On Incentive Compatibility of Deadline Differentiated Pricing for Deferrable Demand

Eilyan Bitar[1] and Yunjian Xu[2]

***Abstract*— A large fraction of the total electric load is comprised of end-use devices whose demand is inherently deferrable in time. While this latent flexibility in demand can be leveraged to absorb variability in supply from renewable generation, the challenge lies in designing incentives to induce the desired response in demand. In the following, we study a novel forward market, where consumers consent to deferred service of pre-specified loads in exchange for a reduced per-unit price for energy. The longer a customer is willing to defer, the larger the reduction in price. The proposed deadline-differentiated forward contract provides a guarantee on the aggregate quantity to be delivered by a consumer-specified deadline. Under the earliest-deadline-first (EDF) scheduling policy, which is shown to be optimal for the supplier, we explicitly characterize differentiated prices yielding an efficient competitive equilibrium between supply and demand. We also show that such prices are incentive compatible (IC) in that every consumer would like to reveal her true deadline type to the supplier, provided that the other consumers are truth-telling.**

***Index Terms*— Incentive Compatibility, Demand Response, Renewable Energy, Pricing Mechanisms, Deadline Scheduling.**

I. INTRODUCTION

As the electric power industry transitions to a greater reliance on intermittent and distributed energy resources, there is an increasing need for flexible resources that can respond dynamically to weather impacts on wind and solar photovoltaic output. These renewable generation sources have limited controllability and production patterns that are intermittent and uncertain. This variability represents one the most important obstacles to the deep integration of renewable generation into the electricity grid. The current approach to renewable energy integration is to balance variability with dispatchable generation. This works at today's modest penetration levels, but it cannot scale, because of the projected increase in reserve generation required to balance the attendant variability in renewable supply [5]. If these increases are met with combustion fired generation, they will both be counterproductive to carbon emissions reductions and economically untenable.

As wind and solar energy penetration increases, how must the assimilation of this variable power evolve, so as to minimize these integration costs, while maximizing the net environmental benefit? Clearly, strategies which attenuate the increase in conventional reserve requirements will be an essential means to this end. One option is to harness the flexibility in consumption on the demand side. As such, significant benefits have been identified by the Federal Energy Regulatory Commission (FERC) [9] in unlocking the value in coordination of demand-side resources to address the growing need for firm, responsive resources to provide supply-demand balancing services (ancillary services) for the bulk power system.

A. Conventional approach

Clearly, there is an opportunity to transform the current operational paradigm, in which supply is tailored to follow demand, to one in which *demand is capable of reacting to variability in supply* – an approach which is generally referred to as demand response (DR) [1]. The challenge lies in *reliably extracting the desired response* from participating demand resources on time scales aligned with traditional bulk power balancing services.

Today, most demand response programs are largely limited to peak shaving applications, with the two most common economic paradigms for customer recruitment and control being: (1) *direct load control* whose capability is procured through a forward transaction (e.g. interruptible load contracts) and (2) *indirect load control* executed as a spot transaction (e.g. retail dynamic pricing). The performance of the latter is based on a *best efforts* basis with no firm guarantee as to how participating demand will adjust in response to signals (e.g. price, system alerts) from the system operator or utility. Moreover, the exposure of demand side resources to dynamic prices may lead to an increase in variability of load beyond nominal load patterns under conventional flat-rate tariffs. As the use of responsive demand shifts from reliability based utility run programs to market-based real-time balancing services, the criteria for performance guarantees become more stringent. In short, resource performance based on best efforts does not provide the level of assurance required to avoid the use of dispatchable generation to manage the electric system.

In the following sections, we propose a market framework that centers on the provisioning of deadline differentiated energy services to end-use customers, whose quality differentiation maps to flexibility in the family of feasible power profiles capable of satisfying said service. By offering a

Supported in part by NSF CNS-1239178, PSERC under sub-award S-52, US DoE under the CERTS initiative. Both authors contributed equally to the submitted work.

1. E. Bitar is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY. eyb5@cornell.edu

2. Y. Xu is with the Department of Engineering Systems and Design, Singapore University of Technology and Design, Singapore. xuyunjian@gmail.com

family of differentiated services – discounted according to quality – the coordinating entity implicitly purchases the right to manage in real-time the delivery of power to participating resources. In this way, it can align it's operational requirements with the vast heterogeneity in end-use customer needs.

B. A differentiated service approach

Flexibility in consumption can be interpreted as a *continuum of feasible power profiles* capable of satisfying the end-use function of a demand resource. A basic question, is how to design a market that enables a coordinating entity the ability to "extract this flexibility" for execution of real-time control applications – e.g., balancing variability in renewable supply?

One possibility resides in the construction of a market for *quality-differentiated electric power services*, where the price to a consumer for receiving a particular service is a monotonic function of the desired quality-of-service (QoS). Naturally, a reduction in QoS is accompanied by a reduction in price. For example, in the concrete setting of deferrable loads, deadline would be a natural specification of QoS. The longer a customer is willing to delay the receipt of a specified quantity of energy, the less that customer pays (per-unit) for said energy. Such markets would require a shift in how we think of electricity – not as an undifferentiated good, but rather as a set of differentiated services from which individual consumers can specify a QoS that best meets their electricity needs. Moreover, as a QoS specification maps directly to a set of feasible power profiles capable of servicing a load, the aggregator can imbed its extraction of flexibility from the demand-side in its delivery of differentiated services with a guaranteed QoS to participating consumers.

The general concept of service differentiation is not new [11]. Many have studied the problem of centrally coordinating the response of a collection of loads for load-following or regulation services – all while ensuring the satisfaction of a pre-specified QoS to individual resources [6], [8], [10], [12], [13], [14]. However, there has been little work in the way of designing market mechanisms that endogenously price the flexibility being offered by the demand side. Several classic [7], [15] and more recent [2] papers have explored the concept of *reliability-differentiated pricing* of interruptible electric power service, where the consumer takes on the risk of interruption in exchange for a reduction in the price for energy. Beyond the apparent issues with moral hazard and difficulty in auditing the delivered reliability of such services, the primary drawback of such an approach stems from the explicit transferal of quantity risk to the demand side, as it requires participating consumers to plan their consumption in the face of uncertain supply – a complex stochastic control problem.

With the aim of alleviating the aforementioned challenges, we analyze a novel market for *deadline differentiated energy services* (initially proposed in [3]), where consumers consent to deferred service of pre-specified loads in exchange for a reduced per-unit price for energy. The forward market for the *deadline differentiated energy service* is described as a three step process. Time is assumed discrete with periods indexed by $k = 0, 1, 2, \ldots$.

Step 1 (Pricing). Prior to period $k = 0$, the supplier announces a bundle of *deadline-differentiated prices*,

$$\mathbf{p} = (p_1, \ldots, p_K)^\top \in \mathbb{R}_+^K$$

The menu stipulates a price p_k (\$/kWh) for energy guaranteed delivery by period k.

Step 2 (Purchasing). Each consumer then purchases a bundle, $\mathbf{a} = (a_1, \ldots, a_K)^\top \in \mathbb{R}_+^K$ (kWh), of deadline-differentiated energy quantities, where a_k denotes the quantity of energy guaranteed delivery to the consumer by the deadline k. We denote the *aggregate demand bundle* (summed over all individual consumer bundles) by $\mathbf{x} = (x_1, \ldots, x_K)^\top \in \mathbb{R}_+^K$. Here, x_k denotes the aggregate quantity requiring delivery by period k.

Step 3 (Delivery). Finally, the supplier must deliver the requested aggregate demand bundle $\mathbf{x}$ subject to deadline constraints on delivery. The supplier is assumed to have *two sources of generation* from which he can service demand:

Intermittent generation. An intermittent supply modeled as a discrete time random process $\mathbf{s} = (s_0, s_1, \ldots, s_{K-1})$, with known distribution. Here, $s_k \in \mathcal{S} \subset \mathbb{R}_+$ (kWh) denotes the energy produced during period k. The intermittent supply is assumed to be zero marginal cost.

Firm generation. A firm supply with a constant marginal cost of production c_0 (\$/kWh).

C. Outline

Building on the basic market model proposed in [3], we present in Sections II-III, refined mathematical models for the demand and supply side, followed by an explicit characterization of incentive compatible prices and scheduling policy that jointly yield an efficient competitive equilibrium between the supplier and consumers in Section IV. In Section V, we present necessary and sufficient conditions for incentive compatibility with an explicit risk-reward interpretation. Finally, we close with brief concluding remarks and directions for future research in Section VI. All formal proofs are omitted due to space constraints.

II. Deferrable Demand Model

Consider now a utility model yielding a consumer preference ordering on deadlines. To capture the effect of consumption deferral on consumer utility, we assume that the *utility* derived from the consumption of a quantity $x \in \mathbb{R}_+$ is non-increasing in the delivery deadline k. Specifically, let $U^k(x) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ denote the utility derived from the consumption of a quantity x by deadline k. Each member of this family of *deadline-differentiated utility functions* is assumed to be *continuous, concave, monotone nondecreasing,* and satisfies:

$$U^1(x) \ \geq \ U^2(x) \ \geq \cdots \geq \ U^K(x) \ \geq \ 0 \quad \forall \, x \in \mathbb{R}_+,$$

where $U^k(0) = 0$ for all k. It follows that the *disutility* incurred by a consumer from deferring the consumption of a quantity x to deadline k is given by $L^k(x) = U^1(x) - U^k(x)$. We make the following simplifying assumptions.

Assumption 1 (Single deadline preference): We assume that each consumer has a *single deadline preference*. More specifically, a consumer with deadline preference k derives no disutility from deferring the consumption of a quantity x till deadline k and derives zero utility for consumption thereafter. Mathematically, this amounts to a family of utility functions satisfying:

$$U(x) := U^1(x) = \cdots = U^k(x) \quad \text{and} \quad U^j(x) = 0$$

for all $j > k$ and $x \in \mathbb{R}_+$.

Assumption 2 (Piecewise linear utility): The marginal utility of consumption is assumed constant at $R > 0$, up to a maximum demand of $q > 0$, after which it becomes zero. More formally, this corresponds to a piecewise linear utility function of the form $U(x) = R \cdot \min\{x, q\}$.

Definition 2.1 (Consumer type): The *type of consumer* i is a triple, $\theta_i = (k_i, R_i, q_i)$, consisting of her *deadline* k_i, *marginal utility* R_i, and *maximum demand* q_i. Let Θ denote the set of all consumer types, which is taken to be finite. ■

Clearly, Consumer i's utility function depends only on her type θ_i, and is defined as

$$U_{\theta_i}(x) \ = \ R_i \cdot \min\{x, q_i\}, \tag{1}$$

where x denotes the consumer's cumulative consumption by deadline k_i.

Assumption 3 (Non-atomic model): To study the aggregate behavior of a *large number of consumers*, we employ a non-atomic model describing a continuum of infinitesimal consumers, indexed by $i \in [0, 1]$.

Definition 2.2 (Type distribution): Let $\mu : \Theta \to [0, 1]$ denote the *distribution of consumer types* over the space Θ. For every $\theta \in \Theta$, there is a $\mu(\theta)$ fraction of consumers of type θ. It follows that $\sum_{\theta \in \Theta} \mu(\theta) = 1$. ■

Definition 2.3 (Consumer action): The *action* of a consumer i is a vector $\mathbf{a}_i = (a_{i,1}, \ldots, a_{i,K})^\top$, where $a_{i,k}$ denotes the amount of electricity that consumer i requests by deadline k. The maximum amount of electricity any consumer can request is $Q = \max_{\theta \in \Theta}\{q\}$. Hence, each consumer's *action space* is restricted to $\mathcal{A} = \left\{\mathbf{a} \in \mathbb{R}^K_+ \ : \ \sum_k a_k \leq Q\right\}$. ■

In the proceeding analysis, we will be concerned with identifying conditions on both the scheduling policy and pricing mechanism that lead to efficient allocations, while inducing consumers to truthfully reveal their underlying type. A truth-telling consumer is defined as follows.

Definition 2.4 (Truth-telling): Given a *deadline differentiated price* bundle $\mathbf{p} \in \mathbb{R}^K_+$, a consumer of type $\theta = (k, R, q)$ is *truth-telling*, if her action $\mathbf{a}$ satisfies

$$a_k = q \cdot \mathbf{1}_{\{R \geq p_k\}} \ \text{ and } \ a_j = 0, \ \forall \ j \neq k, \tag{2}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. In other words, she requests her surplus maximizing quantity $q \cdot \mathbf{1}_{\{R \geq p_k\}}$ to be delivered by her true deadline k. ■

If all consumers are truth-telling, the aggregate demand bundle, $\mathbf{x}$, depends only on the type distribution μ and the price bundle $\mathbf{p}$. It is given by

$$x_j(\mu, \mathbf{p}) = \sum_{\theta \in \Theta} q \cdot \mathbf{1}_{\{k=j, \ R \geq p_j\}} \mu(\theta) \tag{3}$$

for all $j = 1, \ldots, K$, where $\theta = (k, R, q)$.

A. Consumer surplus

We are now in a position to characterize the expected surplus derived by a consumer requesting any feasible demand bundle $\mathbf{a}' \in \mathcal{A}$ – not necessarily equal to her truth-telling bundle $\mathbf{a}$. Before formally characterizing an individual consumer's surplus in our model, we introduce some notation and assumptions. Given a truth-telling demand bundle $\mathbf{x}$, we let $\omega_{k,i}(\mathbf{x}, \mathbf{a}')$ be the random variable denoting the total amount of energy delivered to consumer i by stage k given her reported bundle $\mathbf{a}'$. This random variable, which (naturally) depends on the scheduling policy employed by the supplier, will be formally defined in Section III-A.2 (cf. Eq. (7)). Moreover, we have allowed the random supply $\omega_{k,i}(\mathbf{x}, \mathbf{a}')$ to depend explicitly on the consumer index i, as the supplier may employ a scheduling policy that is consumer index-dependent. The dependency of $\omega_{k,i}(\mathbf{x}, \mathbf{a}')$ on the scheduling policy is made precise in Section III.

Assumption 4: We assume each consumer $i \in [0, 1]$ has knowledge of the distribution on $\omega_{k,i}(\mathbf{x}, \mathbf{a}')$ for all $1 \leq k \leq K$ and $\mathbf{a}' \in \mathcal{A}$.

Consider a consumer i of type $\theta = (k, R, q)$ facing a nonincreasing price bundle $\mathbf{p}$. By monotonicity of prices, said consumer has no incentive to request any quantity before her true deadline k. In other words, $a'_t = 0$ for all $t < k$. Suppose that all consumers other than i are truth-telling. The expected surplus derived under a bundle $\mathbf{a}'$ is given by

$$V_{\theta,i}(\mathbf{x}, \mathbf{a}') \ = \ \mathbb{E}\left\{U_\theta\left(\omega_{k,i}(\mathbf{x}, \mathbf{a}')\right)\right\} \ - \ \sum_{t=1}^{K} p_t a'_t, \tag{4}$$

where expectation is taken over the random supply $\omega_{k,i}(\mathbf{x}, \mathbf{a}')$ to consumer i. We assume that the requested quantities are always supplied by their corresponding deadlines and the total quantity delivered never exceeds the consumer's total demand.

Assumption 5: Given the supplier's delivery commitments, we require that

$$\sum\nolimits_{t=1}^{k} a'_t \ \leq \ \omega_{k,i}(\mathbf{x}, \mathbf{a}') \ \leq \ \sum\nolimits_{t=1}^{K} a'_t, \quad \text{almost surely,}$$

for all $1 \le k \le K$ and $\mathbf{a}' \in \mathcal{A}$. □

Moreover, it follows that under truthful reporting of demand $\mathbf{a}' = \mathbf{a}$, the surplus derived by a type θ consumer simplifies to the deterministic quantity

$$V_{\theta,i}(\mathbf{x}, \mathbf{a}) \; = \; U_\theta(a_k) \; - \; p_k a_k,$$

where, $a_k = q \cdot \mathbf{1}_{\{R \ge p_k\}}$. Given a price bundle $\mathbf{p}$, a consumer i of type $\theta = (k, R, q)$ achieves maximum surplus for any action $\mathbf{a}^* \in \mathcal{A}$ satisfying:

$$\mathbf{a}^* \in \arg\max_{\mathbf{a} \in \mathcal{A}} \; V_{\theta,i}(\mathbf{x}, \mathbf{a}). \tag{5}$$

Notice that $\mathbf{a}^*$ – an explicit function of the price bundle $\mathbf{p}$ – represents consumer i's optimal demand curve under price taking behavior. For simplicity, we've assumed that the maximum surplus is always achieved.

III. Supply Model

In the following, we consider the role of a price-taking supplier, who aims to maximize its expected profit under a given price bundle. As such, his objectives are two-fold:

Scheduling. Causally schedule the allocation of the intermittent supply $\mathbf{s}$ across the deadline differentiated consumer classes, in order to *minimize the expected cost of firm supply* required to ensure demand satisfaction.

Pricing. Based on the optimal scheduling policy, determine an optimal supply curve that specifies the demand bundle it would like to serve at every price bundle $\mathbf{p} \in \mathbb{R}^K_+$. This amounts to computing a price bundle that induces a competitive equilibrium between supply and demand (cf. Definition 4.2).

Observe that the problem of computing an optimal supply curve (cf. Section III-B) (under a price-taking assumption) amounts to a solution of a *two-stage stochastic program*, whose expected recourse cost is the optimal value of a *constrained stochastic control problem* parameterized by the aggregate demand bundle $\mathbf{x}$.

A. Optimal scheduling

When considering the problem of scheduling, it's important to make a distinction between *intra-class* and *inter-class* scheduling policies. Loosely speaking, inter-class scheduling refers to the manner in which available supply is allocated across the deadline differentiated consumer classes, while intra-class scheduling refers to the manner in which available supply is allocated across consumers within a given deadline class. One can readily see that, given a fixed aggregate demand bundle $\mathbf{x}$, the supplier's expected profit depends only on the inter-class scheduling policy and is invariant under the family of feasible intra-class scheduling policies. However, as will be made apparent in Section V, the intra-class scheduling policy employed will have a direct effect on consumer purchase decisions inasmuch as it affects the distribution on each individual's random supply $\omega_{k,i}(\mathbf{x}, \mathbf{a})$.

1) Inter-class scheduling policies: We now characterize the optimal inter-class scheduling policy as a solution to a constrained stochastic optimal control problem. First, we define the system *state* at period k as the pair $(\mathbf{z}_k, s_k) \in \mathbb{R}^K_+ \times \mathcal{S}$, where the vector $\mathbf{z}_k$ denotes the residual demand requirement of the original aggregate demand bundle $\mathbf{x}$ after having been serviced in previous periods $0, 1, \cdots, k-1$.

Define as the *control input* the vectors $\mathbf{u}_k, \mathbf{v}_k \in \mathbb{R}^K_+$, which denote (element-wise in j) the amount of intermittent and firm supply allocated to demand class j at period k, respectively. Naturally then, the state of residual demand evolves according to the difference relation:

$$\mathbf{z}_{k+1} \; = \; \mathbf{z}_k - \mathbf{u}_k - \mathbf{v}_k, \tag{6}$$

$k = 0, \ldots, K-1$, where the process is initialized with $\mathbf{z}_0 = \mathbf{x}$ for $j = 1, \ldots, K$. The delivery deadline constraints manifest in a sequence of nested polytopes $\mathbb{R}^K_+ \supset \mathcal{Z}_1 \supseteq \mathcal{Z}_2 \supseteq \cdots \supseteq \mathcal{Z}_K = 0$ converging to the the origin, where the set $\mathcal{Z}_k$ characterizes the *feasible state space* at stage k. More precisely, we have

$$\mathcal{Z}_k \; = \; \{\mathbf{z} \in \mathbb{R}^K_+ \mid z^j = 0, \;\; \forall\, j \le k \; \text{ and } \; z^j \le x_j, \;\; \forall\, j > k\}.$$

In other words, the feasible state space is such that each demand class is fully serviced by its corresponding deadline. We define as the *feasible input space* at stage k the set of all inputs belonging to the set

$$\begin{aligned} \mathcal{U}_k(\mathbf{z}, s) \; = \; \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^K_+ \times \mathbb{R}^K_+ \mid \mathbf{1}^\top \mathbf{u} \le s \\ \text{and } \; \mathbf{z} - \mathbf{u} - \mathbf{v} \in \mathcal{Z}_{k+1}\}, \end{aligned}$$

which ensures one-step state feasibility and that the total allocation of renewable supply does not exceed availability at the current stage. In characterizing the feasible set of causal scheduling policies, we restrict our attention to policies with *Markovian information structure*, as opposed to allowing the control to depend on the entire history. One can show that this is without loss of optimality for the given problem structure, without explicitly requiring the stochastic process $\mathbf{s}$ to be Markovian. More precisely, we describe the scheduling decision at each stage k by the functions

$$\mathbf{u}_k = \boldsymbol{\mu}_k(\mathbf{z}, s) \;\; \text{and} \;\; \mathbf{v}_k = \boldsymbol{\nu}_k(\mathbf{z}, s),$$

where $\boldsymbol{\mu}_k : \mathcal{Z}_k \times \mathcal{S} \to \mathbb{R}^K_+$ and $\boldsymbol{\nu}_k : \mathcal{Z}_k \times \mathcal{S} \to \mathbb{R}^K_+$.

A *feasible scheduling policy* is any finite sequence of scheduling decision functions

$$\pi = (\boldsymbol{\mu}_0, \ldots, \boldsymbol{\mu}_{K-1}, \boldsymbol{\nu}_0, \ldots, \boldsymbol{\nu}_{K-1})$$

such that $(\boldsymbol{\mu}_k, \boldsymbol{\nu}_k)(\mathbf{z}, s) \in \mathcal{U}_k(\mathbf{z}, s)$ for all $(\mathbf{z}, s) \in \mathcal{Z}_k \times \mathcal{S}$ and $k = 0, \ldots, K-1$. We denote by $\Pi(\mathbf{x})$ the set of all feasible scheduling policies. Throughout the paper, we will suppress the explicit dependency of the feasible policy set on the initial demand bundle $\mathbf{x}$, when it's clear from the context.

2) Intra-class scheduling policies: Recall that an intra-class scheduling policy determines the allocation of available supply within each deadline-differentiated demand class, where the supply available to each demand class is determined by the inter-class policy $\pi \in \Pi$. Formally, we let $\boldsymbol{\lambda}_{k,i}(\mathbf{x}, \mathbf{a}_i) \in \mathbb{R}_+^K$ denote (element-wise in j) the amount of energy delivered (at period k) to consumer i so as to satisfy her demand $a_{i,j}$. We denote the intra-class scheduling policy by

$$\phi = \{(\boldsymbol{\lambda}_{0,i}, \ldots, \boldsymbol{\lambda}_{K-1,i}) \mid i \in [0,1]\},$$

where $\boldsymbol{\lambda}_{k,i} : \mathcal{X} \times \mathcal{A} \to \mathbb{R}_+^K$ for all $i \in [0,1]$ and $k = 0, \ldots, K-1$. Given an inter-class scheduling policy $\pi \in \Pi$, an intra-class scheduling policy ϕ is feasible if and only if it satisfies the following constraints (i) - (iv).

(i) The intra-class scheduling policy ϕ should not deliver any supply allocated to class j to consumers not belonging to class j. That is, for each demand class $j = 1, \ldots, K$, we have that $\lambda_{k,i}^j(\mathbf{x}, \mathbf{a}_i) = 0$ for every consumer i such that $a_{i,j} = 0$.

(ii) No supply can be delivered to consumers belonging to a demand class whose deadline has passed. That is, $\lambda_{k,i}^j(\mathbf{x}, \mathbf{a}_i) = 0$ for every $i \in [0,1]$, and every $1 \le j \le k \le K-1$.

(iii) The total supply allocated to demand class j at time period k must be full utilized, i.e.

$$\int_{[0,1]} \lambda_{k,i}^j(\mathbf{x}, \mathbf{a}_i)\, \varphi(di) \;=\; \mu_k^j + \nu_k^j, \quad \forall\ 0 \le k < j \le K,$$

where we use the Lebesgue integral[1] (with respect to Lebesgue measure φ defined on $[0,1]$), and μ_k^j and ν_k^j denote the amount of intermittent and firm supply, respectively, allocated to demand class j at period k according to the inter-class policy π.

(iv) Each consumer's individual delivery commitments must be met. For $k = 1, \ldots, K$, we have

$$\sum\nolimits_{t=1}^{k} a_{i,t} \;\le\; \omega_{k,i}(\mathbf{x}, \mathbf{a}_i) \;\le\; \sum\nolimits_{t=1}^{K} a_{i,t},$$

where the total energy delivered to consumer i by deadline k is given by

$$\omega_{k,i}(\mathbf{x}, \mathbf{a}_i) = \sum_{t=0}^{k-1} \sum_{j=1}^{K} \lambda_{t,i}^j(\mathbf{x}, \mathbf{a}_i). \tag{7}$$

Notice that for any feasible inter-class scheduling policy $\pi \in \Pi$, it is always possible to satisfy the above constraint.

We denote the set of all feasible intra-class scheduling policies by $\Phi(\pi)$, which are parameterized by a given inter-class policy $\pi \in \Pi$. It is important to note that the supplier's profit depends only on the inter-class scheduling policy. Namely, for any feasible inter-class policy π, all feasible intra-class scheduling policies $\phi \in \Phi(\pi)$ yield the supplier the same expected profit. This follows from the supplier's indifference to supply allocation between consumers within a given demand class. Therefore, in characterizing the optimal scheduling policy for the supplier, we restrict our attention to inter-class policies for the remainder of this section.

3) Supplier profit: We define the *expected profit* $J(\mathbf{x}, \mathbf{p}, \pi)$ derived by a supplier as the revenue derived from an aggregate demand bundle $\mathbf{x}$ less the expected cost of servicing said demand bundle under a feasible inter-class scheduling policy $\pi \in \Pi(\mathbf{x})$. More precisely, let

$$J(\mathbf{x}, \mathbf{p}, \pi) \;=\; \mathbf{p}^\top \mathbf{x} \;-\; Q(\mathbf{x}, \pi), \tag{8}$$

where Q denotes the expected cost of firm generation incurred servicing $\mathbf{x}$ under a feasible policy $\pi \in \Pi$.

$$Q(\mathbf{x}, \pi) \;=\; \mathbb{E} \sum_{k=0}^{K-1} g(\mathbf{v}_k^\pi), \tag{9}$$

where the stage cost is defined as $g(\mathbf{v}) = c_0 \mathbf{1}^\top \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}_+^K$. We write the state and control process as $\{\mathbf{z}_k^\pi\}, \{\mathbf{u}_k^\pi\}, \{\mathbf{v}_k^\pi\}$ to emphasize their dependence on the policy π. We wish to characterize scheduling policies that lead to a minimal expected cost of firm supply.

Definition 3.1: The policy $\pi^* \in \Pi(\mathbf{x})$ is *optimal* if

$$Q(\mathbf{x}, \pi^*) = \inf_{\pi \in \Pi(\mathbf{x})} \mathbb{E} \sum\nolimits_{k=0}^{K-1} g(\mathbf{v}_k^\pi) \tag{10}$$

subject to $\mathbf{z}_{k+1}^\pi = \mathbf{z}_k^\pi - \mathbf{u}_k^\pi - \mathbf{v}_k^\pi$ for $k = 0, \ldots, K-1$.

4) Optimal scheduling policy: We now present the optimal inter-class scheduling policy, π^*.

Theorem 3.2 (Earliest-Deadline-First): Given a demand bundle $\mathbf{x}$, the optimal scheduling policy π^* is given by:

$$\mu_k^{j,*}(\mathbf{z}, s) \;=\; \min\left\{ z^j,\ s - \sum_{i=1}^{j-1} \mu_k^{i,*}(\mathbf{z}, s) \right\}$$
$$\nu_k^{j,*}(\mathbf{z}, s) \;=\; \Big(z^j - \mu_k^{j,*}(\mathbf{z}, s) \Big) \cdot \mathbf{1}_{\{k=j-1\}}$$

for $j = 1, \ldots, K$, $k = 0, \ldots, K-1$, and $(\mathbf{z}, s) \in \mathcal{Z}_k \times \mathcal{S}$.

Qualitatively, the optimal scheduling policy is such that the intermittent supply s_k available at period k is allocated to those unsatisfied demand classes with *earliest-deadline-first* (EDF), while the firm supply is dispatched only as a last resort to ensure demand satisfaction. An interesting feature of the optimal inter-class policy derives from its proof of optimality. Namely, EDF scheduling performs as well as any non-causal policy in the metric of expected firm generation cost. Thus, having prescient information regarding the realization of the intermittent supply process $\mathbf{s}$ cannot not improve the suppliers expected profit beyond that achievable under the causal EDF scheduling.

B. Optimal Pricing

Given the EDF characterization of the optimal inter-class scheduling policy in Theorem 3.2, we are now in a position to characterize the supplier's optimal supply curve, under price taking behavior. To aid our analysis, we define a new *residual*

[1] Note that we have implicitly required here that $\lambda_{k,i}^j(\mathbf{x}, \mathbf{a}_i)$ is Lebesgue integrable with respect to i, over the interval $[0,1]$.

random process $\boldsymbol{\xi} = (\xi_0, \ldots, \xi_K) \in \mathbb{R}^{K+1}$, where a positive residual ($\xi_k > 0$) denotes the amount of intermittent supply leftover after servicing the demand class x_k, according to the EDF inter-class scheduling policy, by its deadline k. A negative residual ($\xi_k \leq 0$) denotes the amount by which the intermittent supply fell short or, equivalently, the quantity of firm supply required to ensure satisfaction of the demand class x_k. Recursively, we have

$$\xi_{k+1} \;=\; \xi_k^+ \;+\; s_k - x_{k+1} \tag{11}$$

for $k = 0, \ldots, K-1$, where $\xi_0 = 0$. Notice that ξ_k depends on both the bundle $\mathbf{x}$ and intermittent supply process $\mathbf{s}$. Using this newly defined process, we arrive at the following compact representation of the minimum expected cost of firm generation (under EDF).

Lemma 3.3 (Expected supplier profit under EDF): The expected profit criterion (8) derived under the optimal inter-class scheduling policy $\pi^* \in \Pi(\mathbf{x})$ (EDF) satisfies the following properties for all $\mathbf{x} \in \mathbb{R}_+^K$.

1) $J(\mathbf{x}, \mathbf{p}, \pi^*)$ is *continuous* and *concave* in $\mathbf{x}$.
2) $Q(\mathbf{x}, \pi^*)$ is *continuous* and *convex* in $\mathbf{x}$.
3) Given a residual process $\boldsymbol{\xi}$ induced by a pair $(\mathbf{x}, \mathbf{s})$, we have that $Q(\mathbf{x}, \pi^*)$ satisfies

$$Q(\mathbf{x}, \pi^*) \;=\; -\mathbb{E} \sum_{k=1}^{K} c_0 \cdot \min\{0, \xi_k\}.$$

Identifying a price bundle that induces a competitive equilibrium between supply and demand requires, first, a characterization of the supplier's optimal supply curve. Namely, given a price bundle $\mathbf{p}$, the supplier computes

$$\mathbf{x}^* \;\in\; \arg\max_{\mathbf{x} \in \mathbb{R}_+^K} \; J(\mathbf{x}, \mathbf{p}, \pi^*), \tag{12}$$

where $\mathbf{x}^*$ denotes the supplier's profit maximizing supply allocation – an explicit function of the price bundle $\mathbf{p}$. By concavity of $J(\mathbf{x}, \mathbf{p}, \pi^*)$, one can readily compute necessary and sufficient conditions for optimality, which are made precise in the following Theorem 3.4.

Theorem 3.4 (Optimal supply curve): An allocation, $\mathbf{x}$, is profit maximizing for a given price bundle, $\mathbf{p}$, if and only if

$$p_k \;=\; \zeta_k(\mathbf{x}), \qquad \text{for all} \quad k = 1, \ldots, K, \tag{13}$$

where $\zeta_k : \mathbb{R}_+^K \to \mathbb{R}_+$ is defined by Eq. (14).

In addition, it is straight forward to show that the optimal supply curve yields nonincreasing prices for any allocation $\mathbf{x} \in \mathbb{R}_+^K$. More precisely,

$$c_0 \;\geq\; \zeta_1(\mathbf{x}) \;\geq\; \zeta_2(\mathbf{x}) \;\geq \cdots \geq\; \zeta_K(\mathbf{x}).$$

This property of price monotonicity is consistent with our initial criterion for constructing this market system. Namely, the longer a customer is willing to defer her consumption in time, the less she is required to pay per unit of energy. In addition, the fact that no price exceeds the marginal cost of firm generation is reassuring, as it ensures that such a market for deferrable electric power service will always provide customers a discount on the firm rate.

IV. Market equilibrium and incentive compatibility

In this section, we show that the equilibrium price bundle in Eq. (14) is incentive compatible. Before providing the main result of this section, we will first introduce some definitions and assumptions that will be useful later.

Definition 4.1 (Incentive compatibility): Given a pair of inter-class $\pi \in \Pi$ and intra-class $\phi \in \Phi(\pi)$ scheduling policies, a price bundle $\mathbf{p}$ is *incentive compatible*, if truth-telling is every consumer's best response, provided that other consumers are truth-telling. Formally, a price bundle $\mathbf{p}$ is incentive compatible, if for all $i \in [0, 1]$ and $\mathbf{x} \in \mathcal{X}$,

$$V_{\theta,i}(\mathbf{x}, \mathbf{a}) \geq V_{\theta,i}(\mathbf{x}, \mathbf{a}'), \quad \text{for all} \;\; \mathbf{a}' \in \mathcal{A}, \quad \theta \in \Theta,$$

where $\mathbf{a}$ is the truth-telling action of a type θ consumer (2) and $\mathbf{x}$ is the truth-telling aggregate demand bundle (3). ■

Definition 4.2 (Competitive equilibrium): Given a distribution of consumer types μ, and a random process for intermittent generation, $\mathbf{s}$, a quantity-price pair $(\mathbf{x}, \mathbf{p})$ constitutes a competitive equilibrium, if there exist feasible inter-class $\pi \in \Pi$ and intra-class $\phi \in \Phi(\pi)$ scheduling policies, such that the following two conditions hold.

1) The price bundle $\mathbf{p}$ is incentive compatible under the scheduling polices (ϕ, π), in the sense of Definition 4.1.
2) The truth-telling aggregate demand bundle $\mathbf{x} = \mathbf{x}(\mu, \mathbf{p})$ (3) and inter-class scheduling policy π maximize the supplier's expected profit (8) under the price bundle $\mathbf{p}$. ■

Assumption 6: We assume that for every $(k, R, q) \in \Theta$, we have $R \geq c_0$.

Theorem 4.3 (Competitive equlibrium): The price schedule in Eq. (14) is incentive compatible for each consumer under the EDF scheduling policy $\pi^* \in \Pi$ and a corresponding feasible intra-class policy $\phi \in \Phi(\pi^*)$ employed by the supplier. It follows that such a price schedule, together with the truth-telling demand bundle it induces, constitute a competitive equilibrium.

We note that the result on incentive compatibility may fail to hold, if the constraint on marginal utility required in Assumption 6 is violated. In the case that $R \leq p_k \leq c_0$, it is easy to see that a consumer i of type (k, R, q) achieves a zero payoff if she is truth-telling; while on the other hand, she may be able to obtain a positive expected payoff by requesting a quantity q with a later deadline $\tilde{k} > k$, if, for example, $p_{\tilde{k}} = 0$ and the distribution on supply is such that she can receive a positive amount of supply by deadline k with positive probability. Moreover, as we will see in the following example, the consumer may still benefit from misreporting its deadline, even if $p_k \leq R \leq c_0$.

$$\zeta_k(\mathbf{x}) \;:=\; c_0 \left[\mathbb{P}(\xi_k \leq 0) \;+\; \sum_{t=k+1}^{K} \mathbb{P}(\xi_1 > 0, \ldots, \xi_{t-1} > 0, \xi_t \leq 0) \right], \qquad \text{for all} \quad k = 1, \ldots, K. \tag{14}$$

Example 1: Consider a scenario where $\mathbb{P}(\xi_{K-1} \leq 0) = 0.5$ and there exists a type $\theta = (K-1, R, q) \in \Theta$ such that $R = 0.6c_0$. Suppose that there is enough intermittent supply at the final period K such that $\mathbb{P}(\xi_K > 0) = 1$, which implies an optimal price of $p_K = 0$ for deadline K. It follows from (14) that $p_{K-1} = \mathbb{P}(\xi_{K-1} \leq 0)c_0 = 0.5c_0$, which yields a truth-telling type θ consumer a payoff of $0.1c_0 q$.

Suppose that whenever $\xi_{K-1} > 0$, the surplus is large enough to satisfy the terminal demand x_K; i.e., $\mathbb{P}(\omega_{K-1,i}(\mathbf{x}, \mathbf{a}') = q) = 0.5$ for every i, where $\mathbf{a}' = (0, \ldots, 0, q)$. By reporting a false bundle $\mathbf{a}'$, the consumer achieves a positive expected payoff of

$$\begin{aligned} &V_{\theta,i}(\mathbf{x}, \mathbf{a}') \\ &\quad = R\mathbb{P}(\xi_{K-1} > 0)\mathbb{E}\{\omega_{K-1,i}(\mathbf{x}, \mathbf{a}') \mid \xi_{K-1} > 0\} - 0 \\ &\quad = \; 0.5qR \;=\; 0.3c_0 q, \end{aligned}$$

which exceeds the payoff obtained from truth-telling. ■

Finally, we note that the result in Theorem 4.3 fails to hold for general concave utility functions. This is intuitive because a consumer with a highly concave utility function may prefer to report a false deadline $\tilde{k}$ that is later than her true deadline k, if she can obtain a fraction of her demand before stage k (but most of her utility) with high probability, at a much lower price $p_{\tilde{k}}$.

V. Incentive compatibility conditions

In this section, we refine the conditions on incentive compatibility for a special case in which the consumer's action space is restricted to be $\mathcal{A} = \{1, \ldots, K\} \times [0, Q]$. In other words, a consumer can only purchase a single energy quantity, $a \in [0, Q]$, and an associated deadline $\tilde{k} \in \{1, \ldots, K\}$ – as opposed to an entire consumption bundle. Consider consumer i of type $\theta = (k, R, q)$. With a slight abuse of notation, we let $\omega_{k,i}(\mathbf{x}, \tilde{k}, a)$ denote the energy delivered to consumer i by deadline k after having requested $(\tilde{k}, a)$. Clearly, the distribution on $\omega_{k,i}(\mathbf{x}, \tilde{k}, a)$ will depend on the intra-class scheduling policy employed by the supplier, which determines the allocation of electricity supply between consumers belonging to the same deadline class.

In the proceeding analysis, we consider *intra-class scheduling policies of the following form.* Namely, the maximum supply available to consumer i does not depend on her reported demand a. Under such a scheduling policy, we can characterize the energy delivered to consumer i reporting $(\tilde{k}, a)$ as

$$\omega_{k,i}(\mathbf{x}, \tilde{k}, a) = \min\{a, \; \rho_{k,i}(\mathbf{x}, \tilde{k})\}, \tag{15}$$

where the random variable $\rho_{k,i}(\mathbf{x}, \tilde{k})$ represents the maximum supply available to consumer i (who takes an action $(\mathbf{x}, \tilde{k})$) by deadline k. As an example, consider an intra-class policy that allocates the supply available to a particular demand class amongst consumers on a first-come first-serve basis - i.e. those consumers arriving earlier within a given class are given priority in terms of service.

Remark 1: We note that the following conditions for incentive compatibility do not depend explicitly on the particular form of inter-class scheduling policy (e.g, EDF) employed by the supplier.

Let $F_{k,i,\tilde{k},\mathbf{x}}(\cdot)$ denote the cumulative distribution function of the random variable $\rho_{k,i}(\mathbf{x}, \tilde{k})$. Its corresponding quantile function is defined by

$$F^{-1}_{k,i,\tilde{k},\mathbf{x}}(y) = \inf\left\{x : F_{k,i,\tilde{k},\mathbf{x}}(x) \geq y\right\}.$$

Assumption 7: We assume that for every $(k, i, \tilde{k}, \mathbf{x})$, $F_{k,i,\tilde{k},\mathbf{x}}(\cdot)$ is strictly increasing over $[0, C_{k,i,\tilde{k},\mathbf{x}}]$, where $C_{k,i,\tilde{k},\mathbf{x}}$ is the minimum positive constant such that $F_{k,i,\tilde{k},\mathbf{x}}(C_{k,i,\tilde{k},\mathbf{x}}) = 1$.

In Proposition 5.1, we derive a necessary and sufficient condition for incentive compatibility under an intra-class scheduling policy of the form (15). Before proceeding, we first introduce some notation that will be useful later. The *conditional value-at-risk* (CVaR) deviation measure $\mathcal{D}_c(\rho_{k,i}(\mathbf{x}, \tilde{k}))$ (cf. [4]) of a random variable $\rho_{k,i}(\mathbf{x}, \tilde{k})$ is defined as

$$\begin{aligned} &\mathcal{D}_c(\rho_{k,i}(\mathbf{x}, \tilde{k})) \\ &= \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k})\right\} - \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k}) \;\middle|\; \rho_{k,i}(\mathbf{x}, \tilde{k}) \leq F_{k,i,\tilde{k},\mathbf{x}}(c)\right\} \\ &= \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k})\right\} - \frac{1}{c}\int_0^{F^{-1}_{k,i,\tilde{k},\mathbf{x}}(c)} \rho \, dF_{k,i,\tilde{k},\mathbf{x}}(\rho). \end{aligned}$$

In words, CVaR deviation measures the distance between the unconditional mean and mean in the $c \in (0, 1]$ probability tail of the distribution.

Proposition 5.1: Suppose that the price bundle, $\mathbf{p}$, is non-increasing and satisfies $R \geq p_1$. For intra-class scheduling policies of the form (15), the price bundle, $\mathbf{p}$, is incentive compatible for consumer i of type $\theta = (k, R, q)$, if and only if

$$\mathcal{D}_c(\rho_{k,i}(\mathbf{x}, \tilde{k})) \;\geq\; \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k})\right\} \;+\; q\left(\frac{p_k - p_{\tilde{k}}}{cR} - 1\right), \tag{16}$$

for all $\mathbf{x} \in \mathcal{X}$ and $\tilde{k} = k+1, \ldots, K$, where

$$c := \min\left\{F_{k,i,\tilde{k},\mathbf{x}}(q), \; \frac{R - p_{\tilde{k}}}{R}\right\}.$$

The IC condition in (16) has an appealing interpretation in terms of an explicit risk-reward trade-off. Namely, a consumer of type $\theta = (k, R, q)$ is disincented from reporting a false deadline $\tilde{k}$, if the risk incurred by reporting a false deadline $\tilde{k}$ exceeds the expected supply by period k plus a scalar, $\alpha_{k,\tilde{k}}$, which is proportional to the associated reduction in expenditure for $\tilde{k} = k+1, \dots, K$,

$$\alpha_{k,\tilde{k}} := q\left(\frac{p_k - p_{\tilde{k}}}{cR} - 1\right).$$

We observe that the scalar $\alpha_{k,\tilde{k}}$ is increasing in $p_k - p_{\tilde{k}}$, which implies that a consumer tends not to report a false deadline $\tilde{k}$, if the price difference between periods k (her true deadline) and $\tilde{k}$ is small.

Example 2 (Uniform Distribution): Suppose that the random variable $\rho_{k,i}(\mathbf{x}, \tilde{k})$ is uniformly distributed over a subset of $[0, 1]$. A straightforward calculation reveals that

$$\mathcal{D}_c(\rho_{k,i}(\mathbf{x}, \tilde{k})) \;=\; \sigma\sqrt{3}(1-c),$$

where σ denotes the standard deviation of $\rho_{k,i}(\mathbf{x}, \tilde{k})$. The incentive compatibility condition in (16) therefore simplifies to

$$\sigma\sqrt{3}(1-c) \;\geq\; \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k})\right\} + q\left(\frac{p_k - p_{\tilde{k}}}{cR} - 1\right), \quad (17)$$

which yields an explicit mean-variance trade-off interpretation of incentive compatibility. We finally note that, for a special case with $c = (R - p_{\tilde{k}})/R$ (cf. definition of c following Eq. (16)), the condition in (17) can be further simplified to

$$\sigma\sqrt{3}\frac{p_{\tilde{k}}}{R} \;\geq\; \mathbb{E}\left\{\rho_{k,i}(\mathbf{x}, \tilde{k})\right\} - q\left(\frac{R - p_k}{R - p_{\tilde{k}}}\right).$$

Essentially, a consumer will be disincented from reporting a false deadline if the standard deviation of supply exceeds the mean, less a constant measuring the reduction in expenditure.

VI. Conclusion

To explore the feasibility of deferrable electricity loads, we propose a novel market for deadline-differentiated energy services that provides a guarantee on the aggregate quantity of energy to be delivered by a consumer-specified deadline. For the earliest-deadline-first (EDF) scheduling policy, which is shown to be optimal for the supplier, we provide a characterization of the deadline-differentiated prices yielding a competitive equilibrium between the supplier and consumers.

Somewhat surprisingly, we show that the supplier's optimal price bundle is incentive compatible, in that every consumer would like to reveal her true deadline to the supplier, provided that the other consumers are truth-telling. This result implies the efficiency of a competitive equilibrium. We also provide incentive compatibility conditions for a special class of intra-class scheduling policies.

The market we have considered in our analysis is single shot. As a natural extension, it would be of interest to explore the dynamic analog of our formulation in which the market is cleared on a recurrent basis. In addition, such a dynamic setting could provide the foundation on which to explore efficient price discovery schemes.

Acknowledgment

The authors would like to thank Paul De Martini, Robert Kleinberg, Kameshwar Poolla, Pravin Varaiya, and Adam Wierman for their helpful discussions.

References

[1] M.H. Albadi, E.F. El-Saadany, "Demand response in electricity markets: an overview," in *Proc. of Power Engineering Society General Meeting*, June 2007, pp. 1-5.

[2] E. Bitar, P. Khargonekar, K. Poolla, R. Rajagopal, P. Varaiya, and F. Wu, "Selling random wind," in *Proc. of 45th Hawaii International Conference on System Science (HICSS)*, 2012.

[3] E. Bitar and S. Low, "Deadline differentiated pricing of deferrable electric power service," in *Proc. of the 51th IEEE Conference on Decision and Control*, 2012.

[4] R. T. Rockafellar, S. Uryasev, and M. Zarabankin, "Deviation Measures in Risk Analysis and Optimization," *Research Report 2002-7*, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL, 2002.

[5] CAISO, "Integration of Renewable Resources: Operational Requirements and Generation Fleet Capability at 20% RPS," California Independent System Operator Report, August 2010.

[6] D.S. Callaway and I.A. Hiskens, "Achieving controllability of electric loads," *Proc. of the IEEE*, vol. 99, pp. 184-199, 2011.

[7] H.P. Chao and R. Wilson, "Priority service: pricing, investment, and market organization," *The American Economic Review*, vol. 77, no. 5, 1987.

[8] S. Chen, T. He, and L. Tong, "Optimal deadline scheduling with commitment", in *Proc. of Allerton Conference on Communication, Control, and Computing*, 2011.

[9] FERC Staff Report, "A national assessment of demand response potential," FERC, June 2009.

[10] A. Krioukov et al., "Design and evaluation of an energy agile computing cluster," Technical Report EECS-2012-13, University of California at Berkeley, January 2012.

[11] S.S. Oren, and S.A. Smith, *Service Opportunities for Electric Utilities: Creating Differentiated Products*, Kluwer Academic Publishers: Boston, 1993.

[12] A. Papavasiliou, and S. Oren, "Supplying renewable energy to deferrable loads: Algorithms and economic analysis," in *Proc. of IEEE Power & Energy Society General Meeting*, 2010.

[13] M. Roozbehani, A, Faghih, M. Ohannessian, and M. A. Dahleh, "The intertemporal utility of demand and price-elasticity of consumption in power grids with shiftable loads," in *Proc. of the 50th IEEE Conference on Decision and Control and European Control Conference*, Orlando, FL, USA, 2011.

[14] A. Subramanian, et al., "Real-time scheduling of deferrable electric loads," in *Proc. of the American Control Conference*, Montreal, Canada, June 2012.

[15] C.W. Tan, and P.P. Varaiya, "Interruptible electric power service contracts," *Journal of Economic Dynamics and Control*, vol. 17, pp. 495-517, 1993.