Generalizing Demand Response Through Reward Bidding

Hongyao Ma
Harvard University
Cambridge, MA, US
hma@seas.harvard.edu

David C. Parkes
Harvard University
Cambridge, MA, US
parkes@eecs.harvard.edu

Valentin Robu
Heriot Watt University
Edinburgh, UK
V.Robu@hw.ac.uk

ABSTRACT

Demand-side response (DR) is emerging as a crucial technology to assure stability of modern power grids. The uncertainty about the cost agents face for reducing consumption imposes challenges in achieving reliable, coordinated response. In recent work, Ma et al. [13] introduce DR as a mechanism design problem and solve it for a setting where an agent has a binary preparation decision and where, contingent on preparation, the probability an agent will be able to reduce demand and the cost to do so are fixed. We generalize this model to allow uncertainty in agents' costs of responding, and also multiple levels of effort agents can exert in preparing. For both cases, the design of contingent payments now affects the probability of response. We design a new, truthful and reliable mechanism that uses a "reward-bidding" approach rather than the "penalty-bidding" approach. It has good performance when compared to natural benchmarks. The mechanism also extends to handle multiple units of demand response from each agent.

Keywords

mechanism design; demand response; reliability bounds

1. INTRODUCTION

The task of maintaining an exact balance of the supply and demand in power systems is increasingly challenging, due to the increasing penetration of intermittent renewable generation [24, 25], and the presence of more volatile types of loads, such as those from electric vehicle charging [21]. This has lead to an increasing interest in *demand-side response* (DR), in which consumers commit to temporarily reduce or shift consumption away from periods where generation capacity does not meet the aggregate demand [15].

In contrast to the operating reserves on the supply side, where the cost and ability for a generator to increase power output can be known with high precision when planning one day ahead, consumers on the demand side face uncertainty about their future costs for reducing consumption. Consider an industrial factory which uses electricity for the production line, transporting raw materials, and cooling. Its ability to respond to a DR event may depend on the production pro-

cedure, time of day when called for DR, customer requests, and weather conditions, thus is highly uncertain. This imposes challenges on selecting and incentivizing a subset of the consumers to meet a total reduction target with high probability (what we call the "global reliability constraint"), without selecting too many consumers to prepare or leading to excessive economic disruption.

Ma et al. [13] introduce reliable DR as a two period mechanism design problem, where the planner is the electricity grid (or DR aggregator) and the agents are consumers interested in offering DR services. In the planning period (period zero) consumers may opt-in to a DR scheme and make reports to the mechanism based on their probabilistic information on their costs and abilities to respond. A subset of these consumers are selected and asked to prepare for demand reduction. Later, in the event DR is required (period one), based on the resolved uncertainty, selected consumers can decide on whether to follow-through and respond to receive a reward, or not to respond and pay a penalty.

Two truthful and reliable "penalty-bidding" mechanisms with fixed rewards are proposed in [13], where agents are selected in decreasing order of their maximum acceptable penalties in the event of non-response. The model on agents' types (each agent can choose to prepare for DR in period zero at a fixed preparation cost, and if so, in period one, she will be able to reduce, with some *fixed* probability, one unit of consumption at a *fixed* opportunity cost), however, does not reflect the reality that with higher rewards and penalties agents will be incentivized to respond with higher probabilities. We generalize the model in the following ways:

(i) Uncertain costs: having prepared in period zero, agents are still uncertain about the costs they will face, and will decide on whether to respond in period one after the actual costs are realized. Agents are more likely to respond when the rewards and penalties are high.

(ii) Multi-effort-level: agents may have multiple levels of effort they can exert when preparing to respond. Higher rewards and penalties may induce a higher preparation level, resulting in a higher probability of responding.

(iii) Multi-unit: each agent is able to reduce a varying amount of consumption, and has probabilistic information about its values for different consumption levels.

The dependence of the global reliability constraint on both agents' types and the payments contingent on agents' responses introduces tensions among selecting a small set of agents, satisfying the reliability constraint, and truthfully eliciting information from the agents. We will see that the

"penalty-bidding" mechanisms [13] do not generalize to achieve reliable, dominant-strategy demand response.

Our main contribution is to design a new, truthful and reliable mechanism for this generalized setting that uses a "reward-bidding" approach. In particular, the mechanism adopts a fixed penalty for non-response for all selected agents, and agents are selected in increasing order of their minimum acceptable rewards given this penalty. Thus, the mechanism implements the idea of reward bidding. The reward offered to a selected agent is large enough that it will choose to prepare to reduce demand, and follow-through with high probability. This is not possible with penalty-bidding mechanisms, because of a subtle interaction between incentive constraints and the need to select a set of agents such that the reliability constraint will be satisfied.

For the multi-unit scenario, we generalize the mechanism by introducing linear payment schedules where rewards and penalties are defined per unit of consumption. We also show that we can handle the possibility of multiple levels of preparation efforts in both the single and multi-unit scenarios.

We demonstrate in simulation that the reward-bidding mechanism achieves close to the first best (i.e. assuming the mechanism knows agent types and therefore how reliable they would be given certain payments) with regard to the number of selected agents. We also benchmark against a spot auction in which a reverse auction is used to achieve a required reduction in consumption. The reward-bidding mechanism achieves the same reliability with lower expected total payments, and much less variance in payments.

1.1 Related Work

Some prior work that considers uncertainty in agent types includes: research on promoting utilization of shared resources in the context of coordination problems [12], maximizing social welfare in a setting with uncertainty about agent actions [18], and maximizing an airline's expected revenue in a setting where passengers have uncertainty about whether or not they may fly and thus refund menus can be useful [7]. Also related is work on dynamic mechanism design with dynamic agent type [3, 4, 5, 16]. But none of this prior work has the objective to satisfy a probabilistic constraint on the joint actions taken by the agents. Most closely related is a paper that introduced the problem of mechanism design for reliable demand response [13]. The present work significantly generalizes this by allowing for uncertain opportunity costs, multiple effort levels, and varying units of possible consumption reduction.

A number of works on demand response have discussed the concept of aggregation of multiple agents, both aggregation of small intermittent generators, and aggregation of uncertain demands [25, 24, 22, 20]. Some of these works propose the use of *scoring rules* to incentivize truthful reports about expected future generation or consumption [23, 1, 20]. Unlike a scoring rule approach, in this work the rewards and penalties of selected agents are determined by the market, from the reports made by other agents. This guarantees that rewards are set such that the selected subset of agents will guarantee the system-wide reliability constraint.

Other prior works on demand response markets (e.g. [11, 9, 17]) consider agents bidding using supply curves, and study the market equilibria or these settings. They do not, however take a mechanism design perspective or guarantee truthful reporting. Pricing mechanisms to incentivize load

shifting have also been studied in [2, 19, 10]. We focus on achieving reliable DR in one future period, whereas analyzing load shifting requires modeling of agents' uncertain valuations for different consumption profiles over extended periods of time.

2. PRELIMINARIES

We now model the single-unit DR problem in which each agent can reduce the same amount of consumption, and defer the model and mechanism for multi-unit DR to Section 4.

Uncertain Costs.

Let $N = \{1, 2, \ldots, n\}$ denote the set of agents, each of which can prepare for demand response ahead at a cost of $c_i > 0$. If an agent prepares for demand reduction, her cost for reducing one unit of consumption will be a random variable V_i with non-negative support, finite expectation and cumulative distribution function (CDF) F_i. V_i represents the uncertain opportunity cost for the loss of electricity, the exact value of which is not realized until later. The pair $\theta_i = (c_i, F_i)$ defines an agent's *type* and is agent i's private information. Let $\theta = (\theta_1, \ldots, \theta_n)$ denote a type profile. We assume in our model that an agent can only respond if she first prepares, and that the opportunity costs of agents are independently distributed.

Note that the discrete single-unit (v_i, p_i, c_i) model proposed in [13], where an agent can reduce one unit of consumption with probability p_i at a cost of v_i if she prepares at the cost of c_i, is a special case of the uncertain cost model, where $V_i = v_i$ with probability p_i and $V_i = -\infty$ (representing the hard constraint) with probability $1 - p_i$.

Reliability Target.

Denote $M \in \mathbb{N}_+$ as the *target capacity reduction* that needs to be achieved. The objective of the planner, the electricity grid or an DR aggregator, is to select a small set of agents to prepare for DR ahead of time and set the proper incentive schemes such that the target reduction is met with probability at least $\tau \in (0, 1)$. (M, τ) is the *system-wide reliability target*. We make a *deep market* assumption that there are enough agents in the economy such that if all are paid a high enough reward, the reliability target can be met. This holds for most real DR markets.

Two-period Mechanisms.

We consider demand response mechanisms that run over two periods with the the following timeline.

Period 0:
- Agents report information to the mechanism, with knowledge of their types.
- The mechanism determines for each selected agent i the period-one reward $r_i \geq 0$ for reducing consumption and penalty $z_i \geq 0$ in case of non-response.
- With the knowledge of r_i and z_i, each selected agent decides whether to prepare for demand response.

Period 1:
- The opportunity costs for responding are realized, and each agent decides whether or not to do so based on r_i, z_i and the realized value of V_i.
- For each selected agent i, the mechanism pays r_i upon response, and charges z_i, otherwise.

We call the pair of action-contingent payments (r_i, z_i) a *payment schedule* for demand response. Note that the mechanism is unable to observe selected agents' choices on preparation or their realized opportunity costs for reducing.

A demand-response mechanism is *dominant strategy incentive compatible* (DSIC) if truthful reporting maximizes each agent's expected utility regardless of the reports of other agents, and conditioned on the agent making rational decisions (see Section 2.1). A demand-response mechanism is *individually rational* (IR) if each agents' expected utility for (truthful) participation is non-negative. Informally, a DSIC mechanism is *truthful*, and we can say that an IR mechanism ensures that agents will choose to participate.

Reward and Penalty Bidding.

Consider agent i facing a fixed penalty for non-response. If her reward for response if zero, she loses in expectation thus will be unwilling to accept the payment schedule. However, if she is offered a million dollars for responding, there is no reason to reject. Intuitively, for any penalty, there is a minimum acceptable reward and similarly, fixing any reward, there would be a maximum acceptable penalty, for the agent to be willing to accept the DR payment schedule.

We now informally state the reward-bidding mechanism that we design in this paper, and also the penalty-bidding mechanism [13]:

Definition (Reward-bidding - informal). Fixing a uniform penalty z, the reward-bidding mechanism selects agents in increasing order of their minimum acceptable rewards until the reliability target is met, and pays each agent the highest minimum acceptable reward that she can claim to still be selected.

Definition (Penalty-bidding - informal). Fixing a uniform reward r, the penalty-bidding mechanism selects agents in decreasing order of their maximum acceptable penalties until the reliability target is met, and pays each agent the lowest maximum acceptable penalty that she can report to still be selected.

Fixing one of r and z is essential for selecting more reliable agents, computing critical payments for truthful information elicitation, and incentivizing higher response probabilities. These are not easily achievable in the general two-dimensional payment space where both r and z depend on agents' reports. We defer the detailed discussions to a full version of this paper.

We now proceed with the analysis of agents' rational decisions, expected utilities, minimum acceptable rewards and reliability in the following section.

2.1 Agents' Decisions, Utilities and Reliability

We first analyze a selected agent' rational decisions on preparation and response when she faces a DR payment schedule (r_i, z_i). Consider the following cases:

1. If the agent does not prepare, she is unable to respond and will be charged the penalty, thus her utility is $-z_i$;
2. If the agent does prepare at a cost of c_i and decides to respond, she gets paid reward r_i but incurs an opportunity cost of V_i, thus her utility is $r_i - V_i - c_i$;
3. If the agent did prepare but decides not to respond, she will be charged the penalty thus her utility is $-z_i - c_i$.

We can see that conditioned on preparation, the utility-maximizing decision in period one would be to respond if and
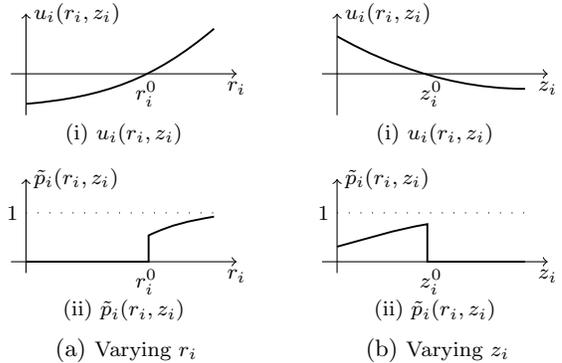
Figure 1: **Expected utility and effective reliability as functions of the reward r_i and the penalty z_i.**

only if (breaking ties in favor of responding) $r_i - V_i - c_i \geq -z_i - c_i \Leftrightarrow V_i \leq r_i + z_i$. Define the reliability of this agent given the payment schedule as the probability with which the agent responds, we have

$$p_i(r_i, z_i) \triangleq \mathbb{P}\left[V_i \leq r_i + z_i\right]. \tag{1}$$

Intuitively, a prepared agent responds only if the opportunity cost is small in comparison with the reward and penalty, and a higher reward or a higher penalty may increase the probability with which the agent responds. The expected utility of a prepared agent at the end of period zero is:

$$u_i(r_i, z_i) = \mathbb{E}\left[(r_i - V_i) \cdot \mathbb{1}\{V_i \leq r_i + z_i\}\right] \\ - z_i \cdot \mathbb{P}\left[V_i > r_i + z_i\right] - c_i, \tag{2}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. Fixing z_i, the expected utilities as a function of r_i are as illustrated in Figure 1(a)(i).

Minimum Acceptable Rewards.

The following lemma states useful properties of the expected utility function. The proofs are straightforward and thus omitted due to the space limit.

Lemma 1. Fixing $z_i \geq 0$, the expected utility function $u_i(r_i, z_i)$ satisfies:

1. $u_i(0, z_i) = -\mathbb{E}\left[V_i \cdot \mathbb{1}\{V_i \leq z_i\}\right] - z_i \mathbb{P}\left[V_i > z_i\right] - c_i < 0$.

2. $\lim_{r_i \to +\infty} u_i(r_i, z_i) = +\infty$.

3. $\frac{\partial}{\partial r_i} u_i(r_i, z_i) = \mathbb{P}\left[V_i \leq r_i + z_i\right] = p_i(r_i, z_i)$.

4. $u_i(r_i, z_i)$ is monotonically increasing and convex in r_i.

5. There exists a unique zero-crossing $r_i^0(z_i)$ s.t. $u_i(r_i^0, z_i) = 0$, see Figure 1(a)(i).

Intuitively, Lemma 1 shows that if an agent is charged a fixed penalty but paid no reward, her expected utility from preparing for DR is negative. As the reward increases, her expected utility continuously increases and crosses zero at some point $r_i^0(z_i)$. This is the *minimum acceptable reward* that the agent needs to be paid for her to be willing to prepare for DR and also pay a penalty of z_i for non-response. Technically, $r_i^0(z_i)$ is a function of z_i but we omit the argument when it is obvious from the context.

With fixed reward r_i, we can prove parallel properties of $u_i(r_i, z_i)$ as a function of the penalty z_i (see Figure 1(b)(i)) but omit the formal statements due to space limit. $u_i(r_i, z_i)$
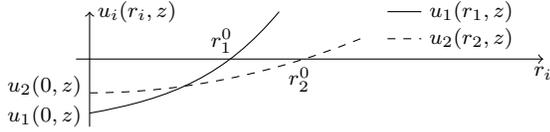
Figure 2: Expected utilities as functions of r_i for agents in Example 1.

is continuously decreasing and convex in z_i, has partial derivative $\frac{\partial}{\partial z_i} u_i(r_i, z_i) = p_i(r_i, z_i) - 1$, and has a unique zero-crossing $z_i^0(r_i)$ representing the agent's *maximum acceptable penalty* given reward r_i.

Preparation Decisions and Effective Reliability.

Fixing $z_i > 0$, if the agent faces a reward smaller than her minimum acceptable reward $r_i < r_i^0(z_i)$ (or equivalently, if an agent faces penalty $z_i > z_i^0(r_i)$ for some reward r_i), her expected utility, preparing or not, would be negative. Thus an agent offered such a payment schedule would not accept and never reduce consumption. When agent i faces $r_i \geq r_i^0(z_i)$ (or equivalently $z_i \leq z_i^0(r_i)$), her expected utility from preparing is non-negative, thus she will accept the payment schedule and choose in period zero to prepare.

Let $X_i(r_i, z_i)$ be a random variable indicating the number of units reduced by agent i, if she is offered the payment schedule (r_i, z_i). $X_i(r_i, z_i)$ is Bernoulli distributed

$$X_i(r_i, z_i) \sim \text{Bernoulli}(\tilde{p}_i(r_i, z_i)) \tag{3}$$

with parameter

$$\tilde{p}_i(r_i, z_i) \triangleq p_i(r_i, z_i) \cdot \mathbb{1}\{r_i \geq r_i^0(z)\}, \tag{4}$$

since she reduces consumption by one unit with probability $p_i(r_i, z_i)$ (see (1)) if and only if she accepts the payment schedule and prepares, which happens when $r_i \geq r_i^0(z)$. We call $\tilde{p}_i(r_i, z_i)$ the *effective reliability* of agent i if offered the payment schedule (r_i, z_i). $\tilde{p}_i(r_i, z_i)$ as a function of the reward r_i and penalty z_i is illustrated in Figure 1(ii).

An important observation from part 3 of Lemma 1 is that $p_i(r_i, z_i)$ relates to the partial derivatives of $u_i(r_i, z_i)$: the more reliable an agent is, the more likely that the agent is going to be paid the reward (not to pay the penalty), thus the faster $u_i(r_i, z_i)$ increases as the r_i increases (the slower $u_i(r_i, z_i)$ decreases as the z_i increases). Thus, an agent's effective reliability is fully determined by her expected utility $u_i(r_i, z_i)$, and $u_i(r_i, z_i)$ fully characterized the parts of an agent's type that is relevant to the DR problem.

Before proceeding to the mechanisms, we look at an example of two agents with uniformly distributed costs.

Example 1. Consider an economy with two agents whose opportunity costs are uniformly distributed: $V_1 \sim \text{U}[0, 8]$, $V_2 \sim \text{U}[0, 20]$ and let the preparation costs be $c_1 = 2, c_2 = 1$. Fixing the penalty $z = 1$ for both agents, the expected utilities computed according to (2) are as illustrated in Figure 2. Solving $u_i(r_i, z) = 0$ we get the minimum acceptable rewards for the two agents: $r_1^0(z) = 5.93$ and $r_2^0(z) = 7.94$

From the distributions of V_i of the two agents, we know that for any common reward and penalty the probability that agent 1 responds is higher. This corresponds to the steeper slope of $u_1(r_1, z)$. In general, agents with smaller minimum acceptable rewards are more likely to have a steeper slope, which corresponds to higher reliability.

3. THE REWARD-BIDDING MECHANISM

We now design a truthful and reliable mechanism for demand response, the *reward-bidding mechanism*, which fixes a uniform penalty for non-response and selects agents in increasing order of their minimum acceptable rewards. Note that reward-bidding is a direct-revelation mechanism, where an agent's critical reward payment is determined using not only the minimum acceptable rewards but also the reliability information reported by the rest of the agents.

We first provide notations. Consider a post-price mechanism where every agent is offered the same payment schedule (r, z). The random variable for the total reduction by all agents given (r, z) is $\sum_{i \in N} X_i(r, z)$, where $X_i(r, z)$ is the number of units reduced by agent i if offered (r, z), as defined in (3). We know from the deep market assumption and the monotonicity of the effective reliability $\tilde{p}_i(r_i, z_i)$ that for any fixed z, there exists a minimum uniform reward $r^{(N)}(z)$ such that the reliability target (M, τ) is met. Formally

$$r^{(N)}(z) \triangleq \min \left\{ r \in \mathbb{R}_+ \text{ s.t. } \mathbb{P}\left[\sum_{i \in N} X_i(r, z) \geq M\right] \geq \tau \right\}. \tag{5}$$

Similarly, for each agent i, we define the minimum sub-economy uniform reward $r^{(N \setminus \{i\})}(z)$ as the minimum amount to offer to all agents but i to achieve the reliability target: $r^{(N \setminus \{i\})}(z) \triangleq \min\{r \in \mathbb{R}_+ \text{ s.t. } \mathbb{P}\left[\sum_{j \in N \setminus \{i\}} X_j(r, z) \geq M\right] \geq \tau\}$. Note that both $r^{(N)}(z)$ and $r^{(N \setminus \{i\})}(z)$ depend on (M, τ), but we omit the arguments when it's clear from the context.

Definition 1. (Reward-Bidding Mechanism with Penalty z) The reward bidding mechanism collects reported type profile $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_n)$, computes for each agent the minimum acceptable reward $\hat{r}_i^0(z)$, and for reliability target (M, τ) the minimum uniform reward $r^{(N)}(z)$ given z, and the minimum sub-economy rewards $r^{(N \setminus \{i\})}(z)$. Then:

- Selection rule (period zero): select all agents that accept the minimum uniform reward $r^{(N)}(z)$, i.e. $x_i(\hat{\theta}) = 1$ if $\hat{r}_i^0(z) \leq r^{(N)}(z)$ and $x_i(\hat{\theta}) = 0$, otherwise.
- Payment rule (evaluated in period zero, payments made in period one): for selected agents, pay reward $r_i(\hat{\theta}) = r^{(N \setminus \{i\})}$ upon demand reduction and charge penalty $z_i(\hat{\theta}) = z$ for non-response. No payment to or from unselected agents.

We now examine the outcome of the reward-bidding mechanism for the economy introduced in Example 1 to show how the reward-bidding mechanism works.

Example 1. (continued) Consider a reliability target $M = 1, \tau = 0.9$ and assume agents report truthfully to the reward-bidding mechanism with penalty $z = 1$. If agent 1 is offered reward $r = 6.2 > r_1^0(z)$, which agent 2 is unwilling to accept, agent 1 accepts the payment schedule, prepares, and reduces with probability $\tilde{p}_1(6.2, 1) = 7.2/8 = 0.9$ and meets the reliability target. Thus $r^{(N)}(z) = 6.2$ and agent 1 is the only selected agent.

In the sub-economy $N \setminus \{1\}$, we can compute that agent 2 needs to be paid at least $r^{N \setminus \{1\}}(z) = 17$ to satisfy the reliability constraint thus agent 1's reward determined by the mechanism would be $r_1 = r^{N \setminus \{1\}}(z) = 17$. With $r_1 = 17$ and $z_1 = 2$ agent 1 actually always reduces consumption thus $M = 1$ is achieved with probability one.

THEOREM 1. *The reward-bidding mechanism is DSIC, IR and always satisfies the reliability target.*

PROOF. We first prove DSIC and IR. Fix an agent i. For all possible reports $\hat{\theta}_i$ that i can make, there are two possible outcomes: to be selected, face payment $(r^{N\setminus\{i\}}, z)$, or not to be selected and face payment zero. Since z is fixed and $r^{N\setminus\{i\}}$ depends only on the reports from the rest of the agents, all payments are independent to i's own report (i.e. agent-independence). To prove DSIC, we only need to show that the mechanism chooses the better outcome between the two for all agents (i.e. agent-maximization, see [14]).

Observe that $r^{N\setminus\{i\}}(z) \geq r^{(N)}(z)$ for all i, since for any r, $\mathbb{P}\left[\sum_{j\in N} X_j(r,z) \geq M\right] \geq \mathbb{P}\left[\sum_{j\in N\setminus\{i\}} X_j(r,z) \geq M\right]$ always holds. For i s.t. $r_i^0(z) \leq r^{(N)}(z)$, the expected utility from the payment schedule $(r^{N\setminus\{i\}}(z), z)$ is therefore non-negative thus getting selected is agent-maximizing. For i s.t. $r_i^0(z) > r^{(N)}(z)$, $r^{N\setminus\{i\}}(z) = r^{(N)}(z)$ holds, since agent i does not accept $(r^{(N)}(z), z)$, thus $\mathbb{P}\left[\sum_{j\in N} X_j(r^{(N)}, z) \geq M\right] = \mathbb{P}\left[\sum_{j\neq i} X_j(r^{(N)}, z) \geq M\right]$. Her expected utility from being selected and face $(r^{N\setminus\{i\}}(z), z)$ is negative, therefore not being selected and getting utility zero is agent-maximizing for her. IR also follows, since all agents gets at least the expected utility of not being selected, which is zero.

What is left to prove is that the mechanism always guarantees the reliability target. This is straightforward, observing that 1) for i s.t. $x_i(\hat{\theta}) = 0$, $r_i^0(z) > r^{(N)}$ thus $\tilde{p}_i(r^{(N)}, z) = 0$, 2) $\mathbb{P}\left[\sum_{i\in N} X_i(r^{(N)}, z) \geq M\right] = \mathbb{P}\left[\sum_{i\in S} X_i(r^{(N)}, z) \geq M\right]$ where $S = \{i \in N \text{ s.t. } x_i(\hat{\theta}) = 1\}$ is the set of all selected agents, 3) $\tilde{p}_i(p_i, z_i)$ is increasing in r_i thus so is the global reliability $\mathbb{P}\left[\sum_{i\in S} X_i(r_i, z) \geq M\right]$ and finally 4) $r_i = r^{(N\setminus\{i\})} \geq r^{(N)}$ for $i \in S$. Thus the probability of achieving the reliability target M can be bounded by:

$$\mathbb{P}\left[\sum_{i\in S} X_i(r_i, z) \geq M\right] = \mathbb{P}\left[\sum_{i\in S} X_i(r^{(N\setminus\{i\})}, z) \geq M\right]$$
$$\geq \mathbb{P}\left[\sum_{i\in S} X_i(r^{(N)}, z) \geq M\right] = \mathbb{P}\left[\sum_{i\in N} X_i(r^{(N)}, z) \geq M\right] \geq \tau.$$

This completes the proof of the theorem. $\square$

3.1 On Penalty Bidding

We now provide some intuition on why penalty-bidding [13] does not generalize to achieve the reliability target in a truthful manner for the uncertain cost scenario. From the above discussions, we know that the reward-bidding mechanism with penalty z selects the smallest set of agents necessary to satisfy the reliability target, in the case that all agents are offered the same penalty z and reward r. Each selected agent is then rewarded the highest minimum acceptable reward that she can report and still be selected—assuming agent i reports some type $\hat{\theta}_i$ such that $r_i^0(z) > r^{(N\setminus\{i\})}$, the $r^{(N)}$ computed based on the new reports becomes $r^{(N\setminus\{i\})}$ and agent i is therefore no longer selected.

What is crucial is that the probability of meeting the target is monotone in the varying part of the payment schedule: for uncertain costs, fixing z_i, and increasing r_i, the effective reliability $\tilde{p}_i(z_i, r_i)$ weakly increases for all agent types. But whereas this is the case for fixed reward and varying penalty under the simple fixed cost model of [13] (i.e. the (v_i, p_i, c_i)

model, fixing r and decreasing z more agents opt-in and non of them becomes less-reliable), it is not the case for penalty-bidding under the uncertain costs model. As is illustrated in Figure 1(b), the effective reliability $\tilde{p}_i(r_i, z_i)$ is first increasing in z_i, however, once z_i exceeds $z_i^0(r)$, the agent no longer accepts the DR payment schedule and the effective reliability drops to zero.

To get a penalty-bidding mechanism to satisfy the global reliability constraint without selecting too many agents, we would need to set the penalty z_i for a selected agent i to be high enough such that the effective reliability is high, but low enough so the payment schedule would not be rejected. This range cannot be easily determined without using agent i's own report, but this would, in turn, violate agent-independence and lose incentive compatibility. In contrast, for reward-bidding where there is no non-monotonicity in the effective reliability. Thus, we only need to guarantee a large enough set of agents are offered high enough rewards, and this can be achieved without using agent i's report to determine her own payments.

3.2 Computation of Reliability and Payments

We now briefly discuss the evaluation of the reliability and the computation of the minimum rewards in (5). Let S be the set of agents s.t. $\tilde{p}_i(r, z) > 0$. The total reduction $\sum_{i\in N} X_i(r, z) = \sum_{i\in S} X_i(r, z)$ is a Poisson-binomial distributed random variable with CDF: $\mathbb{P}\left[\sum_{i\in S} X_i(r, z) \leq k\right] = \sum_{\ell=0}^{k} \sum_{A\in\mathcal{S}_\ell} \prod_{i\in A} \tilde{p}_i(r, z) \prod_{j\in A^c}(1-\tilde{p}_j(r, z))$, where $\mathcal{S}_\ell$ is the set of all subsets of S of cardinality ℓ, and $A^c = S\setminus A$ [6]. We refer readers to [13] for polynomial algorithms for the exact evaluation of Poisson-binomial CDF.

Upper bounds of $r^{(N)}(z)$ and also $r^{(N-\setminus\{i\})}(z)$ can be computed to arbitrary precision by doing a binary search, starting with some very small and very large r. The reliability target is always achieved, and this approximation does not affect the incentives of the agents, since though the computation is not exact, the approximation process is still independent to agent i's own report.

4. MULTI-UNIT GENERALIZATION

The reward-bidding mechanism applies when agents can reduce multiple but fixed number of units of consumption. We generalize agents' type model for the scenario where agents may prefer to reduce a varying amount of consumption depending on the realized values and the payments, and generalize our mechanism to truthfully achieve the reliability target using a linear incentive scheme. The preparation cost and multiple-levels of preparation are not modeled for simplicity of notation, however, the model can be generalized without requiring modifications of the mechanism.

Uncertain Value Functions.

In order to analyze agents' decisions on reducing a varying amount of consumption, we need to consider agents' values for consuming different quantities. Let Ω_i be the set of possible world states of agent i, e.g. the set of all possible orders on cakes for agent i which is a bakery. We assume Ω_i is finite for all i for the simplicity of notation. A world state $\omega_i \in \Omega_i$ is realized with probability $f_i(\omega_i) \in (0, 1)$, and when this is the case, the value agent i derives from consuming $q \in \mathbb{R}_+$ units of electricity is $v_i^{(\omega_i)}(q)$.

We assume that for all $i \in N$, $\omega_i \in \Omega_i$ and $q \geq 0$, $v_i^{(\omega_i)}(q)$

is (A1) weakly increasing, (A2) right continuous, and (A3) bounded from above by some constant $W > 0$ and by $q \cdot T$ for some $T > 0$. Intuitively, (A1) means excessive electricity can be burnt free of cost; (A2) allows the value functions to be discontinuous, (e.g. the agent gets some value only if she can turn on a machine which burns at least q units of electricity), and (A3) prevents agents from being willing to consume infinite amounts of electricity or to pay infinite prices for each unit of electricity.

The set of value functions and the distribution of world states $\theta_i = \{v_i^{(\omega_i)}(q), \ f_i(\omega_i)\}_{\omega_i \in \Omega_i}$ determines an agent's type and is agent i's private information. Each agent knows her type in period zero, but the actual world state and the value of consuming electricity, is not realized until period one. We assume that the realizations of ω_i are independent among agents and cannot be observed by the planner.

Fix the price of each unit of electricity as $t > 0$. In period one, when the realized value function for agent i is $v_i^{(\omega_i)}(q)$, the utility of agent i for consuming q units of electricity is $v_i^{(\omega_i)}(q) - qt$, thus the optimal consumption decision is period one is $q_i^{(\omega_i)}(t) \triangleq \arg\max_{q \in \mathbb{R}_+} v_i^{(\omega_i)}(q) - qt$. The distribution of ω_i induces a distribution of consumption by agent i. Let $Q_i(t)$ be the random variable indicating the number of units consumed by agent i when the price of electricity is t. We know that $Q_i(t)$ takes value $q_i^{(\omega_i)}(t)$ with probability $f_i(\omega_i)$. Let $q_i(t)$ be the expected value of $Q_i(t)$ and $\sigma_i(t)$ be the standard deviation of $Q_i(t)$. We assume that the grid knows $q_i(t)$ and $\sigma_i(t)$ from historical data.

Linear Incentive Payment Schedules.

Consider a linear payment schedule (t, r, z), where electricity costs $t + z$ per unit, however, an agent is paid r per unit if her consumption is below $q_i(t)$. Similar to the above discussion, for any agent i with realized world state ω_i facing the payment schedule (t, r, z), there is an optimal amount of consumption (denoted $q_i^{(\omega_i)}(t, r, z)$) that gives agent i the highest utility $u_i^{(\omega_i)}(t, r, z, q_i^{(\omega_i)}(t, r, z))$. By consuming this amount for every realized world state, the expected utility agent i gets under the linear payment schedule is $u_i(t, r, z) = \sum_{\omega_i \in \Omega_i} f_i(\omega_i) \cdot u_i^{(\omega_i)}(t, r, z, q_i^{(\omega_i)}(t, r, z))$.

Similar to Lemma 1, we can prove parallel properties of $q_i^{(\omega_i)}(t, r, z)$ and $u_i(t, r, z)$ that agents consume less energy as z and r increases, and the expected utilities are increasing in r and decreasing in z. Moreover, there exists a minimum acceptable reward $r_i^0(z)$ such that agent i is willing to accept the additional per-unit cost z instead of getting the standard price schedule $(t, 0, 0)$. Facing a DR payment schedule (t, r_i, z_i), an agent decides to take it iff. $r_i \geq r_i^0(z)$, thus the random variable indicating agent i's consumption $Q(t, r_i, z_i)$ is equal to $Q(t)$ at all times when $r_i < r_i^0$, but takes value $q_i^{(\omega_i)}(t, r_i, z_i)$ with probability $f_i(\omega_i)$ if $r_i \geq r_i^0(z)$.

4.1 Multi-Unit DR Mechanism

For a multi-unit DR mechanism that offers each agent the choice between a DR payment schedule (t, r_i, z_i) and the flat rate (t), to achieve the reliability target (M, τ), we need:

$$\mathbb{P}\left[\sum_{i \in N} Q_i(t, r_i, z_i) \leq \sum_{i \in N} q_i(t) - M\right] \geq \tau \qquad (6)$$

We now analyze the minimum uniform rewards $r^{(N)}(z)$ and the sub-economy minimum rewards $r^{(N\setminus\{i\})}(z)$ to meet

the target when the penalty z is fixed. Define $r^{(N)}(z)$ as the minimum r s.t. (6) can be met when $z_i = z$ for all i. However, defining $r^{(N\setminus\{i\})}(z)$ as the minimum r_i such that $\mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) \leq \sum_{j \neq i} q_j(t) - M\right] \geq \tau$, does not guarantee $\mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) + Q_i(t) \leq \sum_{j \neq i} q_j(t) + q_i(t) - M\right] \geq \tau$ because of the uncertainty in $Q_i(t)$.

In order to compute $r^{(N\setminus\{i\})}(z)$ independent of agent i's own reported information, we need to find the minimum reward r such that for all possible distributions of some random variable Y s.t. $\mathbb{E}[Y] = q_i(t)$ and $\mathrm{std}[Y] = \sigma_i(t)$, $\mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) + Y \leq \sum_{j \neq i} q_j(t) + q_i(t) - M\right] \geq \tau$ always holds. We can prove that such $r^{(N\setminus\{i\})}(z)$ is no smaller than $r^{(N)}(z)$ for all i. We now define the mechanism.

Definition 2. (Multi-unit DR mechanism with Penalty z) The multi-unit DR mechanism collects agents' types, computes $r_i^0(z)$, $r^{(N)}(z)$ and $r^{(N\setminus\{i\})}(z)$, and offers to each agent a flat rate (t) and a DR payment schedule (t, r_i, z) where $r_i = r^{(N\setminus\{i\})}(z)$. Given the offered contracts, each agent selects the her preferred contract, decides on the preparation effort, consumption level, and then the mechanisms pays rewards and charges penalties accordingly.

THEOREM 2. *The Multi-Unit DR Mechanism is DSIC, IR and always guarantees the reliability target.*

The proof of the theorem is similar to that of Theorem 1, and is omitted due to space limitations.

4.2 Computation of Threshold Payments

We now discuss the computation of the minimum sub-economy reward $r^{(N\setminus\{i\})}(z)$. Technically, we are looking for the minimum reward r such that

$$\min_{F_Y} \mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) + Y \leq \sum_{j \neq i} q_j(t) + q_i(t) - M\right] \geq \tau$$

s.t. $\mathbb{E}[Y] = q_i(t)$ and $\mathrm{std}[Y] = \sigma_i(t)$.

The exact computation is not easy, since we would need to analyze the distribution of summation of several random variables and solve a constrained optimization problem in the functional space (i.e. space of all valid distributions). However, we can bound the probability and apply Chebyshev's inequality [8] and show that for all $\epsilon \in \mathbb{R}$,

$$\mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) + Y \leq \sum_{j \neq i} q_j(t) + q_i(t) - M\right]$$

$$\geq \mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) \leq \sum_{j \neq i} q_j(t) - M - \epsilon\right] \cdot \mathbb{P}[Y \leq q_i(t) + \epsilon]$$

$$\geq \left(1 - \frac{\sigma_i^2(t)}{\epsilon^2}\right) \cdot \mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) \leq \sum_{j \neq i} q_j(t) - M - \epsilon\right].$$

By setting $\epsilon(\tau)$ s.t. $1 - \sigma_i^2(t)/\epsilon(\tau)^2 = \sqrt{\tau}$ and looking for minimum r s.t. $\mathbb{P}\left[\sum_{j \neq i} Q_j(t, r, z) \leq \sum_{j \neq i} q_j(t) - M - \epsilon(\tau)\right] \geq \sqrt{\tau}$, we find an upper bound $\bar{r}^{(N\setminus\{i\})}(z)$ of $r^{(N\setminus\{i\})}(z)$ s.t. the reliability constraint is guaranteed to be achieved by offering $(t, \bar{r}^{(N\setminus\{i\})}(z), z)$ to all agents other than i. We can

set r_i to be $\bar{r}^{(N \setminus \{i\})}(z)$ in the multi-unit DR mechanism, and know that the reliability constraint can always be met.

5. MULTIPLE EFFORT LEVELS

We can also allow for agents who can exert multiple levels of preparation effort (these affecting the distribution on period one values). We do this by reducing the multi-effort-level model to the single level model. In this way, the reward-bidding mechanism can be directly applied. This can be done for both the unit-response and multi-unit response scenarios, but because of space limitations we only illustrate the idea in the single-unit case.

Let K_i be the total number of levels of effort that agent i can choose to exert during preparation. If agent i prepares at level k at a cost of $c_i^{(k)}$, her opportunity cost would be distributed according to $F_i^{(k)}$. Agent i's type is described by $\theta_i = (\theta_i^{(1)}, \ldots, \theta_i^{K_i})$ where $\theta_i^{(k)} = (c_i^{(k)}, F_i^{(k)})$. This subsumes the multi-level discrete model, where each agent can prepare at cost $c_i^{(k)}$, which enables her to respond with probability $p_i^{(k)}$ at the cost of $v_i^{(k)}$ for $k = 1, \ldots, K$.

With the same analysis, we know that given agent i prepares at level k, her expected utility is of the form:

$$u_i^{(k)}(r_i, z_i) = \mathbb{E}_{V_i \sim F_i^{(k)}}\left[(r_i - V_i) \cdot \mathbb{1}\{V_i \leq r_i + z_i\}\right]$$
$$- z_i \cdot \mathbb{P}_{V_i \sim F_i^{(k)}}\left[V_i > r_i + z_i\right] - c_i^{(k)}.$$

Since each agent is informed of the payment schedule (r_i, z_i) at period zero, she will choose the preparation effort level that maximizes her expected utility at (r_i, z_i). The equivalent expected utility for agent i facing payment schedule (r_i, z_i) is therefore:

$$\bar{u}_i(r_i, z_i) \triangleq \max_{k=1,\ldots,K_i} u_i^{(k)}(r_i, z_i).$$

As the upper envelope of a set of increasing and convex functions, $\bar{u}_i(r_i, z_i)$ is also increasing and convex. A payment schedule (r_i, z_i) induces an optimal effort $k^* = \text{argmax}_{k=1,\ldots,K_i} u_i^{(k)}(r_i, z_i)$ and the reliability still corresponds to the slope of $\bar{u}_i(r_i, z_i)$:

$$p_i(r_i, z_i) = F_i^{(k^*)}(r_i + z_i) = \frac{\partial}{\partial r_i} u_i^{(k^*)}(r_i, z_i) = \frac{\partial}{\partial r_i} \bar{u}_i(r_i, z_i).$$

since $\bar{u}_i(r_i, z_i) = u_i^{(k^*)}(r_i, z_i)$ in a small neighborhood of r_i.

This implies that the five properties that we proved in Lemma 1 also holds for $\bar{u}_i(r_i, z_i)$, which can be considered as the effective expected utility function of agent i and fully determines the effective reliability of the agent. Therefore, the multi-effort-level scenario can be reduced to the single-effort-level case by setting $u_i(r_i, z_i) = \bar{u}_i(r_i, z_i)$.

Example 2. Consider an agent with two possible effort levels. If she exerts the lower effort level at a cost of $c_i^{(1)} = 1$, she is able to respond with probability $p_i^{(1)} = 0.5$ at an opportunity cost of $v_i^{(1)} = 2$. If she exerts the higher level at a cost $c_i^{(2)} = 4$ the opportunity cost stays the same but her probability of being able to respond is boosted to $p_i^{(2)} = 0.9$.

The expected utilities corresponding to the two effort levels when penalty is fixed at $z = 1$, and the effective expected utility $\bar{u}_i(r_i, z)$ are as illustrated in Figure 3. We know from $\bar{u}_i(r_i, z)$ that with $r_i < (r_i^0)^{(1)}(z) = 5$, the agent does not
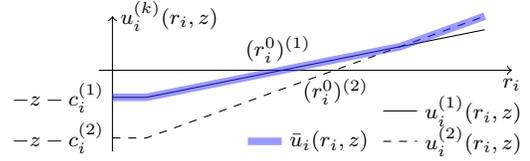
Figure 3: Expected utilities for different effort levels and the upper envelope $\bar{u}_i(r_i, z)$ for Example 2.

accept the payment schedule. For $5 \leq r_i < 8.5$, where $u_i^{(2)}(r_i, z)$ crosses $u_i^{(1)}(r_i, z)$, the agent takes the lower effort level, and therefore responds with probability $\tilde{p}_i(r_i, z) = 0.5$. For $r_i \geq 8.5$, the higher effort level is taken thus agent responds with probability $\tilde{p}_i(r_i, z) = 0.9$.

6. SIMULATION RESULTS

In this section we compare, through numerical simulation, the performance of the reward-bidding mechanism against the best possible outcome (i.e. the first best without private information) and a natural alternative mechanism, the spot auction, in which demand reduction is purchased from agents when needed.

6.1 Comparison with the First Best

We compare the number of agents selected by the reward-bidding mechanism with the "first best", which assumes that the mechanism knows the types of agents and therefore how reliable they would be given certain payments. Throughout this section, we consider agents whose types follow the exponential model. Each agent faces a fixed preparation cost c_i, and contingent on preparation, the opportunity cost V_i is exponentially distributed with parameter λ_i s.t. $\mathbb{E}[V_i] = \lambda_i^{-1}$.

Facing payment schedule (r_i, z_i), a prepared agent responds with probability $1 - e^{-\lambda_i(r_i + z_i)}$, thus the reliability of each agent can be boosted infinitely close to one, and the minimum number of agents needed in the first best would be equal to M. Let the total number of agents be $n = 500$ and the types be i.i.d. uniformly distributed: $c_i \sim \text{U}[0, 1]$ and $\lambda_i^{-1} \sim \text{U}[0, 2]$. We first assume that the grid charges a penalty $z = 1$ in the reward-bidding mechanism and would like to achieve a target reduction $M = 100$.

With τ varying from 0.9 to 0.999, the average number of selected agents over 1000 randomly generated economies is as shown in Figure 4(a). The horizontal axis "log risk" $-\log_{10}(1 - \tau)$ translate $\tau = 0.9$ to 1 and $\tau = 0.999$ to 3. We can see that more agents are selected when the probability target τ increases, and the mechanism does well in comparison with the first best.

Fixing $\tau = 0.98$, the number of agents selected by the reward bidding mechanism with different z is as shown in Figure 4(b). The number of agents selected decreases as z increases, since a higher penalty, and the resulting higher rewards (since agents have higher minimum acceptable rewards when z increases) improves the reliability of each of the selected agents.

6.2 Comparison With the Spot Auction

We now compare the reward-bidding mechanism with a benchmark proposed in [13], the *spot auction*. Without pre-selection in period zero of which agents should invest effort and prepare, the spot auction purchases demand reduction
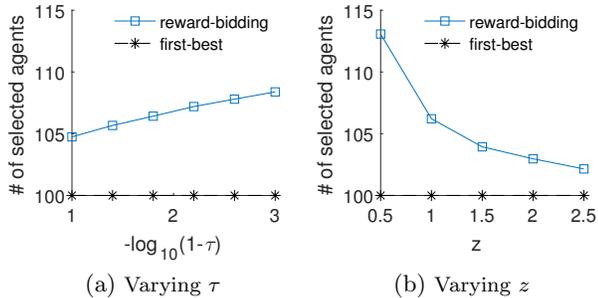
(a) Varying τ (b) Varying z

Figure 4: Comparison between the number of selected agents in reward-bidding and the first best.

(a) Average total cost   (b) Std of total cost

Figure 5: Comparison between the average and standard deviation of total costs.

from agents in period one in the event that DR is required using a simple $(M{+}1)^{\text{st}}$-price auction with a reserve price r, i.e. the reserve sets an upper bound on the reward payment.

Pure Nash Equilibrium on Preparation.

For an agent who prepared, it is a dominant strategy for her to bid in period one the realization of her opportunity cost V_i, since the preparation cost c_i is sunk and the $(M{+}1)^{\text{st}}$-price auction is truthful. What is not straightforward is to decide whether to prepare in period zero. For this, we study economies with exponential type agents where $c_i = c$ for all $i \in N$. Further, we assume these distribution types are known and study the performance of the spot auction under a (complete information) Nash equilibrium of the preparation game.

Assume w.l.o.g. $\lambda_1^{-1} \leq \lambda_2^{-1} \leq \cdots \leq \lambda_n^{-1}$, i.e. agents with smaller indexes face smaller opportunity costs in expectation. We prove through analyzing a threshold structure of the equilibrium that for any reserve $r \geq 0$, there exists a pure Nash equilibrium in which agents prepare if $i \leq m(r)$ and do not prepare otherwise, for some $0 \leq m(r) \leq n$. The full proof is left for an extended version of the paper.

To obtain the (asymmetric) pure-strategy preparation equilibrium, we compute via simulation (over one million realized cost profiles) for each reserve price r how many agents prepare in equilibrium and the resulting probability of achieving the reduction target. The higher the reserve price r, the more agents prepare, the higher the global reliability is and the higher the total payment made to the agents.

Experimental Results.

We now compare the equilibrium outcome of the spot auction with the truthful outcome of the reward-bidding mechanism. Consider a set of $n = 500$ with exponential model types, where the preparation cost $c_i = 2$, expected opportunity cost $\lambda_i^{-1} = i/100$, and the reduction target $M = 100$. For the reward-bidding mechanism, we simply set the penalty $z = 1$. For the spot auction, for each τ ranging from 0.9 to 0.999 we choose the minimum r such that there are enough agents preparing in equilibrium and the reliability target can be met.

The mean and standard deviation (std) of the total costs (reward payments minus collected penalties) of one million instances of agents' realized opportunity costs are shown in Figure 5. Despite the unfair comparison (dominant strategy for reward bidding vs. Nash equilibrium, optimized reserve
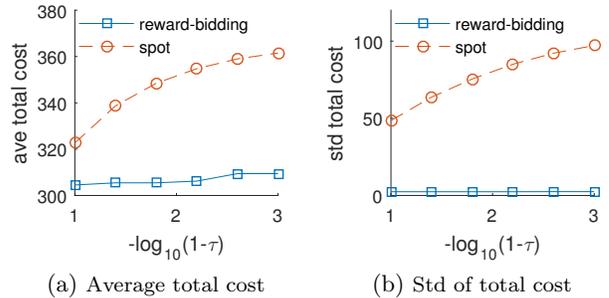
r for spot auction), the total cost under reward-bidding is lower than that of spot auction. Moreover, the standard deviation of the total costs under the spot auction is much higher. This is because under the spot auction, the total cost is low most of the times (when the $(M{+}1)^{\text{st}}$ bid is paid), however, with small probability where there are no more than M bids below r, r is paid to all agents, and this results in a huge variance in the total payments.

A high reserve r is needed in the spot auction because the number of agents preparing hardly increases as r increases, thus r has be large enough such that enough agents' bids fall below r and their reductions are purchased in order to meet the reliability target. When there are too many agents preparing, with high probability agents are getting paid only the $(M{+}1)^{\text{st}}$ bid instead of the high reserve price, and this may not be enough to cover the opportunity cost. As a comparison, the minimum reserve required to achieve $\tau = 0.999$ is $r = 5.76$ under which 101 agents prepare, and under reward-bidding with $z = 1$, there are 103 agents selected and average reward selected agents face (note that the critical rewards differ for different agents) is around 3.02. Both the reward and the penalty help with improving boosting the reliability of each individual agent.

7. CONCLUSIONS

We studied the generalized demand response problem where the design of contingent payments affect the probability of response, and where each agent may reduce multiple units of consumption. We design a new, truthful and reliable mechanism that selects a small number of agents to prepare and does so at low costs when compared to natural benchmarks.

In future work, we plan to understand whether it is possible to (i) design indirect mechanisms with good performance where there is no need for agents to communicate their full types, (ii) meet the reduction target with high probability without reducing too much beyond the target, (iii) optimize total welfare for both demand-side response and supply-side reserves at the same time, while retaining dominant-strategy equilibrium, and (iv) generalize the model and mechanism for demand side response over multiple periods of time.

REFERENCES

[1] C. Akasiadis and G. Chalkiadakis. Agent cooperatives for effective power consumption shifting. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

[2] C. Akasiadis, K. Panagidi, N. Panagiotou, P. Sernani, A. Morton, I. A. Vetsikas, L. Mavrouli, and K. Goutsias. Incentives for rescheduling residential electricity consumption to promote renewable energy usage. In *SAI Intelligent Systems Conference (IntelliSys), 2015*, pages 328–337. IEEE, 2015.

[3] D. Bergemann and J. Välimäki. The dynamic pivot mechanism. *Econometrica*, 78:771–789, 2010.

[4] R. Cavallo, D. C. Parkes, and S. Singh. Optimal coordinated learning among self-interested agents in the multi-armed bandit problem. In *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI'2006)*, Cambridge, MA, 2006.

[5] R. Cavallo, D. C. Parkes, and S. Singh. Efficient mechanisms with dynamic populations and dynamic types. Technical report, Harvard University, 2010.

[6] S. X. Chen and J. S. Liu. Statistical applications of the Poisson-Binomial and Conditional Bernoulli distributions. *Statistica Sinica*, 7:875–892, 1997.

[7] P. Courty and H. Li. Sequential screening. *The Review of Economic Studies*, 67(4):697–717, 2000.

[8] G. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford university press, 2001.

[9] R. Johari and J. N. Tsitsiklis. Parameterized supply function bidding: Equilibrium and efficiency. *Operations Research*, 59(5):1079–1089, 2011.

[10] R. Kota, G. Chalkiadakis, V. Robu, A. Rogers, and N. R. Jennings. Cooperatives for demand side management. In *Proc. 20th European Conference on AI (ECAI'12)*, pages 969–974, 2012.

[11] N. Li, L. Chen, and M. Dahleh. Demand response using linear supply function bidding. *IEEE Transactions on Smart Grid*, 6(4):1827–1838, 2015.

[12] H. Ma, R. Meir, D. C. Parkes, and J. Zou. Contingent payment mechanisms to maximize resource utilization. *arXiv preprint arXiv:1607.06511*, 2016.

[13] H. Ma, V. Robu, N. Li, and D. C. Parkes. Incentivizing Reliability in Demand-Side Response. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, 2016.

[14] N. Nisan. Introduction to mechanism design (for computer scientists). In N. N. et al., editor, *Algorithmic game theory*. Cambridge University Press, 2007.

[15] P. Palensky and D. Dietrich. Demand side management: Demand response, intelligent energy systems, and smart loads. *Industrial Informatics, IEEE Transactions on*, 7(3):381–388, 2011.

[16] D. C. Parkes and S. Singh. An MDP-based approach to Online Mechanism Design. In *Proc. 17th Annual Conf. on Neural Information Processing Systems (NIPS'03)*, 2003.

[17] PJM. Energy & ancillary services market operations. Technical report, PJM, AUGUST 2015.

[18] R. Porter, A. Ronen, Y. Shoham, and M. Tennenholtz. Fault tolerant mechanism design. *Artificial Intelligence*, 172(15):1783–1799, 2008.

[19] S. D. Ramchurn, P. Vytelingum, A. Rogers, and N. Jennings. Agent-based control for decentralised demand side management in the smart grid. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 5–12. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[20] V. Robu, G. Chalkiadakis, R. Kota, A. Rogers, and N. R. Jennings. Rewarding cooperative virtual power plant formation using scoring rules. *Energy*, 117, Part 1:19 – 28, 2016.

[21] V. Robu, E. H. Gerding, S. Stein, D. C. Parkes, A. Rogers, and N. R. Jennings. An online mechanism for multi-unit demand and its application to plug-in hybrid electric vehicle charging. *J. Artif. Intell. Res. (JAIR)*, 48:175–230, 2013.

[22] V. Robu, M. Vinyals, A. Rogers, and N. R. Jennings. Efficient buyer groups for prediction-of-use electricity tariffs. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pages 451–457, 2014.

[23] H. T. Rose, A. Rogers, and E. H. Gerding. A scoring rule-based mechanism for aggregate demand prediction in the smart grid. In *International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2012*, pages 661–668, 2012.

[24] W. Su, J. Wang, and J. Roh. Stochastic energy scheduling in microgrids with intermittent renewable energy resources. *Smart Grid, IEEE Transactions on*, 5(4):1876–1883, 2014.

[25] B. Zhang, R. Johari, and R. Rajagopal. Competition and coalition formation of renewable power producers. *Power Systems, IEEE Transactions on*, 30(3):1624–1632, 2015.