Incentives to Manipulate Demand Response Baselines With Uncertain Event Schedules

Douglas Ellman, *Graduate Student Member, IEEE*, and Yuanzhang Xiao, *Member, IEEE*

***Abstract*—We study baseline-based demand response (DR) programs. In such programs, customers get rebates based on how much they reduce electricity consumption during DR events relative to a "baseline," where this baseline is determined by their consumption during previous non-event days. Customers, or automated controls working on their behalf, can achieve higher DR payments by decreasing consumption during DR events (desired behavior), and by increasing consumption during non-event times (baseline manipulation). Importantly, the customers have imperfect knowledge of when future demand response events will occur. To understand customers' incentives for baseline manipulation, we present a novel multi-stage stochastic dynamic programming model that optimizes customer actions for maximum expected rewards under uncertain event schedules. Analytical results for special cases show fundamental drivers of customer incentives. Simulation results reveal incentives to manipulate baselines and impacts to program performance for a realistic baseline-based demand response program, and how program and customer parameters affect incentives.**

***Index Terms*—Dynamic programming, stochastic optimal control, demand response, demand response baselines, baseline-based demand response, incentive-based demand response, smart grid.**

I. Introduction

DEMAND response (DR) aims to modify customers' electricity consumption in order to improve operation of the electric grid (e.g., reduce costs, improve reliability, reduce emissions). A wide variety of DR mechanisms have been proposed or implemented [1], [2]. Most of these mechanisms fall into three categories: price-based DR, auction-based DR, and incentive-based DR. In price-based DR (e.g., real-time pricing [3], [4]), the price of electricity varies over time, which encourages customers to reduce consumption when prices are high and increase consumption when prices are low. In auction-based DR, customers submit bids (e.g., supply functions [5], utility functions [6], [7]) to compete for payment for DR services. In incentive-based DR (e.g., peak time rebate [8], [9]), customers are paid incentives based on participation or performance in the DR program.

Manuscript received June 12, 2019; revised November 8, 2019 and August 16, 2020; accepted August 31, 2020. Date of publication September 15, 2020; date of current version February 26, 2021. This work was supported by NSF under Grant 1822213. Paper no. TSG-00825-2019. *(Corresponding author: Douglas Ellman.)*

The authors are with the Hawai'i Advanced Wireless Technologies Institute and Department of Electrical Engineering, University of Hawai'i at Mānoa, Honolulu, HI 96822 USA (e-mail: dellman@hawaii.edu; yxiao8@hawaii.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2020.3024208

Day:	1	2	3	4	5	6	7	8	9
Event Status:	NE	NE	E	NE	NE	E	NE	NE	E

Fig. 1. Example of an event schedule for baseline-based DR with a 5-day baseline. E indicates an event day and NE indicates a non-event day. For event day 9 (in orange) the 5 prior non-event days (in blue) are used to determine the baseline load.

An important class of incentive-based DR is baseline-based DR [7]–[13], where incentive payments are based on the difference between baseline consumption (e.g., average consumption during previous days) and event consumption for each DR event. A *DR event* is a specific window of time in which the DR program asks participating customers to reduce energy consumption. *Baseline consumption*, or *baseline* in short, is an estimate of what the customer's energy consumption would have been during that same window of time if it was not a DR event, and is calculated based on the customer's energy consumption on prior days without DR events. Baseline-based DR is used in many utility DR programs, such as [14]. Baselines are also central to DR participation in U.S. wholesale energy markets, as regulated by the Federal Energy Regulatory Commission [15]. Given its prevalence in practice, baseline-based DR will be the focus of our paper.

With the growing use of baseline-based DR, it is important to understand incentives and behavior of participating customers, since actions of customers and control systems acting on their behalf are the primary determinant of overall DR program performance and cost. Baseline-based DR can create complicated customer incentives, including incentives to take action both during DR events (desired behavior) and to manipulate baselines during non-event times. In general, customer incentives can be difficult to predict, because optimal decisions for each time period are interrelated and can depend on the customer's current baseline consumption, customer costs, the demand response program structure, and uncertain future values (such as when DR events will occur).

It is not new that baseline-based DR is susceptible to baseline manipulation. Randomized control trials [16] and legal settlements [17] provide real-world evidence that customers may manipulate baselines. Interested readers can also see [18] for a high-level discussion of the potential for inflating baselines in baseline-based DR. In cases with incentives for baseline inflation, ignoring these behaviors can lead to significant errors in projections of required DR quantities, DR costs, and demand forecasts. As the use of DR for critical grid services grows and as customer responses to DR programs become increasingly sophisticated through automated control systems, these errors could cause significant impacts to the cost and reliability of the grid.

We build on the prior work in the area of baseline-based DR and develop a novel model to identify optimal customer decisions under a baseline-based DR program with an uncertain event schedule (see Section II for a discussion of related work). Our model solves the customer's multi-stage stochastic decision problem with a dynamic programming algorithm, thus providing an optimal policy for how much to reduce or increase electricity consumption based on time, current customer baseline, and whether a DR event is active. The key contribution of our approach is representation of the customer's imperfect knowledge of when DR events will occur over the course of a DR season, which can have significant impacts on the optimal policy and is a primary uncertainty for DR participants. As a demonstration of the value of our approach, we will show non-obvious results based on a real-world DR program (in Section V), including examples of how imposing penalties (i.e., negative payments) for increasing load during DR events affects DR program performance. Our major contributions are as follows:

- To the best of our knowledge, we present the first solutions for optimal customer policies under a variety of baseline-based DR structures with *multi-day* baselines and *multiple uncertain* events over a DR season.
- For a special class of baseline-based DR models, we provide structural results on the customer's optimal policy, which sheds light on the design of the DR program.
- For a general baseline-based DR model, we develop an efficient dynamic programming algorithm by exploiting the structure of the problem, which is capable of providing solutions quickly for extended DR seasons of several months or longer.
- We provide a detailed rigorous analysis of the computational complexity of our algorithm, and conduct empirical analysis of the robustness of our model in the presence of imperfect knowledge about the DR event probabilities.
- We provide sensitivity analysis on baseline determination rules to illustrate how to use our model to inform the design of a baseline-based DR program.

The remainder of this article is organized as follows. Section II discusses related works. Section III describes the model. Section IV provides analytical results for special cases. Section V presents numerical results from a computer implementation of the general model. Section VI summarizes our conclusions.

II. Related Work

There is a vast literature studying DR programs. Here we only discuss works on baseline-based DR programs [7]–[12].

Some works [9], [10], [19] study issues other than baseline manipulation. For example, the work in [9] focuses on learning the optimal payment rule in baseline-based DR programs, the work in [10] compares different baseline structures in terms of metrics such as the accuracy of the baseline, and the work in [19] studies competition between generators in an electricity market with baseline-based demand response. However, the focus of our paper is the study of incentives to manipulate baselines.

Most of the works that address baseline manipulation focus on mechanism design of the DR program [7], [11]–[13]. More specifically, [7], [11], [13] allow customers to self-report baselines, and encourage truthful reporting by excluding a subset of customers from each event, and imposing penalties for deviations from reported baselines for those non-participating customers. In [12] there is a profit-sharing mechanism designed to discourage manipulating baselines for a DR program in which the baseline is the load at the start of the DR event. In contrast, we study a class of existing DR programs where the baseline is defined as a customer's consumption during previous non-event days. The structure of the DR program we study results in a unique feature of the customers' decision making process: they need to determine their consumption in *both* event and non-event days, facing *uncertainty* of when the events will occur. On the contrary, DR payments do not depend on non-event day consumption for the mechanisms in [7], [11]–[13].

The most related work is [8], which studies DR programs with baselines determined by the consumption during non-event days. However, they focus on a two-stage model, where the second stage is the event day with certainty. Note that this work includes a formulation of the general multi-stage model, but it only presents results for the two-stage model without event day uncertainty.

Compared to our conference paper [20], we have added a detailed rigorous analysis of the computational complexity of our algorithm in Section V-A, have added Section V-D to demonstrate the robustness of our results when the estimates of DR event probabilities are imperfect, and have added Section V-E to demonstrate how to use our model to inform the design of a baseline-based DR program.

III. Model

A. System Setup

We study a demand response program that occurs over a demand response season of $T \in \mathbb{N}$ days. At the start of each day $t \in \{1, \dots, T\}$, the utility notifies the customer whether the current day is a demand response *event day*. On event days, the utility desires participating customers to reduce electricity consumption during a pre-specified window of hours (e.g., 2 p.m – 6 p.m. on all event days). We denote the probability that day t is an event day by $p_{E,t} \in [0, 1]$. An event day occurs independently from other days.

In each day t, the customer has a default load level l_t and may deviate from the default load level by a_t, resulting in an actual load level of $a_t + l_t$. The default load level l_t is a model parameter, and the deviation a_t is the customer's decision variable. We model the deviation, instead of the actual load, as the decision variable in order to better represent the customer's decision of whether to increase or decrease the load. We denote the set of available load deviations by $A_t \subset \mathbb{R}$, a compact set including positive and negative values. Therefore, the action set can be either discrete, or continuous and closed and bounded.

On an event day t, the utility gives the customer a rebate for load reduction. We define the rebate function as

$$\begin{aligned} r_t : \mathbb{R} \times A_t &\to \mathbb{R} \\ (\bar{s}_{B,t}, a_t) &\mapsto r_t(\bar{s}_{B,t}, a_t), \end{aligned} \tag{1}$$

which depends on the *baseline* $\bar{s}_{B,t}$ and the actual load $l_t + a_t$. The baseline is determined by the utility according to the actual load over the previous N_B non-event days. We denote the load over the N_B non-event days prior to day t by a vector $\mathbf{s}_{B,t} \in \mathbb{R}^{N_B}$, and write the rule of determining the baseline as a function

$$f : \mathbf{s}_{B,t} \mapsto \bar{s}_{B,t}. \tag{2}$$

We focus on a commonly used class of baseline functions that calculate the average load of the preceding N_B non-event days or the average load over $N'_B < N_B$ days selected based on some rule [10]. We consider two common forms of DR rebates calculated by

$$r_t(\bar{s}_{B,t}, a_t) = r_{DR,t} \cdot (\bar{s}_{B,t} - (l_t + a_t)) \text{ or} \tag{3}$$

$$r_t(\bar{s}_{B,t}, a_t) = r_{DR,t} \cdot (\bar{s}_{B,t} - (l_t + a_t))^+, \tag{4}$$

where $r_{DR,t} \in \mathbb{R}_+$ is the rebate per unit of load reduction and $(\cdot)^+ = \max\{\cdot, 0\}$. When event load is less than baseline load, these rebates are proportional to the difference between baseline load and event load. In (3), the rebate becomes negative when event load is greater than baseline load, representing a penalty to the customer for failing to reduce load. Alternatively, in (4), the rebate is capped above zero, so there is no penalty when event load is greater than baseline load. These forms of rebates are used, for example, in [14], which caps DR performance payments above zero for each aggregation of participants, but allows individual customers within an aggregation to make negative contributions. These rebates are non-increasing in event load and non-decreasing in baseline load, so customers can increase rebates by reducing event load or increasing baseline load.

The customer incurs a cost when deviating from the default load. We define the cost of changing the load on day t as

$$c_t : A_t \to \mathbb{R}. \tag{5}$$

We assume that the cost function is bounded when the action set is discrete, and that the cost function is convex when the action set is continuous. This cost function represents the net financial and discomfort costs to the customer due to deviating from the default load. For example, raising the temperature setting of an air conditioner in the summer would lower energy use and impose a discomfort cost, while lowering the temperature setting would increase energy use and impose the financial cost of buying more energy. A home battery system could be used to charge or discharge to raise or lower energy consumption, imposing a financial cost for extra energy purchases due to efficiency losses in the battery. In general, the cost function will depend on factors including the available customer resources for modifying load, customer preferences, and electricity costs.

The goal of the customer is to maximize its expected total profit (i.e., rebate minus cost) during the demand response season. Therefore, the customer has incentives to reduce the load during the event days if the rebate is large enough to offset the cost of reducing load. Moreover, the customer may also increase the load during non-event days in order to inflate the baseline and thus increase the future rebate, although increasing the load would only result in a higher cost in that non-event day. We are interested in whether and to what extent such baseline manipulation occurs.

B. Dynamic Programming Formulation

We formulate the customer's decision making problem as a stochastic dynamic program over a finite horizon of T days. Next, we describe the key components of the dynamic program in detail.

The state at each day t consists of the baseline state $\mathbf{s}_{B,t} \in \mathbb{R}^{N_B}$ and the event state $s_{E,t} \in \{0, 1\}$. As discussed before, the baseline state $\mathbf{s}_{B,t}$ is a vector of the load levels on the previous N_B non-event days prior to day t. The event state $s_{E,t}$ indicates whether day t is an event day, with $s_{E,t} = 1$ indicating that day t is an event day. We write the complete state for day t as $\mathbf{s}_t = (\mathbf{s}_{B,t}, s_{E,t})$. The action at each day t is $a_t \in A_t$, the deviation from the default load.

The state $\mathbf{s}_t$ and the action a_t determine the next baseline state $\mathbf{s}_{B,t+1}$.

If day t is an event day, the action a_t will not be counted in determining the future baselines. In this case, the baseline state $\mathbf{s}_{B,t+1} = \mathbf{s}_{B,t}$ stays the same, and the event state $s_{E,t+1}$ follows Bernoulli distribution with parameter $p_{E,t+1}$.

If day t is a non-event day, the action a_t will be counted towards determining future baselines. In this case, the baseline state $\mathbf{s}_{B,t+1}$ is a concatenation of the last $N_B - 1$ elements of the previous baseline state $\mathbf{s}_{B,t}$ and the actual load $l_t + a_t$. Mathematically, we have

$$\mathbf{s}_{B,t+1} = \left([\mathbf{s}_{B,t}]_{2:N_B}, l_t + a_t \right), \tag{6}$$

where $[\mathbf{s}_{B,t}]_{2:N_B}$ is the vector of last $N_B - 1$ elements of $\mathbf{s}_{B,t}$. Similar as in the event day, the event state $s_{E,t+1}$ is drawn from the Bernoulli distribution with parameter $p_{E,t+1}$.

In summary, the state transition probability $P(\mathbf{s}_{t+1}|\mathbf{s}_t, a_t)$ can be calculated as

$$\begin{aligned} P(\mathbf{s}_{t+1}|(\mathbf{s}_{B,t}, 1), a_t) = {} & p_{E,t+1}^{s_{E,t+1}} \cdot (1 - p_{E,t+1})^{1-s_{E,t+1}} \\ & \times \mathbb{1}_{\{s_{B,t+1}=s_{B,t}\}}, \end{aligned} \tag{7}$$

and

$$\begin{aligned} P(\mathbf{s}_{t+1}|(\mathbf{s}_{B,t}, 0), a_t) = {} & p_{E,t+1}^{s_{E,t+1}} \cdot (1 - p_{E,t+1})^{1-s_{E,t+1}} \\ & \times \mathbb{1}_{\left\{\mathbf{s}_{B,t+1}=\left([\mathbf{s}_{B,t}]_{2:N_B}, l_t+a_t\right)\right\}}, \end{aligned} \tag{8}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function.

The reward function $u_t : \mathbb{R}^{N_B} \times \{0, 1\} \times A_t \to \mathbb{R}$ is the rebate minus the cost, namely

$$u_t(\mathbf{s}_t, a_t) = s_{E,t} \cdot r_t(f(\mathbf{s}_{B,t}), a_t) - c_t(a_t). \tag{9}$$

Without loss of optimality, we focus on stationary Markov policies, which depend on the current state and time only. We write a stationary Markov policy as

$$\pi = (\pi_1, \ldots, \pi_T) \text{ with } \pi_t : \mathbb{R}^{N_B} \times \{0, 1\} \to A_t, \quad (10)$$

and the set of all Markov policies as Π_M.

Our goal is to find the optimal stationary Markov policy that maximizes the expected total reward, namely solving the following dynamic program

$$\pi^* = \arg\max_{\pi \in \Pi_M} \mathbb{E}^{\pi}\left[\sum_{t=1}^{T} u_t(\mathbf{s}_t, a_t)\right], \quad (11)$$

where the expectation $\mathbb{E}^{\pi}$ depends on the policy π.

The dynamic program (11) can be solved through backward induction.

IV. Analysis

In this section, we focus on a special case of the dynamic program (11), for which we can provide structural results on the customer's optimal policy. Our structural results will provide insights on the customer's decision making process and on the design of the demand response program.

We consider the demand response programs where the baseline is determined as the average load of the previous N_B non-event days, namely

$$\bar{s}_{B,t} = f(\mathbf{s}_{B,t}) = \frac{\sum_{i=1}^{N_B} [\mathbf{s}_{B,t}]_i}{N_B}, \quad (12)$$

where $[\mathbf{s}_{B,t}]_i$ is the i-th element of the vector $\mathbf{s}_{B,t}$. Note that this is a common way to determine the baseline in many demand response programs [10]. We also consider a linear rebate function as in (3), so the customer pays penalties for load increases during events. The linear rebate could apply, for example, to an individual customer in [14] that is part of an aggregation.

Before stating our analytical results, we need to define a few useful quantities. The first one is the myopically optimal action in an event day, namely, the action that maximizes the event day's current reward (9):

$$a_{E,t} = \arg\max_{a_t \in A_t} r_t(\bar{s}_{B,t}, a_t) - c_t(a_t). \quad (13)$$

For each non-event day t, an important quantity, denoted by M_t, is the expected number of future event days whose baselines depend on the action a_t in day t. The baseline for an event day τ depends on the action a_t in day t if the day t load $l_t + a_t$ is one of the entries in the baseline state vector $\mathbf{s}_{B,\tau}$. This occurs when day t is a non-event day prior to event day τ and there are less than N_B non-event days during the period of days from $t+1$ to $\tau - 1$. Based on this, we define a binary indicator of whether day t affects the baseline of an event day τ:

$$m_\tau^t = \begin{cases} 1 & \text{if } s_{E,\tau} = 1 \text{ and } \bar{s}_{B,\tau} \text{ depends on } a_t \\ 0 & \text{otherwise} \end{cases}. \quad (14)$$

Note that m_τ^t is random because it depends on the realization of random event states $s_{E,t+1}, \ldots, s_{E,\tau}$. Finally, we can formally define M_t as

$$M_t = \mathbb{E}\left(\sum_{q=t+1}^{T} m_q^t\right). \quad (15)$$

Given M_t, we define

$$a_{N,t} = \arg\max_{a_t \in A_t} \frac{M_t \cdot r_{DR,t}}{N_B} \cdot a_t - c_t(a_t). \quad (16)$$

Proposition 1: In an event day, the optimal policy π^* chooses the myopically optimal action in that day, namely

$$\pi_t(\mathbf{s}_{B,t}, 1) = a_{E,t}, \ \forall t = 1, \ldots, T. \quad (17)$$

In a non-event day, assuming that the rebate is linear as in (3), and that the baseline is determined according to (12), the optimal policy π^* satisfies that

$$\pi_t(\mathbf{s}_{B,t}, 0) = a_{N,t}, \ \forall t = 1, \ldots, T. \quad (18)$$

Proof: Please see Appendix A. ■

Proposition 1 characterizes the customer's optimal policy.

In event days, the customer will choose an action that myopically optimizes the current reward. This is reasonable because the load in the event day will not be counted towards establishing future baselines. Therefore, the customer will focus on the current reward without worrying about the impact of its action on the future rewards. From the utility's perspective, our result also guarantees the simplicity of designing the rebate scheme: to ensure a certain level of demand response, the utility needs to consider the customer's cost of changing load only, but not other factors such as the probabilities of events, the number of non-event days used to calculate the baseline, and so on.

In non-event days, the customer will choose the action $a_{N,t}$ defined in (16). Note that this is not a myopic action, because the current reward in the non-event day is $-c_t(a_t)$. The objective function in (16) includes a term $\frac{M_t \cdot r_{DR,t}}{N_B} \cdot a_t$, which represents the expected future benefit of increasing the current load by a_t. The benefit comes from the inflated future baselines, which would result in higher future rebates. By exploiting the structure of the problem, we are able to characterize this benefit analytically as $\frac{M_t \cdot r_{DR,t}}{N_B} \cdot a_t$, leading to sharp structural results.

Proposition 1 also suggests that we can solve the dynamic program without using backward induction. In particular, we can avoid the "curse of dimensionality" if the problems in (13) and (16) are easy to solve, which is often the case. For example, when the rebate scheme is linear as in (3), the problem in (13) reduces to a convex optimization problem

$$a_{E,t} = \arg\max_{a_t \in A_t} -r_{DR,t} \cdot a_t - c_t(a_t), \quad (19)$$

which can be solved efficiently (even analytically if the cost function is linear or quadratic). Similarly, we can solve the convex optimization problem in (16) efficiently.

The remaining difficulty lies in how to compute M_t. Note that M_t depends only on the Bernoulli distributions of the event states only, but not on the baseline states, the actions taken, etc. In other words, we can compute M_t based on the probabilities $p_{E,1}, \ldots, p_{E,T}$. For some special cases, we are able to obtain analytical expressions for M_t.

Lemma 1: For a non-event day t, the expected number of event days whose baselines depend on the action a_t in day t can be computed analytically in the following cases.

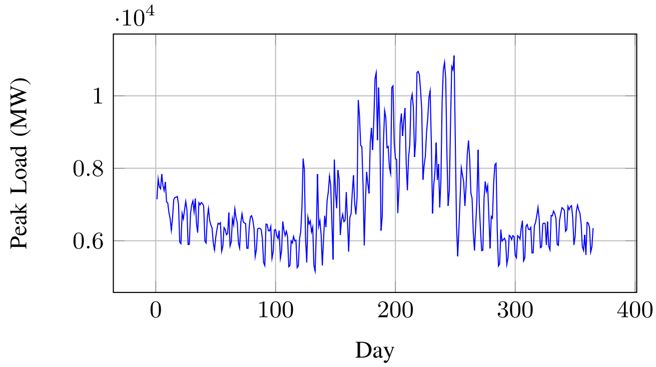

Fig. 2. Daily 2018 peak load for the New York City load zone, which includes the majority of Consolidated Edison's system load.

- When $N_B = 1$, we have

$$M_t = \sum_{s=t+1}^{T} (s-t) \cdot \left(\Pi_{\tau=t+1}^{s} p_{E,\tau}\right) \cdot \left(1 - p_{E,s+1}\right). \quad (20)$$

- When the probabilities of events are the same (i.e., $p_{E,t} = p_E, \ \forall t$), we have

$$M_t = \sum_{r=1}^{T-t} \sum_{s=(r-N_B)^+}^{r-1} \binom{r-1}{s} \cdot p_E^{s+1} \cdot (1-p_E)^{r-s-1}. \quad (21)$$

Proof: Please see Appendix B. ■

V. Model Implementation and Case Study

In this section, we describe a computer implementation of the model and present numerical results from a case study. We model several scenarios based on hypothetical customers in a DR program like Consolidated Edison's "Commercial System Relief Program - Voluntary Option," or "CSRP-V" (see Rider T in [14], [21]–[23]). Consolidated Edison operates a large electric utility serving the New York City area including 3.4 million electric customers. Consolidated Edison's system peak load in 2018 was 13 GW. Daily 2018 peak loads for the New York City load zone (which includes the majority of Consolidated Edison's system load) are shown in Fig. 2. CSRP-V is a baseline-based DR program with day-ahead event notice, thus matching the structure of our model. Results for the selected scenarios demonstrate how our model can reveal non-obvious relationships between program design, customer characteristics, and customer actions.

A. Model Implementation and Computational Complexity

We implemented a backward induction algorithm to solve the finite horizon dynamic program (11) for cases with discrete states and actions. The implementation can solve the optimal policy for a wide variety of DR program designs, including linear rebates with penalties as in (3) or payments capped above zero as in (4). Customer default loads, customer costs for each action, and DR rebate rates can also vary from day to day. The implementation also supports arbitrary values of probabilities of event for each day, thus providing solutions for cases beyond the special cases considered in (20) and (21).

The standard backward induction algorithm works as follows. For the last day T in the horizon, we solve for the optimal policy in day T by maximizing the reward in that day, namely

$$\pi_T(\mathbf{s}_T) = \arg\max_{a \in A_t} u_T(\mathbf{s}_T, a), \quad (22)$$

and obtained the value function defined as

$$V_T(\mathbf{s}_T) = \max_{a \in A_T} u_T(\mathbf{s}_T, a). \quad (23)$$

Then for each day $t = T-1, T-2, \ldots, 1$, we work backwards and solve for each day t's optimal policy by

$$\pi_t(\mathbf{s}_t) = \arg\max_{a \in A_t} \ u_t(\mathbf{s}_t, a) + \sum_{\mathbf{s}_{t+1}} P(\mathbf{s}_{t+1}|\mathbf{s}_t, a) V_{t+1}(\mathbf{s}_{t+1}), \quad (24)$$

where the second term in the objective is the expected future reward, and V_{t+1} is the value function at day $t+1$ defined as

$$V_{t+1}(\mathbf{s}_{t+1}) = \max_{a \in A_{t+1}} \ u_{t+1}(\mathbf{s}_{t+1}, a) + \sum_{\mathbf{s}_{t+2}} P(\mathbf{s}_{t+2}|\mathbf{s}_{t+1}, a) V_{t+2}(\mathbf{s}_{t+2}). \quad (25)$$

Now we discuss the computational complexity of the standard backward induction algorithm. Denote the number of states by N_s and the number of actions by N_a. As we can see from (24), the main computational complexity comes from evaluation of the expected future reward, which involves N_s multiplications. Since we need to evaluate this term for all pairs of states and actions, the total complexity is $O(N_s^2 N_a)$.

However, we can use the special structure in the state transition probabilities (7) and (8) to reduce the complexity. Since the transition of baseline states is deterministic, we can simplify (24) to

$$\pi_t\left(\mathbf{s}_{B,t}, 1\right) = \arg\max_{a \in A_t} \ \left\{u_t(\mathbf{s}_t, a) + p_{E,t+1} V_{t+1}\left(\mathbf{s}_{B,t}, 1\right) + \left(1 - p_{E,t+1}\right) V_{t+1}\left(\mathbf{s}_{B,t}, 0\right)\right\} \quad (26)$$

for event days, and

$$\pi_t\left(\mathbf{s}_{B,t}, 0\right) = \arg\max_{a \in A_t} \left\{u_t(\mathbf{s}_t, a) + p_{E,t+1} V_{t+1}\left(\left([\mathbf{s}_{B,t}]_{2:N_B}, l_t + a\right), 1\right) + \left(1 - p_{E,t+1}\right) V_{t+1}\left(\left([\mathbf{s}_{B,t}]_{2:N_B}, l_t + a\right), 0\right)\right\} \quad (27)$$

for non-event days. As we can see from (26) and (27), the computation of the future reward is simplified to a multiplication of two terms. Therefore, the computational complexity of our algorithm is $O(N_s N_a)$. By reducing the complexity by a factor of N_s, which scales exponentially as N_B, our implementation offers significant speedup.

Our implementation achieves this speedup by clever indexing of states to take advantage of the structure of the problem.

Specifically, the steps are as follows. We assign each possible value of the actual load an actual load index, i_{l+a}, which ranges from 0 to $N_{l+a} - 1$. We represent each baseline state by a series of N_B actual load indices, starting from the most recent day's load to the oldest load in the baseline. We consider this representation of the baseline state to be a base-N_{l+a} number, which we convert to a base-10 number to give the baseline state index, i_B. We can then discard the full representation of the baseline states, and work with just the indices. If the current day is a non-event day, the current day's baseline state index is $i_{B,t}$ and the current day's actual load index is $i_{l+a,t}$, then the next day's baseline state index is $i_{B,t+1} = \lfloor i_{B,t}/N_{l+a} \rfloor + i_{l+a,t} \cdot (N_{l+a})^{(N_B-1)}$, where $\lfloor \cdot \rfloor$ is the floor operator. Alternatively, if the current day is an event day, then the baseline state does not change and $i_{B,t+1} = i_{B,t}$. These calculations allow us to quickly determine the indices of the elements of V_{t+1} that we use in (26) and (27).

B. Description of Scenarios

We define eight representative scenarios for different types of customers, load, event probability profiles, rebate mechanisms, and customer costs. We investigate various performance metrics under these scenarios, in order to verify insights from our analytical results and reveal new findings that are difficult to obtain analytically. We discuss the results obtained under the assumption that we know the true event probabilities in Section V-C, and results obtained under imperfect estimates of event probabilities in Section V-D.

To facilitate understanding the discussion of scenarios and results, we define the terms in the first column of Table I. *Baseline Manipulation Allowed* indicates whether the set of available customer actions includes the possibility of increasing load, which allows for baseline inflation. *Probability of Event* indicates which type of daily event probability profile was used: *flat* for constant probability each day, matching the probability assumption in (21), or *spike* for constant (lower) probability most days, except for two consecutive days with higher event probabilities. Note that probabilities based on Consolidated Edison system load are considered in Sections V-D and V-E. *Negative Payments* indicates whether DR payments include penalties for negative DR quantities as in (3) or no penalties as in (4). *Default Load* indicates whether the customer default load has a constant value each day, or whether it varies from day to day. *Baseline Type* indicates whether baseline load is calculated as the average load in the top 5 out of the prior 10 non-event days (*5 in 10*), or as the average load in the prior 5 non-event days (*5 in 5*). *Customer Costs* indicates whether the customer costs of modifying load take on standard values or higher cost values, which are described later in this section. *Expected True DR (kWh)* is the expectation of the amount of load reduction in the event days, summed over all DR events in the DR season. *Expected Apparent DR (kWh)* is the expectation of the difference between baseline load and event load, summed over all DR events. *Expected DR Payments ($)* is the expected value of the DR rebates (3) or (4), summed over all DR events. *Expected Customer Costs ($)* is the expectation of costs of modifying load (5) summed over all days in the DR season. *Expected Customer Net Benefits ($)* is the difference between *Expected DR Payments ($)* and *Expected Customer Costs ($)*. *Expected Payment per Unit True DR ($/kWh)* is *Expected DR Payments ($)* divided by *Expected True DR (kWh)*.

We model customers participating in CSRP-V for the summer DR season (roughly 150 days). CSRP-V pays $3 per kWh of demand reduction during a pre-assigned 4-hour window on event days. For direct participants, payments are capped above zero, but for participants represented by an aggregator, negative event performance is counted and netted against contributions from other customers. The utility advises to expect 3 events per year, on peak load days. CSRP-V offers a "5 in 10" baseline methodology, where baseline load is average load over the 5 highest load days out of the preceding 10 similar days. We also consider an alternative baseline of the average load over the prior 5 similar days, or "5 in 5." For the "5 in 10" and "5 in 5" baseline scenarios, we include 10 or 5 days, respectively, before the start of the DR season in order to establish an initial baseline.

Scenario 1 models a customer who cannot manipulate baselines. The customer has three choices for actions and costs each day: (1) no load change at no cost; (2) reduce load by 1 kWh at the cost of $0.02; or (3) reduce load by 2 kWh at the cost of $2.02. This represents a customer who can use a battery to shift 1 kWh of energy (paying for 10% efficiency losses at the $0.20/kWh electricity rate), and can reduce appliance use to lower consumption 1 kWh more (costing $2/kWh for lost utility). The customer has no additional insight about when events will occur, so the probability of event each day is the total number of expected events per season divided by the number of days in the season (i.e., 3/150 = 0.02). The customer is not represented by an aggregator, so negative payments are not allowed. We set the default load to be the same each day.

Scenario 2 adds two additional action and cost options: (4) increase load by 1 kWh at the cost of $0.02; or (5) increase load by 2 kWh at the cost of $0.22. This represents a customer who can use a battery to shift 1 kWh of load (again paying for 10% losses), and can use appliances to increase load by an additional 1 kWh at the electricity rate of $0.20/kWh.

Scenario 3 is a modification of Scenario 2, representing a customer with more information about when events are likely to occur. This represents a customer that predicts event based on weather forecasts. This scenario includes two consecutive days with a high probability (i.e., 0.5) of events, and the remaining days with a lower event probability. The lower event probability is properly set so that the expected number of events is still 3.

Scenarios 4 through 8 include the alternative "5 in 5" baseline. Scenario 4 is identical to scenario 2 except for the alternative baseline type. Scenarios 5 through 8 include large day-to-day variations in default load, where the deviations are greater than the available customer load modifying actions. Scenarios 5 and 6 compare the impact of variable default load with negative payments not allowed and negative payments allowed, respectively. Scenarios 7 and 8 are identical to scenarios 5 and 6, except that the customer has a less efficient battery with 25% round trip losses, so the initial 1 kWh of reduction or increase costs $0.05 instead of $0.02.

TABLE I
SIMULATION SCENARIOS AND RESULTS

Scenario Number	1	2	3	4	5	6	7	8
Baseline Manipulation Allowed	no	yes	yes	yes	yes	yes	yes	yes
Probability of Event	flat	flat	spike	flat	flat	flat	flat	flat
Negative Payments	no	no	no	no	no	yes	no	yes
Default Load	flat	flat	flat	flat	variable	variable	variable	variable
Baseline Type	5 in 10	5 in 10	5 in 10	5 in 5	5 in 5	5 in 5	5 in 5	5 in 5
Customer Costs	standard	standard	standard	standard	standard	standard	higher	higher
Expected True DR (kWh)	6.0	6.0	6.0	6.0	3.7	6.0	2.9	6.0
Expected Apparent DR (kWh)	6.0	9.0	10.1	9.0	6.3	8.9	2.9	8.8
Expected DR Payments ($)	18.0	27.0	30.3	27.0	18.9	26.6	8.8	26.4
Expected Customer Costs ($)	6.1	7.6	8.6	9.0	6.4	9.0	3.3	13.3
Expected Customer Net Benefits ($)	11.9	19.4	21.8	17.9	12.5	17.6	5.4	13.1
Expected Payment per Unit True DR ($/kWh)	3.0	4.5	5.1	4.5	5.2	4.4	3.0	4.4

C. Results Under True Event Probabilities

We first assume that we know the true event probabilities. We summarize numerical results in Table I, and illustrate them in Figure 3, which shows the mean optimal policy for event and non-event days. The optimal policies shown in Fig. 3 are characterized by the optimal level of load reduction if a given day is an event day (dashed blue lines) and the optimal level of load inflation if a given day is a non-event day (solid red lines). In general, the optimal policy may depend on the baseline state. We plot the average optimal policies over all baseline states, which enables visualizing the policy when the number of states is large. The results reveal several findings, some of which corroborate our analytical results, and some of which address more complicated cases that would be difficult to analyze manually.

Scenario 1 shows that when manipulating baselines is not possible, the true DR quantity (load reductions due to actions during events) matches the apparent DR quantity (the difference between the potentially inflated baseline load and the event load), so that the utility payment per unit of true DR matches the CSRP-V program incentive rate ($3/kWh). Alternatively, in Scenario 2, the customer has the ability and incentive to inflate baseline load. This case has the same true DR quantity, but a larger apparent DR quantity due to the inflated baseline, thus resulting in 63% greater customer net benefits and a 50% higher utility cost per unit true DR. The differing optimal policies for Scenarios 1 and 2 is illustrated by plots (a) and (b) in Fig. 3, which shows load increases in non-event days for Scenario 2, but no changes to load in non-event days for Scenario 1.

Scenario 2 also shows that even with poor information about when events will occur, customers can have incentive to take inexpensive actions to inflate baselines (such as load shifting with a battery). With better information about event probabilities, as in Scenario 3, customers can have incentive to take more expensive actions to inflate baseline load (such as turning on more appliances) during the days prior to the days with high probability of events (as illustrated by plot (c) in Fig. 3). This is consistent with intuition gained from our analytical results for simpler programs: the expected number of event days affected by the current action, namely M_t, is higher during these days, and therefore the optimal load levels are higher according to (16). Scenario 3 has even greater apparent DR, higher customer profits, and 70% higher utility cost per unit true DR compared to Scenario 1 with no baseline inflation.

Scenario 4, with a "5 in 5" baseline, has nearly identical values of true DR, apparent DR, and utility cost, to Scenario 2, with a "5 in 10" baseline. Despite these similarities, the "5 in 10" baseline has the advantage of less demand inflation on non-event days. Note that the "5 in 10" baseline is beyond the scope of the analytical results we provided because it does not follow (12), so this illustrates how the computer implementation can generate non-obvious results for cases that have not been solved analytically.

Scenario 5 shows that varying default load can reduce incentives for both event response and baseline inflation, as illustrated by plot (e) in Fig. 3. This is because when payments are capped above zero and default load is much higher than the baseline, customer actions may not be able to increase the payment above zero. However, variable load does not have a major impact on true or apparent DR quantities when negative payments are allowed (note the similar results for Scenarios 4 and 6 in Table I). This is consistent with our analytical results, because neither (13) nor (16) depends on the load level l_t when the DR rebate is linear and allowed to be negative.

Scenarios 5–6 and Scenarios 7–8 indicate that capping payments at zero reduces incentives for both event response and baseline inflation. However, it has opposing effects on the utility cost per unit true DR when the customer's costs of increasing load changes. Specifically, under lower customer costs (Scenarios 5–6), the cost per unit true DR decreases when we allow negative payments (i.e., from $5.2/kWh in Scenario 5 to $4.4/kWh in Scenario 6); under higher customer costs (Scenarios 7–8), the cost per unit true DR increases when we allow negative payments (i.e., from $3.0/kWh in Scenario 7 to $4.4/kWh in Scenario 8). Therefore, the impact of penalty on the program performance is non-obvious and hard to predict analytically. This emphasizes the importance of our computational solution to the general model.

D. Imperfect Knowledge of Event Probabilities

The results shown thus far in this article are based on the assumption that the customer knows the probabilities of events for each day in the DR season. In practice, the customer does not have perfect information about these event probabilities. This section demonstrates one approach that companies or systems working on behalf of customers could use to estimate

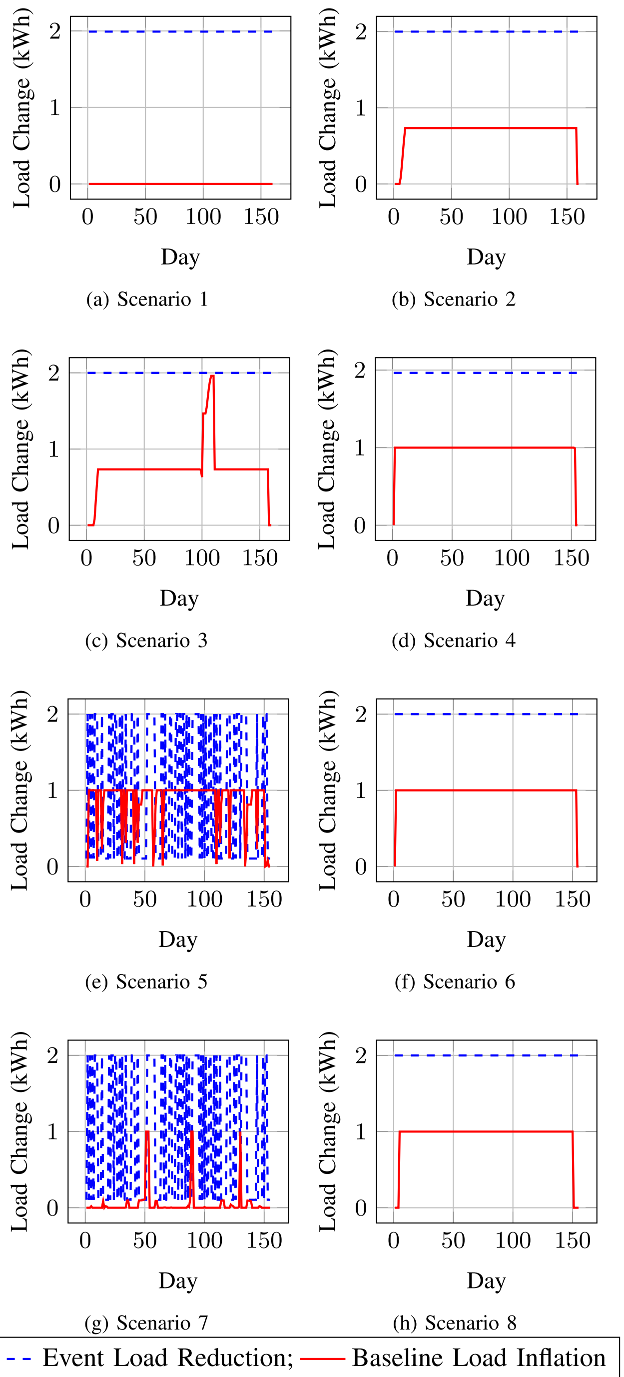

Fig. 3. Average optimal policy by day for scenarios 1-8. Dashed blue lines show the optimal level of load reduction on event days, as an average over all baseline states. Solid red lines show the optimal level of load inflation on non-event days, as an average over all baseline states.

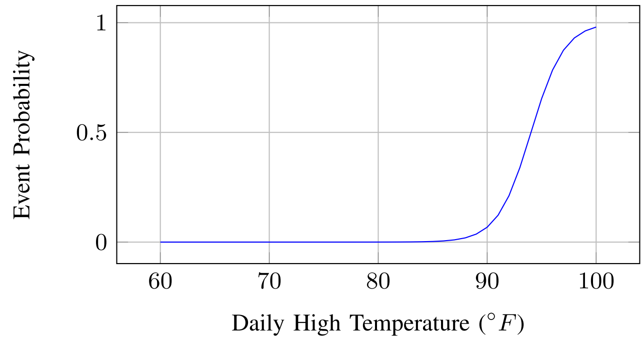

Fig. 4. Event probability versus daily high temperature, determined from logistic regression of historical temperature and Consolidated Edison system load.

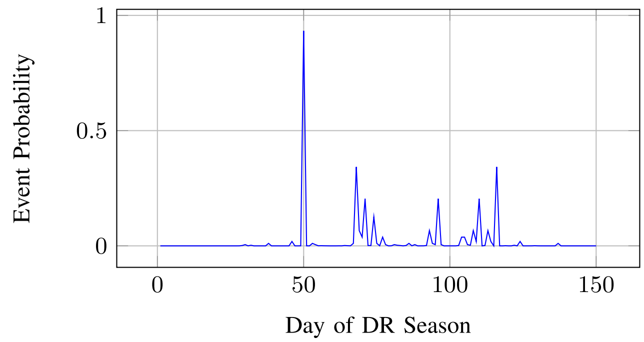

Fig. 5. Daily event probabilities for New York City Typical Meteorological Year 3 weather data, based on logistic regression model from Fig. 4.

event probabilities. Then we analyze the customer benefits of manipulating baselines based on imperfect event probability forecasts, compared to a simpler DR response strategy with no baseline manipulation.

Many DR programs are designed to respond to system load conditions. Since system load is typically correlated with weather, the probability of having DR events is also correlated with weather. When we have historical data of weather and system load or DR events, we can perform a regression to learn the mapping from weather conditions to event probabilities. Then, this mapping can be used to create forecasts of future DR event probabilities based on weather forecasts.

Fig. 4 shows the result of a simple logistic regression model of probability of event versus daily peak temperature based on Con Edison 2018 system load data and 2018 New York City daily peak temperatures. In this analysis, we assign DR events to the 3 peak load days of 2018 (based on the expectation of 3 events per year). For the purposes of this study, we assume that the mapping from temperature to event probability shown in Fig. 4 is the “true” mapping. We leave to future work the task of improving mapping of weather to event probability based on additional weather parameters (such as humidity), more years of data, and more data about DR event occurrence.

Next, we assess the performance of our DR optimization model based on a typical year’s weather in New York City and various qualities of event probability forecasts. In this case, the actual realized weather corresponds to a typical year of weather for JFK Airport in New York City (based on Typical Meteorological Year 3 data), and the “true” event probabilities are calculated based on that weather data and the regression model in Fig. 4. Fig. 5 shows the true event probabilities.

Now we can demonstrate the robustness of our results under imperfect knowledge of event probabilities. We run our algorithm to obtain the customer’s optimal policy given *erroneous estimates of event probabilities*. We investigate two types of erroneous estimates. The first type of erroneous estimates come from the logistic regression method with

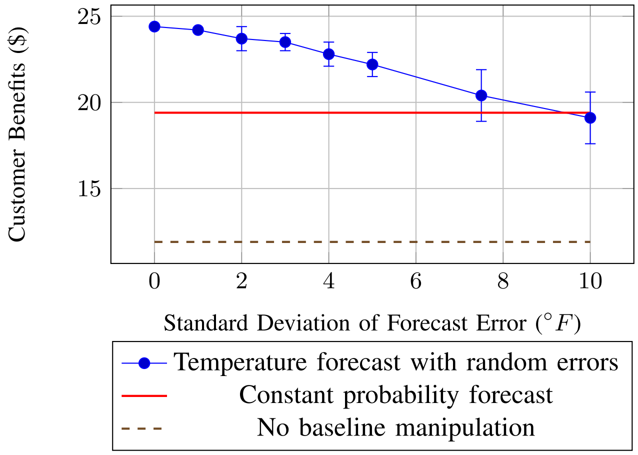

Fig. 6. Customer benefits under different types of event forecasts. Blue dots show benefits under event forecasts based on temperature forecasts with Gaussian random errors of varying standard deviations. The solid red line shows benefits using a constant probability forecast. The dashed line shows benefits if the customer does not manipulate baselines. Both the red line and the dashed line are horizontal, because the standard deviation of temperature forecast errors is not applicable to the constant probability event forecast and does not affect the policy without baseline manipulation.

erroneous weather forecasting as input. We included results under erroneous estimates with different levels of estimation errors (resulting from inaccurate weather forecasting with 1-degree to 10-degree errors). The second type of erroneous estimates is a "constant event probability estimation", where everyday has the same event probability, calculated as the expected number of events per year divided by the number of days in a year. This constant probability forecast represents a scenario where the customer does not use any sophisticated method to estimate event probabilities at all. Then we evaluate the policies computed based on erroneous estimates of event probabilities *under the true event probabilities*. In Fig. 6, we can see that the performance of the policies is robust to the estimation errors. For example, under 5-degree errors in weather forecasting, the policy can still achieve 90% of the performance of the optimal policy computed given true event probabilities (the optimal policy corresponds to the left-most blue dot with zero forecast error). Therefore, the policy is close to optimal even when the event probability estimates have moderate errors.

The dashed horizontal line show optimal customer benefits when customer resources are not used to manipulate baselines. Comparing this to the other cases, we see that a baseline manipulation strategy based on a constant probability forecast has significant customer benefits relative to the no baseline manipulation case. If reasonably accurate weather forecasts are available, customers can achieve even greater benefits by manipulating baselines. Therefore, the qualitative conclusion of this article, namely that the customer has incentives to increase the baseline load, is robust to imperfect knowledge on the event probabilities. This observation emphasizes the importance of understanding customer behavior and designing programs that discourage undesired incentives.

In summary, we have demonstrated one feasible approach to estimate event probabilities using logistic regression, and that empirically, our results are robust to imperfect event probability forecasts. A rigorous approach with theoretical guarantees on robustness (e.g., robust optimization) is an important future direction of research.

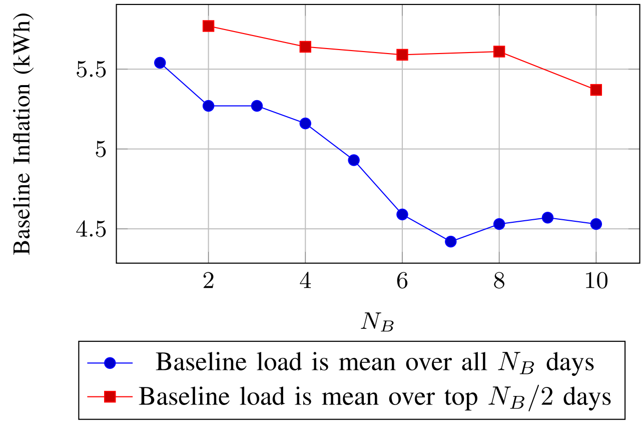

Fig. 7. Baseline load inflation for various baseline determination rules.

E. Sensitivity Analysis of Baseline Determination Rules

We conduct a sensitivity analysis on baseline determination rules (2), where we vary the number N_B of non-event days used for computing the baseline, and change how to compute the baseline, namely as an average over all N_B days or the $N_B/2$ days with maximum load. We demonstrate how our model can inform the design of a baseline-based DR program. In all simulations, we use the event probabilities in Fig. 5.

Fig. 7 shows the level of baseline inflation for various baseline types. We see that using all N_B days to calculate the baseline load leads to less baseline inflation than using the top $N_B/2$ days. This makes sense, because using the top $N_B/2$ out of the N_B days to calculate baseline load allows customers to achieve a given level of baseline inflation with less total load increase over the N_B days, thus making baseline inflation cheaper and more attractive.

We also see that the level of baseline inflation is generally higher when N_B is smaller. However, the trend is not monotone. In general we would expect two competing effects as we increase N_B. First, as N_B increases, to achieve the same level of baseline inflation, the total amount of load increase during the N_B non-event days increases, which increases the cost and makes baseline inflation less attractive. Second, under higher N_B, increasing load in a single baseline day is expected to inflate baseline load in more event days, making baseline inflation more attractive. The overall result of the two competing effects depends on many factors, such as the profile of event probabilities, the rebate rate, and the cost function. Thus, it is important to have our computational solution to help determine the optimal N_B based on the various parameters.

VI. Conclusion

Our baseline-based DR model can identify optimal customer behaviors under a variety of baseline-based DR program parameters and a wide variety of customer parameters, thus revealing customer incentives to artificially inflate baselines when the schedule of DR events is uncertain. Analytical results provide some fundamental insights into the drivers of optimal customer decisions. Numerical results show that incentives for baseline manipulation may exist in real-world DR programs, and that the level and impact of baseline manipulation can depend on a number of factors, in ways that would be difficult to predict without our model. Since the impacts of manipulating baselines can have significant and non-obvious impacts on the cost and effectiveness of baseline-based DR programs, customer incentives should be considered during DR program design and in utility planning and operations. An important future work is to optimize the DR program (e.g., the rebate function) given the customer's decisions.

Appendix A
Proof of Proposition 1

We first prove that on the event day, the user chooses the myopically optimal action $a_{E,t}$. This can be proved by looking at the Bellman equation:

$$\pi_t(\mathbf{s}_{B,t}, 1) = \arg\max_{a\in A_t} u_t\big[(\mathbf{s}_{B,t}, 1), a\big] + \mathbb{E}\{V_{t+1}(\mathbf{s}_{t+1})\} \quad (28)$$

where V_{t+1} is the value function at $t+1$. Because the load a_t on the event day t will not be counted in future baselines, the baseline state remains the same, namely $\mathbf{s}_{B,t+1} = \mathbf{s}_{B,t}$. Therefore, the future expected value can be rewritten as

$$\mathbb{E}\{V_{t+1}(\mathbf{s}_{t+1})\} = \mathbb{E}_{s_{E,t+1}}\big\{V_{t+1}\big(\mathbf{s}_{B,t}, s_{E,t+1}\big)\big\}. \quad (29)$$

Since the action a_t does not affect either the probability distribution of the event state $s_{E,t+1}$ or the baseline state $\mathbf{s}_{B,t}$, it does not affect the future expected value $\mathbb{E}\{V_{t+1}(\mathbf{s}_{t+1})\}$. As a result, the optimal action $\pi_t(\mathbf{s}_{B,t}, 1)$ should maximize the current payoff $u_t[(\mathbf{s}_{B,t}, 1), a_t]$, which is exactly defined as $a_{E,t}$ in (13).

Now we prove that on the non-event day, the user chooses her action according to (16). We prove this by directly computing the expected total future reward, instead of by looking at the Bellman equation. At a non-event day t, under state $(\mathbf{s}_{B,t}, 0)$, the expected total reward is

$$u_t\big[(\mathbf{s}_{B,t}, 1), a_t\big] + \mathbb{E}\left\{\sum_{\tau=t+1}^{T} u_\tau(\mathbf{s}_\tau, a_\tau)\right\}, \quad (30)$$

where the expectation is taken over the random future states $\mathbf{s}_{t+1}, \ldots, \mathbf{s}_T$. Note that the expected total future reward $\mathbb{E}\{\sum_{\tau=t+1}^{T} u_\tau(\mathbf{s}_\tau, a_\tau)\}$ implicitly depends on the current action a_t, because the future baseline states depend on the current load a_t.

Since the baseline is determined as the average load during the previous N_B non-event days as in (12), given a sequence of realized event states $s_{E,t+1}, \ldots, s_{E,T}$ and a sequence of actions $a_{t+1}, \ldots, a_T$, the baseline states evolve deterministically. In other words, the randomness comes from the event states only. Therefore, the expected total future payoff from day $t+1$ to day T can be calculated as

$$\mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\left\{\sum_{\tau=t+1}^{T} u_\tau\big[\big(\mathbf{s}_{B,\tau}, s_{E,\tau}\big), a_\tau\big]\right\}$$
$$= \mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\left\{\sum_{\tau=t+1}^{T} s_{E,\tau}\cdot r_\tau\big[f\big(\mathbf{s}_{B,\tau}\big), a_\tau\big] - c_\tau(a_\tau)\right\}$$
$$= \mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\left\{\sum_{\tau=t+1}^{T} s_{E,\tau}\cdot r_\tau\big(\bar{s}_{B,\tau}, a_\tau\big) - c_\tau(a_\tau)\right\}, \quad (31)$$

where the baseline $\bar{s}_{B,\tau}$ is determined by

$$\bar{s}_{B,\tau} = \frac{\sum_{i=1}^{N_B}\big[\mathbf{s}_{B,\tau}\big]_i}{N_B}. \quad (32)$$

To determine the optimal action in day t, we only need to consider the components in (31) that depend on a_t. First, we do not need to consider the cost $c_\tau(a_\tau)$ for $\tau > t$. Second, we only need to consider the reward $r_\tau(\bar{s}_{B,\tau}, a_\tau)$ if day τ is an event day (i.e., $s_{E,\tau} = 1$) and if the action a_t plays a role in determining the baseline $\bar{s}_{B,\tau}$. Therefore, it is useful to define binary random variables m_τ^t that indicate whether day τ has an event whose baseline $\bar{s}_{B,\tau}$ depends on load a_t. Formally, we have

$$m_\tau^t = \begin{cases} 1 & \text{if } s_{E,\tau} = 1 \text{ and } \bar{s}_{B,\tau} \text{ depends on } a_t \\ 0 & \text{otherwise} \end{cases}. \quad (33)$$

Furthermore, since the reward is linear, namely

$$r_\tau\big(\bar{s}_{B,\tau}, a_\tau\big) = r_{DR,\tau}\cdot\big(\bar{s}_{B,\tau} - a_\tau\big), \quad (34)$$

the contribution of the action a_t in the above reward is

$$r_{DR,\tau}\cdot\frac{a_t}{N_B} \quad (35)$$

if $m_\tau^t = 1$. Combining this observation with the definition of m_τ^t, we can write the components in (31) that depend on the action a_t as

$$\mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\left\{\sum_{\tau=t+1}^{T} m_\tau^t\cdot r_{DR,\tau}\cdot\frac{a_t}{N_B}\right\}$$
$$= r_{DR,\tau}\cdot\frac{a_t}{N_B}\cdot\mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\left\{\sum_{\tau=t+1}^{T} m_\tau^t\right\}. \quad (36)$$

Note that the expectation $\mathbb{E}_{s_{E,t+1},\ldots,s_{E,T}}\{\sum_{\tau=t+1}^{T} m_\tau^t\}$ is actually the expected number of event days whose baselines depend on the load a_t in day t, denoted by M_t. Therefore, the part of the expected future payoff that depends on the action a_t is

$$r_{DR,\tau}\cdot\frac{a_t}{N_B}\cdot M_t. \quad (37)$$

In addition, in a non-event day t, the action affects the cost $c_t(a_t)$ only.

In summary, the action a_t affects

$$-\,c_t(a_t) + r_{DR,\tau}\cdot\frac{a_t}{N_B}\cdot M_t. \quad (38)$$

The optimal action $\pi_t(\mathbf{s}_{B,t}, 0)$ that maximizes the expected total payoff should be $a_{N,t}$ as defined in (16).

APPENDIX B
PROOF OF LEMMA 2

In (33), we defined binary random variables m_τ^t to indicate whether day τ has an event whose baseline $\bar{s}_{B,\tau}$ depends on load a_t. Note that m_τ^t is always 0 for $\tau \le t$, because actions can only impact baselines for future days.

The expectation of the number of event days whose baselines depend on the action a_t in day t, denoted by M_t, is then

$$M_t = \mathbb{E}\left[\sum_{q=t+1}^{T} m_q^t\right] = \sum_{q=t+1}^{T} \mathbb{E}\left[m_q^t\right]. \tag{39}$$

Now we calculate M_t for the special cases in Lemma 2.

1) *The Special Case of $N_B = 1$:* If the baseline depends only on the load in the previous non-event day, we have $m_\tau^t = 1$ if and only if $m_q^t = 1$ for all $t < q < \tau$. In other words, if day τ is an event day whose baseline depends on action a_t, all the days between day t and day τ must be event days. Therefore, the possible sequences of the indicators $m_{t+1}^t, \ldots, m_T^t$ are

$$m_q^t = \begin{cases} 1 & q \le \tau \\ 0 & q = \tau + 1 \ , \quad \tau = t+1, \ldots, T, \\ * & q > \tau + 1 \end{cases} \tag{40}$$

where $*$ means that m_q^t can be either 0 or 1. For a given $\tau \in [t+1, T]$, the above sequence happens with probability

$$\left(\Pi_{q=t+1}^{\tau} p_{E,q}\right) \cdot \left(1 - p_{E,\tau+1}\right). \tag{41}$$

In addition, there are $\tau - t$ event days whose baselines depend on a_t in this sequence. Thus, the expected number of event days whose baselines depend on a_t is

$$\sum_{\tau=t+1}^{T} (\tau - t) \cdot \left(\Pi_{q=t+1}^{\tau} p_{E,q}\right) \cdot \left(1 - p_{E,\tau+1}\right). \tag{42}$$

2) *The Special Case Where the Probabilities of Events are the Same (i.e., $p_{E,t} = p_E, \ \forall t$):* Now, we calculate $\mathbb{E}[m_q^t]$ for this case. To assist in the calculation, we define a sample space $\Omega = \{0, 1\}^{q-t}$, the Cartesian product of $q - t$ copies of the set {0,1}. Each element of the sample space represents a sequence of 1's for events and 0's for non-events for each day from $t + 1$ through q, and the sample space includes every possible sequence of event and non-event days. The digits of each element of Ω are independent Bernoulli(p_E) random variables, because the probability of event each day is p_E and is independent of the other days. This sequence determines whether day q is an event day whose baseline depends on a_t. Specifically, day q will be an event day whose baseline depends on a_t if and only if there are less than N_B 0's in the first $q - t - 1$ digits of the sequence (i.e., less than N_B non-event days in days $t + 1$ through $q - 1$), and the final digit of the sequence is a 1 (i.e., day q is an event day).

We can partition the sample space based upon the number of events that occur from day $t + 1$ through day $q - 1$. We define D_s to be the logical event corresponding to s DR events occurring during those days. The minimum possible value of s is 0, and the maximum value is $q-t-1$. By the law of total expectation, we have

$$\mathbb{E}\left[m_q^t\right] = \sum_{s=0}^{q-t-1} \mathbb{E}\left[m_q^t \mid D_s\right] \cdot \mathbb{P}[D_s]. \tag{43}$$

We know that if there are N_B or more non-event days from day $t + 1$ through day $q - 1$, then a_t will not be included in the baseline for day q, and m_q^t will be 0. In order to have less than N_B non-event days, there must be at least $q - t - N_B$ event days, or if $q - t - N_B$ is negative (i.e., there are less than N_B days from $t+1$ through $q-1$) the minimum number of event days is 0. We can concisely express the minimum number of event days over days $t+1$ through $q-1$ to ensure there are less than N_B non-event days as $(q-t-N_B)^+$, where $(\cdot)^+ = \max\{\cdot, 0\}$. Using this, we discard the terms that equal zero in (43) to get

$$\mathbb{E}\left[m_q^t\right] = \sum_{s=(q-t-N_B)^+}^{q-t-1} \mathbb{E}\left[m_q^t \mid D_s\right] \cdot \mathbb{P}[D_s]. \tag{44}$$

We also know that if there are less than N_B non-event days from day $t + 1$ through day $q - 1$, then a_t will be included in the baseline for day q. In this case, m_q^t will be 1 if day q is an event day (with probability p_E), or will be 0 otherwise. Based on this, we see that $\mathbb{E}[m_q^t \mid D_s] = p_E$ over all values of s in the summation in (44). Using this, we have

$$\mathbb{E}\left[m_q^t\right] = \sum_{s=(q-t-N_B)^+}^{q-t-1} p_E \cdot \mathbb{P}[D_s]. \tag{45}$$

Now, we calculate $\mathbb{P}[D_s]$. $\mathbb{P}[D_s]$ is the probability of having s DR events out of days $t + 1$ through $q - 1$, which includes $q - t - 1$ days. This is equivalent to the probability that a Binomial$(q - t - 1, p_E)$ random variable takes on the value s. Thus

$$\mathbb{P}[D_s] = \binom{q-t-1}{s} \cdot p_E^s \cdot (1 - p_E)^{q-t-s-1} \tag{46}$$

Substituting (46) into (45), substituting the result into (39), and using the substitution $r = q - t$ gives the final result (21).

REFERENCES

[1] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, “A survey on demand response programs in smart grids: Pricing methods and optimization algorithms,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 152–178, 1st Quart., 2015.

[2] B. Chew, B. Feldman, D. Ghosh, and M. Surampudy, “2018 utility demand response market snapshot,” Smart Electric Power Alliance, Washington, DC, USA, Rep., 2018. [Online]. Available: https://sepapower.org/resource/2018-demand-response-market-snapshot/

[3] A. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, “Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid,” *IEEE Trans. Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec. 2010.

[4] N. Li, L. Chen, and S. H. Low, “Optimal demand response based on utility maximization in power networks,” in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Detroit, MI, USA, Jul. 2011, pp. 1–8.

[5] Y. Xu, N. Li, and S. H. Low, “Demand response with capacity constrained supply function bidding,” *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1377–1394, Mar. 2016.

[6] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong, "Advanced demand side management for the future smart grid using mechanism design," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1170–1180, Sep. 2012.
[7] D. Muthirayan, D. Kalathil, K. Poolla, and P. Varaiya, "Mechanism design for demand response programs," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 61–73, Jan. 2020.
[8] J. Vuelvas and F. Ruiz, "Rational consumer decisions in a peak time rebate program," *Elect. Power Syst. Res.*, vol. 143, pp. 533–543, Feb. 2017.
[9] K. Khezeli and E. Bitar, "Risk-sensitive learning and pricing for demand response," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6000–6007, Nov. 2018.
[10] P. Antunes, P. Faria, and Z. Vale, "Consumers performance evaluation of the participation in demand response programs using baseline methods," in *Proc. IEEE Grenoble Conf.*, Jun. 2013, pp. 1–6.
[11] J. Vuelvas, F. Ruiz, and G. Gruosso, "Limiting gaming opportunities on incentive-based demand response programs," *Appl. Energy*, vol. 225, pp. 668–681, Sep. 2018.
[12] D. G. Dobakhshari and V. Gupta, "A contract design approach for phantom demand response," *IEEE Trans. Autom. Control*, vol. 64, no. 5, pp. 1974–1988, May 2019.
[13] D. Muthirayan, E. Baeyens, P. Chakraborty, K. Poolla, and P. P. Khargonekar, "A minimal incentive-based demand response program with self reported baseline mechanism," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2195–2207, May 2020.
[14] Con Edison. (2019). *Commercial Demand Response Program Guidelines*. [Online]. Available: https://www.coned.com/-/media/files/coned/documents/save-energy-money/rebates-incentives-tax-credits/smart-usage-rewards/smart-usage-program-guidelines.pdf?la=en
[15] Federal Energy Regulatory Commission. (2011). *Order No. 745, Demand Response Compensation in Organized Wholesale Energy Markets*. [Online]. Available: https://www.ferc.gov/sites/default/files/2020-04/OrderNo.745.pdf
[16] F. Wolak. (2006). *Residential Customer Response to Real-Time Pricing: The Anaheim Critical-Peak Pricing Experiment*. [Online]. Available: https://escholarship.org/uc/item/3td3n1x1
[17] J. Pierobon. (2013). *Two FERC Settlements Illustrate Attempts to "Game" Demand Response Programs*. [Online]. Available: https://www.energycentral.com/c/ec/ferc-settlements-illustrate-attempts-game-demand-response-programs
[18] S. Borenstein. (2014). *Money for Nothing?* [Online]. Available: https://energyathaas.wordpress.com/2014/05/12/money-for-nothing/
[19] J. Vuelvas and F. Ruiz, "A novel incentive-based demand response model for cournot competition in electricity markets," *Energy Syst.*, vol. 10, no. 1, pp. 95–112, 2019.
[20] D. Ellman and Y. Xiao, "Customer incentives for gaming demand response baselines," in *Proc. IEEE 58th Conf. Decis. Control (CDC)*, 2019, pp. 5174–5179.
[21] Consolidated Edison Company of New York Inc. (2019). *Schedule For Electricity Service*. [Online]. Available: https://www.coned.com/_external/cerates/documents/elecPSC10/electric-tariff.pdf
[22] Con Edison. (2018). *Demand Response Forum Presentation*. [Online]. Available: https://www.coned.com/-/media/files/coned/documents/save-energy-money/rebates-incentives-tax-credits/smart-usage-rewards/demand-response-forum.pdf
[23] Consolidated Edison Company of New York Inc. (2018). *Energy Efficiency and Demand Management Procedure–General: Calculating Customer Baseline Load*. [Online]. Available: https://www.coned.com/-/media/files/coned/documents/save-energy-money/rebates-incentives-tax-credits/smart-usage-rewards/customer-baseline-load-procedure.pdf?la=en

Douglas Ellman (Graduate Student Member, IEEE) received the A.B. degree in physics from Princeton University, Princeton, NJ, USA, in 2009, and the M.S. degree in technology and policy from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering from the University of Hawai'i at Mānoa, Honolulu, HI, USA, and a Research Assistant with the Hawai'i Advanced Wireless Technologies Institute. His research interests include applying optimization and machine learning to the integration of distributed energy resources and renewable energy into electric power systems.

Yuanzhang Xiao (Member, IEEE) received the B.E. and M.E. degrees in electronic engineering from Tsinghua University in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from UCLA in 2014. He has been an Assistant Professor with the University of Hawai'i at Mānoa since 2017. He was a Postdoctoral Fellow with Northwestern University from 2015 to 2017. His research interests include reinforcement learning, game theory, and optimization, with applications in smart grids, wireless communications, and socio-technological networks.