, adopter home values are moving downward over time. Because *pbc* is distributed randomly in the Random Fitted model, in the simulation adopter home values are converging on the population mean, which is over $200,000 lower than the adopter mean. Thus, in the Random Fitted model the randomness associated with the initial distributions of agent variables drove down the global and local (neighborhood) average adopter wealth toward to the population averages. This model, thus, reflects convergence to the mean rather than a price-driven shift in adopter demographics. These observations further stress the importance of validating the models using multiple criteria, whenever possible based on the available data.

3.4. Economic Only

Fitting results are shown in Fig. 2. Since this model does not have a social component, it involves only two parameters in the fitting process (Table 2). As such, the Economic Only model is quite

[7] Note that the relatively better performance of the Random Fitted case in temporal validation even relative to the Base-case (discussed above), despite weaker performance in all other validation measures is due to one instance of quarterly over-prediction in installations in the Base-case. This occurs in Q4 2011, due to a sharp decrease in the dollar per Watt installed costs of PV. The price decline results in a surge in installations in the Base-case greater than that observed in the empirical data. With this quarter removed, the Base-case RMSE is 45.79, and the Random Fitted RMSE is 67.29.

lean. RMSE in the cumulative quarterly number of solar installations was calculated as 153.36. The RMSE from marginal quarterly installations was fairly low (Table 3), and the major structural features of the adoption curve were able to be reproduced (Fig. 2). This demonstrates that economic factors such as rebates and prices are central in driving solar technology adoption as well as determining the general shape of the diffusion curve. Regarding spatial validation, the average spatial arithmetic error in solar adopter density over the study area was high relative to the more complex models discussed above. The relatively poor spatial performance was also reflected in the low fuzzy numerical similarity and the wavelet correlation coefficient, which were similar to the Simple Environment model. Demographic validation was also fairly weak, evidenced by the RMSE which was the highest of the four models presented here. While this error is somewhat decreased when focusing on *local* demographic patterns (Dem: Fuzzy κ and Dem: r^w in Table 3), a clear pattern of overestimating adopter home values remains. Thus, while the fit of the Economic Only model was quite good on a temporal basis, the overall validation metrics suggest that other underlying process like social interactions are important in increasing the spatial and demographic fidelity of the model.

4. Discussion

It is clear that the solar adoption process is driven by more than economics alone. However, the fact that the Economic Only model – which uses only two parameters and includes only a financial criterion in agent decision rules – exhibits the three major structural components of the empirical diffusion curve (moderate growth early, a period of very slow growth from Q2 2009 to Q1 2001, and a period of rapid growth 2011–Q2 2013), reveals two important findings. First, it suggests that the variables used in the economic module of the model to estimate a household's ability to afford solar (home value, the tree cover, and insolation) are good predictors of the solar installation decision. Second, a simple economic model is likely adequate if the goal of the study is only to predict adoption levels over time. This finding is somewhat surprising given the significant non-technical barriers associated with solar adoption beyond the system cost. However, to move to models that are predictive of complex disaggregated features such as spatial and demographic patterns, social and attitudinal components need to be integrated into the modeling effort. Using the spatial r^w and the RMSE in median adopter home value validation criteria, lack of incorporation of these components is associated with a 36% lower spatial accuracy and 42% higher demographic error.

Given the importance of the solar ABM's economic component in determining the shape of the diffusion curve, it is not surprising that random initialization in the Random Fitted model fails to reproduce structural features in the time series, even when using fitted distributions. The Random Fitted model also does not reproduce well the empirical spatial distributions, and is out-performed by the Economic Only case, which, while greatly simplified, is grounded in empirical distributions. The Simple Environment model replicates the temporal adoption patterns well, and adequately reproduces spatial structures for the most part. However, it is clear that the model places too much emphasis on the wealth of agents, as opposed to the generation potential of the system. The varied terrain, steep slope, and high tree cover in some parts of the study area make them difficult areas for solar adoption. But that is not captured well by the Simple Environment model, which only uses a city-level aggregate value for insolation. Clearly, the economics of solar depend on the generation potential of the system – however, when studying spatially-resolved adoption patterns, our results emphasize the importance of estimating generation potential at the household level rather than using aggregated (e.g. city or zip code level) metrics.

The use of insolation at the household level in the Base-case model generated significant improvement across all validation metrics. Examination of Fig. 2 shows that the increased match relative to the Simple Environment model comes mostly in the first two years of the simulation. The increased realism from the addition of the heterogeneous system performance results in some agents in high sunlight areas to adopt earlier in the Base-case. Besides, the spatial structure generated in the Base-case is also much improved compared to the Simple Environment due a more accurate (and localized) accounting of solar economics. Using household-level insolation yields a spatial correlation that is 54% higher (increase of 0.30) than that of the Simple Environment model. Further, the error in quarterly median adopter home value is 47% lower, and the spatial structure of adopter demographics is also improved (10% higher fuzzy similarity; 9% higher wavelet correlation). Overall, the Base-case generates emergent phenomena with high fidelity across the board, which is strongly supportive of the model's validity.

4.1. Potential applications in utility program design and infrastructure planning

As noted in Section 1, one of the main strengths of *validated* agent-based models is the flexibility they offer for virtual policy experiments [24,45,46]. In particular, the ABM framework developed in this paper makes it possible to conduct a range of policy simulations to aid incentive program design and infrastructure planning. Below we describe some specific application studies that may be performed using the ABM framework we have developed.

4.1.1. Predictive modeling

Knowing the likely scale and pattern of adoption at a future date (say 10 years later) is important for utilities in overall generation planning and for identifying needs for any upgrades in the distribution infrastructure [10]. As such, the most obvious and quite valuable application of the developed model is for predictive modeling. Key inputs required for such predictions include the trajectory of projected installed prices, electricity prices, and federal and local incentives. It is often useful to build various scenarios regarding future price and incentive trajectories, because there is inherent uncertainty about these variables (for example, Hsu [92] uses a system dynamics model to study the impact of varying levels and types of subsidies on PV deployment). Given these inputs and scenarios, the model can be used to predict the scale and rate of adoption (similar to Fig. 2), the spatial distribution of adoption (similar to Fig. 3), and the demographic patterns of adoption (similar to Fig. 4).

4.1.2. Cost-effectiveness of rebates

Markets for novel, clean consumer energy technologies such as solar PV and plug-in hybrid vehicles (PEVs) are characterized by a number of externalities, including environmental externalities (for example, emissions of greenhouse gases and other local pollutants), peer effects [93], learning by doing, and spillovers [94]. To internalize some of those externalities, jurisdictions at all levels – national, state, local governments – across the world have either already embarked upon or are contemplating subsidy programs to incentivize the adoption clean energy technologies. The cost-effectiveness (i.e., maximum adoption under a given budget constraint) of an incentive program is one of the most important factors that policymakers and program designers consider [95]. Our ABM framework is well suited for such cost-effectiveness studies. For example, a number of recent empirical findings on optimal subsidy design find that larger rebates early on in the subsidy program and declining over time may be more cost-effective, especially when peer effects and learning-by-doing effects are strong [95,96]. The spatio-temporal implications of such aggregate findings on subsidy program design could be assessed in full-scale ABM simulations.

4.1.3. Locationally targeted rebates

Congestion in the transmission and distribution (T&D) system due to load-pockets and other factors [56] makes the electricity generated by solar PV have a different value for electric utilities based on the location and time of generation [97]. The potential to offer different rebates in different locations based on *value* potentially increases the ability of utilities for better grid management and could allow for greater penetration levels, while enhancing grid operations and reliability. Similarly, better rebate targeting could enable targeted adoption and increased generation capacity in certain locations on the distribution grid, allowing the utility to defer costly infrastructure upgrades [98]. Using the developed framework a range of *targeted rebate* scenarios could be explored through ABM simulation experiments to study the temporal and locational changes in adoption resulting from changes in the rebate. Knowing this would allow utility planners to set the rebates at efficient levels, by matching marginal costs (the rebate) with the marginal benefits (value provided by a new unit of the solar PV deployed, for example, due to locational benefits).

4.1.4. Targeted information dissemination

A large body of the empirical and practice-oriented literatures identify information gaps as a major barrier to the wider adoption of new energy technologies [14,21,15]. In a related vein, contagion models have been successfully used to study the spread of information and epidemics [99,100]. A key finding from that literature is that high centrality nodes are most effective in spreading information [101]. As regards new energy technology adoption, while centrality may be strongly associated with the ability to spread information, the impact of specific nodes (households) will be moderated by how quickly they are able to spread information out to other *economically capable* nodes. The key point here is that just the spread of information is not sufficient to "infect" others to adopt a technology; financial capability to adopt is equally, if not more important. Thus it is likely that the *combined effect* of information dissemination campaigns and subsidies likely has a greater effect than either enacted in isolation [15,102]. The ABM framework developed here could be used to study both the temporal and scale (steady-state reach) effects of various information seeding strategies that try to rapidly disseminate information by "seeding" the well-connected agents in the network.

5. Conclusion

Agent-based modeling offers a unique ability to model the decision-making of heterogeneous agents with bounded rationality. However, this requires careful model design, implementation, and validation. The rigor of the modeling process will determine the robustness and applicability of any insights gleaned from ABM. Using a detailed, theoretically- and empirically-grounded agent-based model for residential solar PV adoption, in this paper we have shown how the inclusion of social and attitudinal components in the model, in addition to economic factors alone, impact model fit and validity. Solar PV adoption is a complex process, including agent-level economic, attitudinal, and social factors. Models that include all of these components fit the data better and generate more realistic emergent patterns. Most importantly, we find that models that focus solely on the financial aspects of agents' decision-making do well in generating the rate of adoption and the cumulative adoption curve, but do not perform well on spatial and demographic patterns of adoption. The incorporation of agents' attitudinal aspects and social interactions become critical when seeking high spatial and demographic accuracy of the model.

Furthermore, we find it necessary to use multiple validation checks in assessing the performance of the different models considered. This is especially important when modeling systems with multi-layered environmental and social interactions that impact agent behavior, as in the case of solar PV adoption. In general, more granular data and less reliance on random distributions for describing the system increases model fit and validity. The important point here is that the validity of a model critically depends upon the underlying data and the level of abstraction: the same model may perform well for highly-aggregated outcomes (i.e., scale of adoption), while doing poorly in describing the structural aspects (e.g., spatial, demographic, etc.) of the system at lower levels of abstraction.

Acknowledgements

We thank Jay Zarnikau and Eugenio Arima for very helpful suggestions on earlier versions of this work. We thank the anonymous reviewers for their constructive feedback, which helped improve the manuscript. We acknowledge funding support from the U.S. Department of Energy under its Solar Energy Evolution and Diffusion Studies (SEEDS) program within the SunShot Initiative (Award Number DE-EE0006129). Data assistance from Austin Energy and the City of Austin is gratefully acknowledged. We also gratefully acknowledge computing resources made available by the Texas Advanced Computing Center (TACC). VR acknowledges support from the Elspeth Rostow Memorial Fellowship and from the Policy Research Institute (PRI) at the LBJ School of Public Affairs (UT Austin). All remaining errors are ours alone.

References

[1] de Vries P, Aarts H, Midden CJH. Changing simple energy-related consumer behaviors: how the enactment of intentions is thwarted by acting and non-acting habits. Environ Behav 2011;43(5):612–33.
[2] Dietz T, Stern PC, Weber EU. Reducing carbon-based energy consumption through changes in household behavior. Daedalus 2013;142(1):78–89.
[3] Faber JF. Behavioural climate change mitigation options and their appropriate inclusion in quantitative longer term policy scenarios. CE Delft; 2012.
[4] Pisello AL, Asdrubali F. Human-based energy retrofits in residential buildings: a cost-effective alternative to traditional physical strategies. Appl Energy 2014;133(0):224–35.
[5] Shavel I, et al. Exploring natural gas and renewables in ERCOT. Part iii: The role of demand response, energy efficiency, and combined heat & power. Technical report, The Brattle Group; 2014.
[6] Zhao J, Kucuksari S, Mazhari E, Son Y. Integrated analysis of high-penetration PV and PHEV with energy storage and demand response. Appl Energy 2013;112(0):35–51.
[7] Moshari A, Yousefi GR, Ebrahimi A, Haghbin S. Demand-side behavior in the smart grid environment. In: Innovative smart grid technologies conference Europe (ISGT Europe), 2010 IEEE PES; October 2010. p. 1–7.
[8] Davito B, Tai H, Uhlaner R. The smart grid and the promise of demand-side management. McKinsey on Smart Grid; 2010. p. 38–44.
[9] Zheng M, Meinrenken CJ, Lackner KS. Agent-based model for electricity consumption and storage to evaluate economic viability of tariff arbitrage for residential sector demand response. Appl Energy 2014;126(0):297–306.
[10] Ipakchi A, Albuyeh F. Grid of the future. Power Energy Mag IEEE 2009;7(2):52–62.
[11] Wilkerson JT, Cullenward D, Davidian D, Weyant JP. End use technology choice in the National Energy Modeling System (NEMS): an analysis of the residential and commercial building sectors. Energy Econ 2013;40(C):773–84.
[12] MJ Bradley & Associates. Electric vehicle grid integration in the US, Europe, and China: challenges and choices for electricity and transportation policy. Regulatory Assistance Project; 2013.
[13] Kemp R, Volpi M. The diffusion of clean technologies: a review with suggestions for future diffusion analysis. J Cleaner Prod 2008;16(1, Suppl. 1):14–21.

[14] Wilson C, Dowlatabadi H. Models of decision making and residential energy use. Ann Rev Environ Resour 2007;32:169–203.
[15] Stern PC. What psychology knows about energy conservation. Am Psychol 1992;47(10):1224.
[16] Conlisk J. Why bounded rationality? J Econ Lit 1996;34(2):669–700.
[17] Gigerenzer G, Selten R. Bounded rationality: the adaptive toolbox. MIT Press; 2002.
[18] Rubinstein A. Modeling bounded rationality, vol. 1. MIT press; 1998.
[19] Rai V, Robinson SA. Effective information channels for reducing costs of environmentally-friendly technologies: evidence from residential PV markets. Environ Res Lett 2013;8(1). 014044(1-8).
[20] Faiers A, Neame C. Consumer attitudes towards domestic solar power systems. Energy Policy 2006;34(14):1797–806.
[21] Margolis R, Zuboy J. Nontechnical barriers to solar energy use: review of recent literature. National Renewable Energy Laboratory, US Department of Energy; 2006.
[22] M Krause R, Carley SR, Lane BW, Graham JD. Perception and reality: public knowledge of plug-in electric vehicles in 21 US cities. Energy Policy 2013;63:433–40.
[23] Epstein JM. Generative social science: studies in agent-based computational modeling. Princeton University Press; 2006.
[24] Macal C, North M. Tutorial on agent-based modelling and simulation. J Simul 2010;4(3):151–62.
[25] Crooks AT, Castle CJE. Agent-based models of geographical systems. The integration of agent-based modelling and geographical information for geospatial simulation. Springer; 2012. chapter 12, pp. 219–252.
[26] Gatti D, Gaffeo E, Gallegati M. Complex agent-based macroeconomics: a manifesto for a new paradigm. J Econ Interact Coord 2010;5(2):111–35.
[27] Manson SM. Bounded rationality in agent-based models: experiments with evolutionary programs. Int J Geogr Inf Sci 2006;20(9):991–1012.
[28] J North M, Collier NT, Ozik J, Tatara ER, Macal CM, Bragen M, et al. Complex adaptive systems modeling with repast simphony. Complex Adapt Syst Model 2013;1(1):1–26.
[29] Cantono S, Silverberg G. A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies. Technol Forecast Soc Change 2009;76(4):487–96.
[30] Gunther M, Stummer C, Wakolbinger L, Wildpaner M. An agent-based simulation approach for the new product diffusion of a novel biomass fuel. J Oper Res Soc 2011;62:12–20.
[31] Kiesling E, Gunther M, Stummer C, Wakolbinger L. Agent-based simulation of innovation diffusion: a review. Central Euro J Oper Res 2012;20(2): 183–230.
[32] Laciana CE, Rovere SL. Ising-like agent-based technology diffusion model: adoption patterns vs seeding strategies. Phys A: Stat Mech Appl 2011;390(6):1139–49.
[33] Lee T, Yao R, Coker P. An analysis of UK policies for domestic energy reduction using an agent based tool. Energy Policy 2014;66(0):267–79.
[34] Mazhari E, Zhao J, Celik N, Lee S, Son YJ, Head L. Hybrid simulation and optimization-based design and operation of integrated photovoltaic generation, storage units, and grid. Simul Model Practice Theory 2011;19(1):463–81.
[35] Robinson S, Stringer M, Rai V, Tondon A. GIS-integrated agent-based modeling of residential solar pv diffusion. In: 32nd USAEE/IAEE North American conference proceedings; July 2013.
[36] Schwarz N, Ernst A. Agent-based modeling of the diffusion of environmental innovations: an empirical approach. Technol Forecast Soc Change 2009;76(4):497–511.
[37] Schwoon M. Simulating the adoption of fuel cell vehicles. J Evol Econ 2006;16(4):435–72.
[38] Sopha BM, Klöckner CA, Hertwich EG. Exploring policy options for a transition to sustainable heating system diffusion using an agent-based simulation. Energy Policy 2011;39(5):2722–9.
[39] Sopha BM, Klockner CA, Hertwich EG. Adoption and diffusion of heating systems in Norway: coupling agent-based modeling with empirical research. Environ Innov Soc Trans 2013;8(0):42–61.
[40] Tran M. Agent-behaviour and network influence on energy innovation diffusion. Commun. Nonlinear Sci. Numer. Simul. 2012;17(9):3682–95.
[41] Van Vliet O, Vries B, Faaij A, Turkenburg W, Jager W. Multi-agent simulation of adoption of alternative fuels. Transport Res Part D: Transport Environ 2010;15(6):326–42.
[42] Zhang T, Gensler S, Garcia R. A study of the diffusion of alternative fuel vehicles: an agent-based modeling approach. J Product Innov Manage 2011;28(2):152–68.
[43] Zhang T, Nuttal WJ. An agent-based simulation of smart metering technology adoption. Int J Agent Technol Syst 2012;4:17–38.
[44] McCoy D, Lyons S. Consumer preferences and the influence of networks in electric vehicle diffusion: an agent-based microsimulation in Ireland. Energy Res Soc Sci 2014;3:89–101.
[45] Crooks AT, Heppenstall AJ. Agent-based models of geographical systems. Introduction to agent-based modelling. Springer; 2012. chapter 5, pp. 85–105.
[46] Alam SJ, Geller A. Networks in agent-based social simulation. In: Heppenstall AJ, Crooks AT, See LM, Batty M, editors. Agent-based models of geographical systems. Springer; 2012. p. 199–216.
[47] Durlauf SN. Complexity, economics, and public policy. Politics Philos Econ 2012;11(1):45–75.
[48] Windrum P, Fagiolo G, Moneta A. Empirical validation of agent-based models: alternatives and prospects. J Artif Socit Soc Simul 2007;10(2):8.
[49] Grimm V, Revilla E, Berger U, Jeltsch F, Mooj WM, Railsback SF, et al. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 2005;310:987–91.
[50] Werker C, Brenner T. Empirical calibration of simulation models. Technical report. Papers on Economics and Evolution; 2004.
[51] Heppenstall A, Evans A, Birkin M. Using hybrid agent-based systems to model spatially-influenced retail markets. J Artif Societ Soc Simul 2006;9(3).
[52] North MJ, Macal CM. Managing business complexity: discovering strategic solutions with agent-based modeling and simulation. Oxford University Press; 2007.
[53] Rand W, Rust RT. Agent-based modeling in marketing: guidelines for rigor. Int J Res Market 2011;28(3):181–93.
[54] Smajgl A, Barreteau O. Empirical agent-based modeling: challenges and solutions. Springer; 2014.
[55] Barbose G, Darghouth NR, Weaver S, Feldman D, Margolis R, Wiser R. Tracking US photovoltaic system prices 1998–2012: a rapidly changing market. Progress in photovoltaics: research and applications; 2014.
[56] Blackburn G, Magee C, Rai V. Solar valuation and the modern utility's expansion into distributed generation. Electricity J 2014;27(1):18–32.
[57] Rai V, Sigrin B. Diffusion of environmentally-friendly technologies: buy vs lease decisions in residential PV markets. Environ Res Lett 2013;8(1). 014022(1–8).
[58] Drury E, Miller M, Macal CM, Graziano DJ, Heimiller D, Ozik J, et al. The transformation of southern california's residential photovoltaics market through third-party ownership. Energy Policy 2012;42:681–90.
[59] Rai V, Robinson SA. Agent-based modeling of energy technology adoption: empirical integration of social, behavioral, economic, and environmental factors, in preparation. <http://bit.ly/1rVSuB8>.
[60] Rai V, McAndrews K. Decision-making and behavior change in residential adopters of solar PV. In: Proceedings of the world renewable energy forum. Denver (CO); May 2012.
[61] Armitage CJ, Conner M. Efficacy of the theory of planned behaviour: a meta-analytic review. Brit J Soc Psychol 2001;40(4):471–99.
[62] Glasman LR, Albarracan D. Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. Psychol Bull 2006;132(5): 778.
[63] Ajzen I. The theory of planned behavior. Organ Behav Hum Dec Process 1991;50(2):179–211.
[64] Madden TJ, Ellen PS, Ajzen I. A comparison of the theory of planned behavior and the theory of reasoned action. Pers Soc Psychol Bull 1992;18(1):3–9.
[65] Zhang T. Agent-based simulation of consumer purchase decision-making and the decoy effect. J Bus Res 2007;60(8):912–22.
[66] Kniveton D, Smith C, Wood S. Agent-based model simulations of future changes in migration flows for Burkina Faso. Global Environ Change 2011;21(S1):S34–40.
[67] Richetin J, Sengupta A, Perugini M, Adjali I, Hurling R, Greetham D, et al. A micro-level simulation for the prediction of intention and behavior. Cognit Syst Res 2010;11(2):181–93.
[68] Kaufmann P, Stagl S, Franks DW. Simulating the diffusion of organic farming practices in two new EU member states. Ecol Econ 2009;68(10): 2580–93.
[69] Deffuant G, Neau D, Amblard F, Weisbuch G. Mixing beliefs among interacting agents. Adv Complex Syst 2000;3(1):87–98.
[70] Hegselmann R, Krause U. Opinion dynamics and bounded confidence models, analysis, and simulation. J Artif Societ Soc Simul 2002;5(3).
[71] Meadows M, Cliff D. Reexamining the relative agreement model of opinion dynamics. J Artif Societ Soc Simul 2012;15(4).
[72] Brown DG, Page S, Riolo R, Zellner M, Rand W. Path dependence and the validation of agent-based spatial models of land use. Int J Geogr Inf Sci 2005;19(2):153–74.
[73] Maio GR, Bell DW, Esses VM. Ambivalence and persuasion: the processing of messages about immigrant groups. J Exp Soc Psychol 1996;32(6):513–36.
[74] Clark JK, Wegener DT, Fabrigar LR. Attitudinal ambivalence and message-based persuasion: motivated processing of proattitudinal information and avoidance of counterattitudinal information. Pers Soc Psychol Bull 2008;34(4):565–77.
[75] Fu P, Rich PM. A geometric solar radiation model with applications in agriculture and forestry. Comput Electron Agric 2002;37(1):25–35.
[76] Pons X, Ninyerola M. Mapping a topographic global solar radiation model implemented in a GIS and refined with ground data. Int J Climatol 2008;28(13):1821–34.
[77] Fu P, Rich PM. Design and implementation of the solar analyst: an arcview extension for modeling solar radiation at landscape scales. In: Proceedings of the 19th annual ESRI user conference. San Diego (USA); 1999.
[78] Marion B, Anderberg M, George R, Gray-Hann P, Heimiller D. PVWATTS version 2–enhanced spatial resolution for calculating grid-connected pv performance. In: Proceedings of the 2001 NCPV program review meeting. Lakewood (CO); 2001. p. 143–4.
[79] Watts D, Strogatz S. Collective dynamics of small-world networks. Nature 1998;393(6684):440–2.
[80] Deffuant G, Amblard F, Weisbuch G, Faure T. How can extremism prevail? A study based on the relative agreement interaction model. J Artif Societ Soc Simul 2002;5(4).

[81] Deffuant G, Weisbuch G. Probability distribution dynamics explaining agent model convergence to extremism. Soc Simul: Technol Adv New Disc 2008:43–60.
[82] Franks DW, Noble J, Kaufmann P, Stagl S. Extremism propagation in social networks with hubs. Adapt Behav 2008;16(4):264–74.
[83] Deffuant G, Huet S, Amblard F. An individual based model of innovation diffusion: mixing social value and individual benefit. Am J Sociol 2005;110(4):1041–69.
[84] Deffuant G, Huet S, Bousset JP, Henriot J, Amon G, Weisbuch G, et al. Agent-based simulation of organic farming conversion in Allier département. Complex Ecosyst Manage: Theory Practice Multi-agent Syst 2002:158–87.
[85] Hagen A. Fuzzy set approach to assessing similarity of categorical maps. Int J Geogr Inf Sci 2003;17(3):235–49.
[86] Hagen A, Straatman B, Uljee I. Further developments of a fuzzy set map comparison approach. Int J Geogr Inf Sci 2005;19(7):769–85.
[87] Briggs WM, Levine RA. Wavelets and field forecast verification. Mon Weather Rev 1997;125(6):1329.
[88] Casati B, Ross G, Stephenson DB. A new intensity-scale approach for the verification of spatial precipitation forecasts. Meteorol Appl 2004;11(2):141–54.
[89] Macintosh A, Wilkinson D. Searching for public benefits in solar subsidies: a case study on the Australian government's residential photovoltaic rebate program. Energy Policy 2011;39(6):3199–209.
[90] Nelson T, Simshauser P, Kelley S. Australian residential solar feed-in tariffs: industry stimulus or regressive form of taxation? Econ Anal Policy 2011;41(2):113–29.
[91] Shepard D. A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM national conference. ACM; 1968. p. 517–24.
[92] Hsu Chiung-Wen. Using a system dynamics model to assess the effects of capital subsidies and feed-in tariffs on solar PV installations. Appl Energy 2012;100:205–17.
[93] Graziano M, Gillingham K. Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. J Econ Geogr 2014. http://dx.doi.org/10.1093/jeg/lbu036.
[94] Popp D, Newell RG, Jaffe AB. Energy, the environment, and technological change. Handbook Econ Innov 2010;2:873–937.
[95] Dong C. Technology diffusion policy design: cost-effectiveness and redistribution in California solar subsidy programs. PhD thesis. The University of Texas at Austin; 2014.
[96] Van Benthem A, Gillingham K, Sweeney J. Learning-by-doing and the optimal solar policy in California. Energy J 2008:131–51.
[97] Borenstein S. The market value and cost of solar photovoltaic electricity production. Center for the Study of Energy Markets; 2008.
[98] Hoff Thomas E, Wenger Howard J, Farmer Brian K. Distributed generation: an alternative to electric utility investments in system capacity. Energy Policy 1996;24(2):137–47.
[99] Young HP. Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. Am Econ Rev 2009;99(5):1899–924.
[100] Kuperman M, Abramson G. Small world effect in an epidemiological model. Phys Rev Lett 2001;86:2909–12.
[101] Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. Science 2009;323(5916):892–5.
[102] Ashraf N, Jack BK, Kamenica E. Information and subsidies: complements or substitutes? J Econ Behav Org 2013;88(0):133–9. Asian Institutional Economics.