Games and Economic Behavior 121 (2020) 482–505

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Optimal multi-unit mechanisms with private demands [☆]

Nikhil R. Devanur [a], Nima Haghpanah [b,*], Alexandros Psomas [c]

[a] Amazon, Seattle, WA 98109, USA
[b] Pennsylvania State University, University Park, PA 16802, USA
[c] Google Research, Mountain View, CA 94043, USA

ARTICLE INFO

Article history:
Received 21 September 2018
Available online 2 April 2020

Keywords:
Mechanism design
Multi-unit demands
Deterministic mechanisms

ABSTRACT

A seller can produce multiple units of a single good. The buyer has constant marginal value for each unit she receives up to a demand, and zero marginal value for units beyond the demand. The marginal value and the demand are drawn from a distribution and are privately known to the buyer. We show that under natural regularity conditions on the distribution, the optimal (revenue-maximizing) selling mechanism is deterministic. It is a price schedule that specifies the payment based on the number of units purchased. Further, under the same conditions, the revenue as a function of the price schedule is concave, which in turn implies that the optimal price schedule can be found in polynomial time. We give a more detailed characterization of the optimal prices when there are only two possible demands.

1. Introduction

What is the optimal selling strategy for a monopolist who can produce multiple units of a good or service? For instance, how should a cloud computing platform such as Amazon EC2 sell units of virtual machines, a cloud storage provider such as Dropbox sell units of storage, and a cellphone service providers like AT&T sell units of cellular data? We study this problem in a mechanism design setting with multiple dimensions of private information and provide simple conditions on the distribution of buyers' preferences with two implications. First, optimal mechanisms are deterministic. Second, the optimal price schedule can be computed efficiently. These results add a natural instance to the few multi-dimensional settings where optimal mechanisms are known to be structurally and computationally simple.

We consider the following mechanism design problem. There is a single good that can be produced in any number of units. There is a single buyer who has a linear valuation for consuming any number of units of the good up to a demand, but does not value consuming units beyond the demand.[1] The type of the buyer, consisting of her value per unit and her demand, is drawn at random from a distribution, and is known privately to the buyer. The buyer's utility is her value for the units received minus payment to the seller.

☆ A one-page abstract based on an earlier version of this paper appeared in the proceedings of the ACM Conference on Economics and Computation, 2017.

* Corresponding author.
E-mail addresses: iam@nikhildevanur.com (N.R. Devanur), haghpanah@psu.edu (N. Haghpanah), alexpsomi@cs.berkeley.edu (A. Psomas).

[1] In the examples mentioned above, it is perhaps more realistic to assume that a buyer has decreasing marginal values for units of the product. For example, the first minute of phone calls will be used for most urgent calls, the second unit for less urgent ones, and so on. The literature on multi-unit mechanisms usually assumes decreasing marginal values (Maskin et al. (1989); Ausubel et al. (2014)). Our model is a first approximation where the marginal value is fixed until the buyer is saturated, at which point the marginal value becomes zero. The assumption is made for tractability.

https://doi.org/10.1016/j.geb.2020.03.007
0899-8256/

A mechanism is a menu of lotteries. A lottery consists of a price and a distribution over the number of units. That is, once the buyer purchases the lottery by paying its price, she receives a random number of units from the distribution. The buyer purchases a lottery from the menu to maximize her expected utility (or purchases no lotteries and gets utility of zero). A mechanism is deterministic if each lottery in the menu gives the buyer a deterministic number of units. Below is a simple example that shows that a mechanism that uses randomization can obtain higher revenue than any deterministic mechanism. We represent the type of a buyer with a pair (v, d), where v is the per unit valuation and d is the demand.

Example 1 *(Deterministic mechanisms are not optimal).* Suppose that there are three uniformly distributed types $t_1 = (1, 3)$, $t_2 = (1, 2)$ and $t_3 = (6, 1)$. That is, for instance, type t_1 has a marginal value of 1 per unit for the first 3 units it receives (and has marginal value of zero beyond that). Consider the mechanism (which indeed is optimal for this instance) that offers the buyer to choose among the following lotteries:

(a) Pay 3. Receive 3 units. Or,
(b) Pay 1.5. Receive 2 units with probability $\frac{3}{4}$ and 0 units otherwise.

Types t_1 and t_3 buy 3 units, whereas type t_2 buys the lottery. Expected revenue is $\frac{7.5}{3}$.[2] We argue in Appendix A.1 that the highest revenue from deterministic mechanisms is $\frac{7}{3}$, obtained by offering 2 units for a price of 2 or 3 units for a price of 3.

It appears that the optimal mechanism is usually randomized for small examples with discrete support. This phenomenon is quite common. While Myerson (1981) and Riley and Zeckhauser (1983) show that the optimal mechanism is deterministic if all types have the same demand, randomized mechanisms become optimal even for slight generalizations. Optimal mechanisms are often structurally and computationally complex and may not even have a finite description (Daskalakis et al. (2017); Hart and Nisan (2013); Pavlov (2011); Hart and Reny (2015)). On the other hand, deterministic mechanisms are simple and commonly used in practice. Hence it is important to understand conditions that lead to optimality of deterministic mechanisms, and their implications on the computation of optimal mechanisms. In this paper we offer two insights in this regard.

Our contributions Our first contribution is to identify two conditions that collectively guarantees that there exists an optimal mechanism that is deterministic. The first condition, decreasing marginal revenue (DMR), requires that the revenue function associated with each demand is concave. The revenue function of a demand d maps each price p to the expected revenue of offering one unit of the good at price p only to types with demand d. That is, the revenue function of demand d is $p \cdot (1 - F_d(p))$, where F_d denotes the distribution of the values v conditioned on demand d. Requiring the revenue function to be concave is equivalent to requiring marginal revenue to be decreasing, hence the choice of name for the condition. The second condition requires that the supports of all distributions F_d have the same highest value. We show with examples that both conditions are indeed needed in the sense that either of them alone does not guarantee the optimality of deterministic mechanisms.

Our approach to showing optimality of deterministic mechanisms is as follows. In two steps, we show that any mechanism can be converted to a deterministic one with higher revenue. First, we convert a mechanism so that a type with highest valuation and demand d receives a deterministic allocation of d units. The first step uses the assumption that the supports of all conditional distributions have the same highest value. Second, we argue that a mechanism resulting from the first step can be converted to a deterministic mechanism. In particular, we remove all non-deterministic allocations from the mechanism, and allow types to choose only among the remaining deterministic allocations. This conversion weakly decrease the utility function of the mechanism, while keeping the utility of the highest types fixed. The DMR condition implies that such a conversion increases revenue.

Our second contribution is to show that assuming DMR, the optimal deterministic mechanisms can be found in polynomial time. To optimize over deterministic mechanisms, a first question is how a mechanism is to be represented. Direct representations can be conveniently used to formulate the problem of optimizing over *randomized* mechanisms as a linear program. The trouble with deterministic allocation constraints is that they turn the problem into an integer program. On the other hand, a deterministic mechanism does have a natural representation as a vector of prices. The difficulty is now that to express revenue of a given price vector, one needs to keep track of the decisions of all types. For example, if all types have a demand of either 1 or 2, then given a price of 3 for one unit and 4 for two units, all types with demand 2 buy either two or zero units, whereas given a price of 1 for one unit and 4 for two units, some such type, e.g. a type $(v, d) = (2, 2)$, buys one unit. A naive approach to optimizing over deterministic mechanisms would then be to partition the space of all price vectors based on the induced choices of all types, optimize prices within each region, and then compare prices between regions. This approach is quite computationally demanding since the number of possible regions is prohibitively large.

[2] The utilities of the two options are 0 and 0 respectively for type t_1, -1 and 0 for t_2, and 3 and 3 for t_2. We assume that ties are broken to maximize revenue.

We show that assuming DMR, revenue as a function of the price vector is concave. This implies that optimal prices can be found efficiently using standard convex optimization approaches such as gradient descent, ellipsoid, or other cutting plane methods (Khachiyan (1980); Vaidya (1989); Tat Lee et al. (2015)). Note that the definition of DMR requires concavity as a function of a *single price*, and does not immediately imply concavity as a function of the *vector of prices*. The same instance as in Example 1 can be used to show that, without imposing DMR, revenue as a function of the price vector need not be concave.

Example 2 *(Revenue function may not be concave).* Consider the instance from Example 1. Let us denote a vector of prices by $\mathbf{p} = (p_1, p_2, p_3)$, where p_i is the price of i units. Consider price vectors $\mathbf{p}^1 = (2, 2, 3)$, $\mathbf{p}^2 = (3, 3, 3)$, and their convex combination $\mathbf{p}^3 = \frac{1}{2}\mathbf{p}^1 + \frac{1}{2}\mathbf{p}^2 = (2.5, 2.5, 3)$. The revenue of $\mathbf{p}^1$ is $\frac{7}{3}$, the revenue of $\mathbf{p}^2$ is 2, and the revenue of $\mathbf{p}^3$ is $\frac{5.5}{3}$.[3] This violates concavity since the convex combination of revenues of $\mathbf{p}^1$ and $\mathbf{p}^2$ with weights $\frac{1}{2}$ each is $\frac{6.5}{3}$, strictly greater than the revenue of $\mathbf{p}^3$. Hence revenue as a function of the price vector is not concave.

We use our analysis of the revenue function to obtain two additional results. First, we characterize the optimal prices (of deterministic mechanisms) for the case of two demands 1 and 2, assuming DMR. In particular, we show that the optimal *monopoly prices* can be used as a guide to calculate the optimal prices. An optimal monopoly price for a distribution F_d is a price $\hat{p}_d$ that maximizes $p \cdot (1 - F_d(p))$. We show that optimal prices and optimal monopoly prices lie within the same region, as discussed above, in the space of all prices. For instance, if $\hat{p}_1 > \hat{p}_2$, then the same will hold for the optimal prices. Therefore, in this case, all types buy two units and the optimal prices can be calculated using simple first order conditions.

As another corollary of our analysis, we design approximately optimal pricing schemes for a dynamic version of our problem with demand uncertainty. In particular, we consider a seller who is facing a sequence of T buyers. The type of each buyer is drawn identically and independently from a distribution, but the distribution is unknown to the seller. If the distribution satisfies the DMR assumption, then we construct a dynamic pricing scheme such that the average revenue per round converges to the optimal revenue (i.e., if the seller knew the distribution) at the rate of $\frac{k^{9.5}}{\sqrt{T}}$, where k is the highest demand in the support of the distribution. We prove this result by using the concavity of revenue to convert our problem into a "convex bandits" problem (Agarwal et al. (2011); Bubeck et al. (2017)).

1.1. Related work

Preston McAfee and McMillan (1988) were the first to study optimality of deterministic mechanisms for selling multiple products. Among other results, they claimed that a condition similar to our DMR condition implies optimality of deterministic mechanisms. (Thanassoulis (2004) and Manelli and Vincent (2006) show via examples that the claim is incorrect.) Manelli and Vincent (2006) identify conditions (different than DMR) for optimality of deterministic mechanisms for selling two or three non-identical products. On the other hand, numerous papers show that deterministic mechanisms are not generally optimal or even approximately optimal (Hart and Nisan (2017), Thanassoulis (2004), Manelli and Vincent (2006), Briest et al. (2015)). We contribute to this literature by finding optimality conditions of deterministic mechanisms for a different generalization of the single product problem, namely selling multiple units instead of multiple products.

Malakhov and Vohra (2009) consider an auction design setting with multiple buyers with multi unit demands, but make two assumptions: (1) that the buyers cannot report a higher demand, and (2) that the Myerson virtual value is monotone in both the value and the demand.[4] They formulate the problem as a linear program and interpret it as a shortest path problem on a lattice.

Fiat et al. (2016) solve a problem related to ours with an assumption similar to the first assumption above, that buyers cannot report a higher demand. They consider what they call the "FedEx" problem, which too has a 2 dimensional type space, where one of them is a value v, and the other is a "deadline" d. In their model, all times earlier than d have the same valuation and times later than d have a zero valuation, whereas in our model, the valuation stays the same for higher d but degrades as d decreases. Fiat et al. (2016) characterize optimal mechanisms using duality and are not concerned with optimality of deterministic mechanisms.

2. The model and main results

We study a multi-unit mechanism design setting with a single buyer with private demand. There is a single item, any number of units of which can be produced at zero cost. The type (v, d) of the buyer specifies her (per unit) value

[3] Given prices $\mathbf{p}^1$, t_1 chooses 3 units and pays 3, t_2 chooses 2 units and pays 2, and t_3 chooses 2 units and pays 2, giving an expected revenue of 7/3. Given prices $\mathbf{p}^2$, t_1 chooses 3 units and pays 3, t_2 takes the outside option and pays 0, and t_3 chooses 3 units (or 1 or 2) and pays 3, giving an expected revenue of 2. Given prices $\mathbf{p}^3$, t_1 chooses 3 units and pays 3, t_2 takes the outside option and pays 0, and t_3 chooses 1 units (or 2) and pays 2.5, giving an expected revenue of $\frac{5.5}{3}$.

[4] The Myerson virtual value given a distribution with CDF F and PDF f is $\phi(v) = v - \frac{1-F(v)}{f(v)}$. In our case, we define the virtual value of a type (v, d) by applying the same definition using the marginal distribution on v, conditioned on d, and denoted by F_d and f_d. $\phi(v, d) = v - \frac{1-F_d(v)}{f_d(v)}$.

$v \in [\underline{v}_d, \bar{v}_d]$ and her demand $d \in \{1, 2, \cdots, k\}$.[5] The utility of type (v, d) for receiving $i \in \mathbb{Z}_+$ units of the item and paying $p \in \mathbb{R}$ is $v \cdot \min\{i, d\} - p$. That is, the type has marginal utility of v per unit for the first d units it receives, and after that it becomes satiated. Let $\mathcal{T}$ denote the set of all types. The type of the buyer is randomly drawn and is privately known to the buyer. Let q_d denote the probability that a type has demand d. Let f_d and F_d denote the probability density and the cumulative density functions of the distribution of values conditioned on demand d. Assume f_d is differentiable and that $f_d(v) > 0$ for all $v \in [\underline{v}_d, \bar{v}_d]$.

By the revelation principle, we restrict attention to direct mechanisms. A (direct) mechanism (A, p) is a pair of functions, an allocation rule $A : \mathcal{T} \to \Delta(\mathbb{Z}_+)$ and a payment rule $p : \mathcal{T} \to \mathbb{R}$.[6] Note that the mechanism may be randomized, that is, the allocation $A(v, d)$ is a random variable.[7]

A mechanism is incentive compatible (IC) if each type maximizes its utility by honestly reporting its true type. Formally, a mechanism (A, p) is IC if for all types (v, d) and (v', d'),

$$\mathbb{E}\left[v \min\{A(v,d), d\} - p(v,d)\right] \geq \mathbb{E}\left[v \min\left\{A(v',d'), d\right\} - p(v',d')\right],$$

where the expectation is taken over the randomization of the mechanism. The mechanism is individually rational (IR) if each type gets a non-negative utility from truthtelling.[8] That is, for all (v, d),

$$\mathbb{E}\left[v \min\{A(v,d), d\} - p(v,d)\right] \geq 0.$$

An IC and IR mechanism is optimal if it maximizes the expected revenue

$$\mathbb{E}\left[p(v,d)\right]$$

over all IC and IR mechanisms, where the expectation is taken over the randomization of types.

Our main theorems use the following assumptions on the distribution of types.

Definition 1. The conditional distributions have *decreasing marginal revenue* (DMR) if $v \cdot (1 - F_d(v))$ is concave in $v \in [\underline{v}_d, \bar{v}_d]$ for all d, and have *identical highest values* if there exists $\bar{v}$ such that $\bar{v}_d = \bar{v}$ for all d.

To explain the DMR condition, consider the distribution of types conditioned on demand d. The revenue of making a take-it-or-leave-it offer of one unit of the item at price v such types is

$$R_d(v) := v \cdot (1 - F_d(v)). \tag{1}$$

This is because any type (v', d) with $v' \geq v$ would take the offer, and any type (v', d) with $v' < v$ would leave the offer, resulting in payment v with probability $1 - F_d(v)$. Thus DMR, which requires concavity of the revenue function $v \cdot (1 - F_d(v))$, is equivalent to the fact that the marginal revenue, $1 - F_d(v) - v f_d(v)$, is non-increasing in v, hence the choice of name. Monotonicity of marginal revenue is in turn closely related to the usual definition of regularity of F_d, which requires that marginal revenue divided by $f_d(v)$, i.e., $\frac{1-F_d(v)}{f_d(v)} - v$, is monotone non-increasing. We provide a more detailed comparison between DMR and regularity in Section 6.1.

Our first theorem shows that if the conditional distributions have DMR and identical highest values, then the optimal mechanism is deterministic. A mechanism (A, p) is deterministic if for all (v, d), $A(v, d)$ is supported on $\{i\}$ for some $i \in \{1, \ldots, k\}$. We say that the optimal mechanism is deterministic if there exists an optimal mechanism that is deterministic.

Theorem 1. *If the conditional distributions have DMR and identical highest values, then the optimal mechanism is deterministic.*

By the taxation principle, a mechanism can be represented as a menu of lotteries. A lottery is a pair of a probability distribution over $\mathbb{Z}_+$ and a price, corresponding to a randomized allocation and payment. The buyer chooses the lottery that maximizes her expected utility from the menu. The menu may be of infinite size.

A deterministic mechanism can be represented as a menu of deterministic lotteries (i, p_i) for $i \in \{1, \ldots, k\}$, in which the buyer can get i units by paying p_i. Given the menu, a type (v, d) chooses i to maximize her utility $v \min\{i, d\} - p_i$. Let $\mathbf{p}$ denote the vector of prices $(p_1, \ldots, p_k)$. We assume without loss of generality that $p_1 \leq p_2 \leq \cdots \leq p_k$.[9] We denote by $\mathrm{Rev}(\mathbf{p})$ the (expected) revenue of this mechanism. Our second main theorem states that if the distribution of types has DMR, then revenue is concave as a function of $\mathbf{p}$.

---

[5] The assumption that $d \in \{1, 2, \cdots, k\}$ can be relaxed to $d \in \{d_1, d_2, \cdots, d_k\}$ for arbitrary d_i.

[6] $\Delta(\mathbb{Z}_+)$ is the set of probability distributions over $\mathbb{Z}_+$.

[7] By linearity of expectation, we may assume without loss of generality that the payment rule is deterministic.

[8] A stronger notion of individual rationality is to require the utility of every type to be non-negative for any realization of the random choices of the mechanism. In Appendix A.2, we show that any IR mechanism can be converted to one that satisfies this stronger individual rationality requirement.

[9] If $p_i > p_{i+1}$, then no type would choose to receive i units. Consider an alternative mechanism with prices $\mathbf{p}'$ such that $p'_j = p_j$ for all $j \neq i$, and $p'_i = p_{i+1}$. The new mechanism satisfies the property that $p'_i \leq p'_{i+1}$, and the two mechanisms have the same revenue.

Theorem 2. *If the conditional distributions have DMR, then* Rev($\mathbf{p}$) *is concave.*

We use concavity of the revenue function to obtain two additional results. First, we identify optimal deterministic mechanisms with two demands in Section 4.2. Second, we show in Section 4.3 that optimal deterministic mechanisms can be calculated efficiently using standard tools from convex optimization.

2.1. Both conditions of Theorem 1 are needed

We show with examples that both conditions of Theorem 1 are needed, in the sense that each one of them alone does not guarantee that deterministic mechanisms are optimal. In the first example, the conditional distributions do not have DMR. In the second example, the conditional distributions have non-identical highest values. Both examples are obtained by approximating the distribution in Example 1. Since there is a gap between the optimal revenue and the optimal deterministic revenue in Example 1, then by a continuity argument such a gap will also exist for distributions that approximate the distribution in Example 1 closely enough. We defer the detailed calculations to Appendix A.3. In our proof of Theorem 1, we will point out the exact place where either of the two conditions is used.

Example 3 *(No DMR; identical highest values).* The conditional distributions of Example 1 for $d = 1, 2$ and 3 are point masses at 1, 1 and 6 respectively. Replace them with normal distributions $\mathcal{N}(1, \sigma)$, $\mathcal{N}(1, \sigma)$ and $\mathcal{N}(6, \sigma)$, truncated at 0 and V, for some $V > 6$. Note that as σ goes to zero, the distributions approach those in Example 1. We show that for σ small enough, the optimal mechanism is not deterministic.

The example above does not imply that DMR cannot be weakened. The example does, however, show that DMR cannot be replaced with regularity, or even monotone hazard rate. The truncated normal distributions satisfy the monotone hazard rate condition. The monotone hazard rate condition requires the function $\frac{1-F(v)}{f(v)}$ to be monotone non-increasing, and is a stronger condition than regularity. Note that in the special case where $k = 1$, it is known from Myerson (1981) and Riley and Zeckhauser (1983) that deterministic mechanisms are optimal without any restrictions on the distribution.

The second example shows that the DMR condition alone does not guarantee the optimality of deterministic mechanisms.

Example 4 *(DMR; non-identical highest values).* Replace the conditional distributions of Example 1 for $d = 1, 2$ and 3 with uniform distributions over $[1 - \epsilon, 1 + \epsilon]$, $[1 - \epsilon, 1 + \epsilon]$, and $[6 - \epsilon, 6 + \epsilon]$. The conditional distributions have DMR but non-identical highest values. Note that as ϵ goes to zero, the distributions approach those in Example 1. For small enough ϵ, the optimal mechanism is not deterministic.

3. Optimal mechanisms

We start by exploring some structural properties of optimal mechanisms. We then use these properties to simplify the optimization problem and prove the first main theorem regarding optimality of deterministic mechanisms.

3.1. Structural properties

Allocating only the demanded number of units We first argue that it is without loss of generality to assume that any mechanism (A, p) allocates only the demanded number of units. In such a mechanism, a type who reports a demand of d either receives d units with some probability, or none at all. That is, for each (v, d), the random variable $A(v, d)$ is supported on $\{0, d\}$.

Lemma 1. *For any IC and IR mechanism $(\hat{A}, p)$, there exists a mechanism (A, p) that allocates only the demanded number of units, is IC and IR, and has the same expected revenue as $(\hat{A}, p)$.*

Proof. Define $A(v, d) = d$ with probability $\mathbb{E}\left[\min\left\{\hat{A}(v, d), d\right\}\right]/d$, and $A(v, d) = 0$ otherwise. The two mechanisms (A, p) and $(\hat{A}, p)$ have the same expected revenue since they have the same payment rule.

Note for future reference that for any $a, d, d' \in \mathbb{Z}_+$,

$$\begin{aligned}\min\left\{d, d'\right\} \frac{\min\left\{a, d'\right\}}{d'} &= \min\left\{\frac{\min\left\{d, d'\right\}}{d'} \times a, \frac{\min\left\{d, d'\right\}}{d'} \times d'\right\} \\ &\le \min\left\{a, \min\left\{d, d'\right\}\right\} \\ &\le \min\{a, d\}, \end{aligned} \tag{2}$$

with equality if $d = d'$ (if $d = d'$, both sides are equal to $\min\{a, d\}$).

We now argue that the mechanism (A, p) is IC and IR. If a type (v, d) reports (v', d'), it receives d' units with probability $\Pr[A(v', d') = d']$, and zero units otherwise. Thus its utility is

$$\begin{aligned}
& v \times \min\left\{d, d'\right\} \times \Pr[A(v', d') = d'] - p(v', d') \\
&= v \times \min\left\{d, d'\right\} \times \frac{\mathbb{E}\left[\min\left\{\hat{A}(v', d'), d'\right\}\right]}{d'} - p(v', d') \\
&= v \times \mathbb{E}\left[\min\left\{d, d'\right\} \frac{\min\left\{\hat{A}(v', d'), d'\right\}}{d'}\right] - p(v', d').
\end{aligned}$$

By (2), the utility of type (v, d) from reporting (v', d') in mechanism (A, p) is

$$\leq v \times \mathbb{E}\left[\min\left\{\hat{A}(v', d'), d\right\}\right] - p(v', d'),$$

with equality if $d' = d$. Note that the above expression is the utility of type (v, d) from reporting (v', d') in mechanism $(\hat{A}, p)$. Thus, the utility of (v, d) from reporting (v', d') in mechanism (A, p) is no larger than reporting (v', d') in mechanism $(\hat{A}, p)$, and the utility of truthtelling is equal in the two mechanisms. As a result, the mechanism (A, p) is IC and IR. □

Note that Lemma 1 does not imply that *every* optimal mechanism must allocate only the demanded number of units. We use Lemma 1 to simplify the optimization problem, and then argue that there exists a deterministic mechanism that obtains the same revenue as the optimal mechanism that allocates only the demanded number of units. The deterministic mechanism may indeed give $d' < d$ units to type (v, d).

We represent a mechanism that allocates only the demanded number of units by a pair (w, p) in which (v, d) receives d units with probability $w_d(v)$, receives 0 units otherwise, and pays $p(v, d)$. The utility of type (v, d) from reporting (v', d') can be written as $v \min\left\{d, d'\right\} w_{d'}(v') - p(v', d')$.

Local IC constraints are sufficient We now simplify the IC constraints. In particular, we show that it is sufficient to consider only a subset of "local" IC constraints that imply all other constraints. The first set of local constraints are "horizontal" constraints, in which a type (v, d) reports (v', d). The standard analysis from the theory of mechanism design states that a mechanism (w, p) satisfies the horizontal constraints if and only if w_d is monotone non-decreasing for all d, and the payment rule satisfies the payment identity à la Myerson (see, for example, Section 3 in Börgers (2015)). Formally, let U denote the indirect utility function of a mechanism (w, p) that specifies the utility of truthtelling for each type,

$$U_d(v) := v d w_d(v) - p(v, d). \tag{3}$$

Then the payment rule must satisfy

$$p\,(v, d) = v d w_d(v) - d \int\limits_{\underline{v}_d}^{v} w_d\,(z)\, dz - U_d(\underline{v}_d). \tag{4}$$

The second set of local constraints are the "diagonal" constraints, in which a type either increases or decreases its demand by one unit, and in each case, misreports its value to a certain value. In particular, it is sufficient to consider deviations of (v, d) to $(v, d-1)$ or $(vd/(d+1), d+1)$. To state our claim regarding simplification of IC constraints formally, let $u\left(v, d \to v', d'\right)$ denote the utility of type (v, d) from reporting (v', d').

Proposition 1. *A mechanism (w, p) is IC if and only if for all d and v,*

1. *w_d is monotone non-decreasing and $p(v, d)$ satisfies Equation (4), and*
2. *$U_d(v) \geq u\,(v, d \to v, d-1)$, and*
3. *$U_d(v) \geq u\left(v, d \to v\frac{d}{d+1}, d+1\right)$.*

Furthermore, an IC mechanism is IR if $U_d(\underline{v}_d) \geq 0$ for all d.

Proof. We start by proving the first statement regarding incentive compatibility. The equivalence of property 1 with horizontal IC constraints is standard and is omitted. Properties 2 and 3 are a subset of IC constraints and must be satisfied by any IC mechanism.

We next argue that the three properties imply incentive compatibility. Note that property 2 implies that

$$\begin{aligned}U_d(v) &\geq u(v, d \to v, d-1)\\&= v(d-1)w_{d-1}(v) - p(v, d-1)\\&= U_{d-1}(v).\end{aligned}$$

Applying this inequality inductively, we have that

$$U_d(v) \geq U_{d'}(v) \tag{5}$$

for all $d' \leq d$. Horizontal incentive compatibility implies that for $d' \leq d$,

$$U_{d'}(v) \geq u(v, d' \to v', d') = vd'w_{d'}(v') - p(v', d') = u(v, d \to v', d') \tag{6}$$

Note that (5) and (6) together imply that $U_d(v) \geq u(v, d \to v', d')$ for all $d' \leq d$ and v, v'. That is, a type does not benefit from reporting a lower demand.

To show that a type (v, d) does not benefit from reporting (v', d') for $d' \geq d$, note that property 3 implies that

$$\begin{aligned}U_d(v) &\geq u(v, d \to v\frac{d}{d+1}, d+1)\\&= vdw_{d+1}(v\frac{d}{d+1}) - p(v\frac{d}{d+1}, d+1)\end{aligned}$$

multiplying and dividing the first term by $d+1$,

$$\begin{aligned}&= v\frac{d}{d+1}(d+1)w_{d+1}(v\frac{d}{d+1}) - p(v\frac{d}{d+1}, d+1)\\&= U_{d+1}(v\frac{d}{d+1}).\end{aligned}$$

Applying this inequality inductively, we have that for all $d' \geq d$,

$$U_d(v) \geq U_{d'}(v\frac{d}{d'}). \tag{7}$$

From horizontal incentive compatibility we have

$$\begin{aligned}U_{d'}(v\frac{d}{d'}) &\geq u(v\frac{d}{d'}, d' \to v', d')\\&= v\frac{d}{d'}d'w_{d'}(v') - p(v', d')\\&= u(v, d \to v', d').\end{aligned} \tag{8}$$

Putting (7) and (8) together, we conclude that $U_d(v) \geq u(v, d \to v', d')$ for all $d' \geq d$ and v, v', which completes the proof of incentive compatibility.

We now prove the second statement regarding individual rationality. Note from the payment identity (4) that the utility of type (v, d) from truthtelling is

$$d\int_{\underline{v}_d}^{v} w_d(z)\,dz + U_d(\underline{v}_d) \geq U_d(\underline{v}_d).$$

We conclude that if the IR constraint is satisfied for the type $(\underline{v}_d, d)$, then it is also satisfied for all types (v, d). □

The global-to-local reduction of Theorem 1 is related to use in the FedEx problem. Syntactically, for the FedEx problem just the first 2 constraints above are sufficient, but the semantics are different.[10]

[10] In the FedEx problem the d's are the deadlines, and a larger d signifies an inferior product, whereas in our problem a larger d is a superior product. That the IC constraints still look the same for misreporting a lower d is due to the other difference between the problems: utility scales linearly with d in our problem, but remains constant in the FedEx problem. Thus in both problems, the valuation for an item of type $d' < d$ is the same for types (v, d) and (v, d').

3.2. The mathematical program

We now write a mathematical program that uses the properties discussed above to capture the optimal mechanism design problem. We write the program in terms of the indirect utility functions. The following lemma is a reformulation of Theorem 1 in terms of the indirect utility function U. Define $U'_d(v) = \frac{d}{dv}U_d(v)$.

Lemma 2. *A mechanism (w, p) with indirect utility U defined in (3) is IC if and only if*

1. *U_d is convex and $U'_d(v) = dw_d(v)$ if U'_d exists, and*
2. *$U_d(v) \geq U_{d-1}(v)$, and*
3. *$U_d(v) \geq U_{d+1}(v\frac{d}{d+1})$.*

The following lemma expresses expected revenue in terms of the indirect utility function using integration by parts à la Myerson. Let q_d denote the probability that demand is equal to d. Recall that the revenue function is defined as follows $R_d(v) = v \cdot (1 - F_d(v))$. Let $R''(v) = \frac{d^2}{dv^2}R(v)$ be the second derivative of the revenue function.

Lemma 3. *If a mechanism (w, p) is IC, then the expected revenue is*

$$\mathbb{E}[p(v,d)] = \sum_d \left(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) R''_d(v) dv \right) q_d.$$

Using Lemma 2 and Lemma 3, we now restate the optimal mechanism design problem.

$$\max_U \sum_d \left(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) R''_d(v) dv \right) q_d \tag{9}$$

subject to :

$$\begin{aligned}
&U_d(v) \text{ is convex} && \forall d \in \{1, \dots, k\} \\
&d \geq U'_d(v) \geq 0 && \forall d \in \{1, \dots, k\}, \forall v \\
&U_d(v) \geq 0 && \forall d \in \{1, \dots, k\} \\
&U_d(v) \geq U_{d-1}(v) && \forall d \in \{2, \dots, k\} \\
&U_d(v) \geq U_{d+1}\left(v\frac{d}{d+1}\right) && \forall d \in \{1, \dots, k-1\}
\end{aligned}$$

The objective of the mathematical program, the expected revenue, is written in terms of the indirect utility function using Lemma 3. Let us discuss the constraints. The first two constraints represent horizontal IC constraints and feasibility of the mechanism and follow from Lemma 2. In particular, the second constraint combines the requirement that $U'_d(v) = dw_d(v)$ with the feasibility constraint $0 \leq w_d(v) \leq 1$. The third constraint is the IR constraint. The last two constraints represent the diagonal constraints and follow from Lemma 2.

3.3. Optimality of deterministic mechanisms

We now turn to the proof of the main theorem, Theorem 1, that a deterministic mechanism is the optimal solution to the above revenue maximization program. To prove our main result, we utilize the DMR property through the following lemma. The lemma allows us to compare the revenue of mechanisms by comparing their indirect utilities pointwise. In particular, the lemma states that by lowering the utilities of all types while keeping the utility of types with the highest value fixed, we can improve the revenue of a mechanism. The observation is standard and is frequently made in the mechanism design literature (e.g., Preston McAfee and McMillan (1988)).

Lemma 4. *Consider two IC mechanisms with indirect utilities U and $\hat{U}$, such that $U_d(v) \geq \hat{U}_d(v)$ for all types, with equality for $v = \underline{v}_d$ for all d. If the distribution of types has DMR, then the revenue of the mechanism with indirect utility $\hat{U}$ is at least as high as the revenue of the mechanism with indirect utility U.*

Proof. The proof follows directly from the expression of revenue in Lemma 3. Since $U_d(v) \geq \hat{U}_d(v)$ for all types, with equality for $v = \underline{v}_d$, we have

$$v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} \leq v f_d(v) \hat{U}_d(v)|_{\underline{v}_d}^{\bar{v}_d}$$

The DMR property, which requires concavity of R_d, is equivalent to $R_d''(v) \le 0$. Together with the assumption that $U_d(v) \ge \hat{U}_d(v)$, we have $U_d(v)R_d''(v) \le \hat{U}_d(v)R_d''(v)$. Thus revenue of the mechanism with indirect utility U is

$$\sum_d \Big(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) R_d''(v) dv \Big) q_d$$
$$\le \sum_d \Big(v f_d(v) \hat{U}_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} \hat{U}_d(v) R_d''(v) dv \Big) q_d,$$

which is the revenue of the mechanism with indirect utility $\hat{U}$. □

The main theorem, Theorem 1, follows immediately from the following two lemmas. The first lemma uses the property that the conditional distributions have identical highest values, whereas the second lemma uses the fact that they have DMR.

The first lemma states that we can improve the revenue of any mechanism by assigning deterministic allocations to types with highest values, assuming that conditional distributions have identical highest values. More precisely, if $\bar{v} = \bar{v}_d$ for all d, then type $(\bar{v}, d)$ is deterministically assigned d units. The intuition is as follows. Consider any IC mechanism and a type $(\bar{v}, d)$. Construct a new mechanism in which the type $(\bar{v}, d)$ receives d units deterministically, and set the payment of $(\bar{v}, d)$ in the new mechanism such that $(\bar{v}, d)$ is indifferent between its assignments in the old and the new mechanism. Let the allocation and payment of each other type be identical in the two mechanisms. We make two claims. First, this modification increases revenue. This is because the new mechanism gives type $(\bar{v}, d)$ its most desirable allocation, and therefore can charge it more while giving it the same utility. Second, the modified mechanism is incentive compatible. This is because the type $(\bar{v}, d)$, among all other types, is willing to pay the highest amount to switch from its old allocation to a deterministic allocation of d units. If the price for this change (the difference in payments) is such that $(\bar{v}, d)$ is indifferent, no other type would be willing to take the new allocation. Note that the lemma below does not require the DMR condition.

Lemma 5. *Assume that $\bar{v} = \bar{v}_d$ for all d. For any IC and IR mechanism, there exists an IC and IR mechanism, with revenue at least as large, in which type $(\bar{v}, d)$ deterministically receives d units for all d.*

Proof. Fix any IC and IR mechanism $(\hat{w}, \hat{p})$ with induced utility $\hat{U}$. Construct a mechanism (w, p) as follows. For each demand d, define $w_d(\bar{v}) = 1$ and $p_d(\bar{v}) = \bar{v}d - \hat{U}_d(\bar{v})$. For all types (v, d) with $v < \bar{v}$, let the allocation and payment be identical to that of $(\hat{w}, \hat{p})$, that is $w_d(v) = \hat{w}_d(v)$ and $p_d(v) = \hat{p}_d(v)$.

First, notice that the revenue of the mechanism (w, p) is no lower than the revenue of $(\hat{w}, \hat{p})$. Indeed, note that

$$\begin{aligned} p_d(\bar{v}) &= \bar{v}d - \hat{U}_d(\bar{v}) \\ &= \bar{v}d - (\bar{v}d\hat{w}_d(\bar{v}) - \hat{p}_d(\bar{v})) \\ &= \bar{v}d(1 - \hat{w}_d(\bar{v})) + \hat{p}_d(\bar{v})). \end{aligned} \tag{10}$$

Thus $p_d(\bar{v}) \ge \hat{p}_d(\bar{v})$ while payments of all other types remain the same.

Second, we argue that the mechanism (w, p) is IR by showing that the utility of truthtelling in the two mechanisms is the same, that is

$$U_d(v) = \hat{U}_d(v). \tag{11}$$

Indeed, for any type $(\bar{v}, d)$ we have $U_d(\bar{v}) = \bar{v}d - (\bar{v}d - \hat{U}_d(\bar{v})) = \hat{U}_d(\bar{v})$. And for any type (v, d) where $v < \bar{v}$, the allocation and the payment remains the same and thus $U_d(v) = \hat{U}_d(v)$ trivially.

Third, we argue that the mechanism (w, p) is incentive compatible. Notice that the utility of $(\bar{v}, d)$ from reporting (v', d') is the same in the two mechanisms. Thus we only need to show that a type (v, d) has no incentive to misreport to $(\bar{v}, d')$. The utility from misreporting is

$$\begin{aligned} u(v, d \to \bar{v}, d') &= v \min(d, d') - p_{d'}(\bar{v}) \\ &= v \min(d, d') - \bar{v}d'(1 - \hat{w}_{d'}(\bar{v})) - \hat{p}_{d'}(\bar{v}), \end{aligned} \tag{12}$$

where the second equality followed from (10). We now use the assumption that $\bar{v} = \bar{v}_d$ for all d. This assumption, together with the facts that $1 - \hat{w}_{d'}(\bar{v}) \ge 0$ and $d' \ge \min(d, d')$, implies that

$$\bar{v}d'(1 - \hat{w}_{d'}(\bar{v})) \ge v \min(d, d')(1 - \hat{w}_{d'}(\bar{v})).$$

Substituting the above inequality into (12), we have

$$u(v, d \to \bar{v}, d') \le v \min(d, d') - v \min(d, d')(1 - \hat{w}_{d'}(\bar{v})) - \hat{p}_{d'}(\bar{v})$$
$$= v \min(d, d')\hat{w}_{d'}(\bar{v}) - \hat{p}_{d'}(\bar{v}).$$

The last expression is the utility that type (v, d) would obtain from misreporting type $(\bar{v}, d')$ in the mechanism $(\hat{w}, \hat{p})$. By incentive compatibility of $(\hat{w}, \hat{p})$, the above expression is at most the utility that type (v, d) gets from truthtelling, $\hat{U}_d(v)$. Therefore, we conclude that

$$u(v, d \to \bar{v}, d') \le \hat{U}_d(v) = U_d(v),$$

where the last equation followed from (11). Thus the mechanism is incentive compatible, and the lemma follows. □

Notice that Lemma 5 does not hold with non-identical highest values. Indeed, in Example 1, each of the three types has the highest value among all types with that demand. However, deterministic mechanisms are not optimal.

The next lemma shows that for any mechanism in which all types $(\bar{v}_d, d)$ deterministically receive d units, there exists a deterministic mechanism with revenue at least as large. The lemma requires DMR, but not identical highest values. The intuition is that by removing all non-deterministic allocations from the mechanism, the utility of every type would weakly decrease, while the utility of a type $(\bar{v}_d, d)$ stays the same since it still has access to its most desirable assignment (the deterministic assignment). Lemma 4 can then be used to argue that the revenue of a deterministic mechanism is weakly higher.

Lemma 6. *Consider any IC and IR mechanism where a type $(\bar{v}_d, d)$ deterministically receives d units, for all d. If the distribution of types has DMR, then there exists a deterministic IC and IR mechanism with revenue at least as large.*

Proof. Fix any type with highest value $(\bar{v}_d, d)$ that deterministically receives d units. Consider the menu representation of the mechanism: it offers, among other lotteries, deterministic allocations of d units, for all d. Now construct an alternative menu that only offers such deterministic allocations, in addition to the choice of zero units at price zero. The alternative menu contains k choices of deterministic allocations of 1 to k units, plus the choice of zero units. Note that the utility function of the alternative mechanism is pointwise (weakly) smaller than the utility function of the original mechanism, since each type faces a smaller menu of choices. Furthermore, the utility of type $(\bar{v}_d, d)$ remains the same for all d, since the deterministic allocations that they chose in the original mechanism are still available in the alternative mechanism. By Lemma 4, the revenue of the alternative mechanism is no lower than the revenue of the original mechanism. □

We are now ready to complete the proof of Theorem 1.

Proof of Theorem 1. Consider any IC and IR mechanism. By Lemma 5, there exists a mechanism with revenue at least as large, where any type $(\bar{v}, d)$ deterministically receives d units. By Lemma 6, there exists a deterministic mechanism with revenue at least as large. Notice that an optimal deterministic mechanism, that is a vector of prices $\mathbf{p}$ that maximizes $\text{Rev}(\mathbf{p})$, exists. The optimal deterministic mechanism is IC and IR, and has no less revenue than any other mechanism, and is thus optimal. □

4. The optimal (deterministic) prices

Equipped with Theorem 1, we now turn to the problem of identifying optimal prices in deterministic mechanisms. In Section 4.1, we study revenue as a function of prices of deterministic allocations (henceforth simply referred to as “the prices”), and prove Theorem 2. We use the properties of the revenue function in the subsequent subsections. In Section 4.2, we identify optimal prices for the case of two demands. In Section 4.3, we argue that concavity of revenue can be used to compute optimal prices efficiently. The next section also uses concavity to develop approximately optimal dynamic pricing schemes with demand uncertainty.

4.1. Structure of the revenue function

In this section study the structure of revenue as a function of prices of deterministic allocations, culminating with the proof that revenue is concave (Theorem 2). Recall that the demands in the support of the distribution are $1, \dots, k$, and that for all $i \in \{1, \dots, k\}$, p_i denotes the price for the bundle of i units, and $\mathbf{p} = (p_1, \dots, p_k)$ denotes the vector of all prices. Without loss of generality, we may assume that the domain of $\mathbf{p}$ is such that

$$0 \le p_1 \le p_2 \le \dots \le p_k.$$

With this, we may assume that a type (v, i) only buys a bundle of j units for $j \le i$, and gets utility $vj - p_j$. We restate Theorem 2 for convenience.

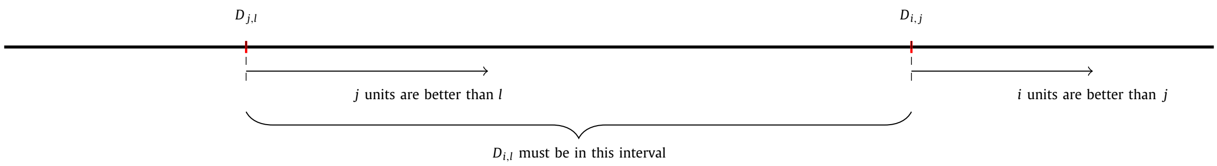

Fig. 1. The threshold $D_{i,l}$ is in the interval $[D_{j,l}, D_{i,j}]$.

Fig. 2. An example where $p_5 = 8$, $p_4 = 6$, $p_3 = 2$, $p_2 = 1$ and $p_1 = 1/2$. The $D_{i,j}$s define the intervals in which a buyer (v,i) purchases different number of units.

Theorem 2. *If the conditional distributions have DMR, then* Rev($\mathbf{p}$) *is concave.*

Characterizing optimal bundles Given a price vector $\mathbf{p}$, the revenue is determined by what the utility maximizing bundle is for each type. To analyze this, we first consider when a given type prefers a bundle of j units to one of l units, for $j \neq l \in \{1,\dots,k\}$. The following quantity turns out to be the threshold at which the preference changes.

$$\forall j,l \in \{1,\dots,k\} : j > l, \quad D_{j,l} \triangleq \frac{p_j - p_l}{j-l} \quad . \tag{13}$$

For convenience, we also define $D_{j,0} \triangleq p_j/j$ for all $j \in \{1,\dots,k\}$. The next lemma states precisely what it means for $D_{j,l}$ to be a threshold: if $v > D_{j,l}$ then a buyer with type (v,i) strictly prefers a bundle of j units to a bundle of l units, where $i \geq j > l$. If $v = D_{j,l}$ the buyer is indifferent between the two options.

Lemma 7. *For all $i \geq j > l \in \{1,\dots,k\}$, a buyer of type (v,i) prefers a bundle of j units to a bundle of l units if and only if $v > D_{j,l}$. Both bundles are equally preferable precisely when $v = D_{j,l}$.*

Proof. The buyer prefers j units over l units if and only if $vj - p_j > vl - p_l$. Rearranging, we get the lemma. $\square$

Before we proceed further, we note the following property for future reference. The threshold $D_{i,l}$ where a buyer switches between preferring i units to l units is between thresholds $D_{i,j}$ and $D_{j,l}$, for all $i \geq j \geq l \in \{1,\dots,k\}$. Intuitively, if $D_{i,j} > D_{j,l}$, then the buyer prefers j units to l units when $v > D_{j,l}$, and i to j units when $v > D_{i,j}$. Therefore, for the buyer to be indifferent between i and l, her value must be in the interval $[D_{j,l}, D_{i,j}]$. See Fig. 1.

Lemma 8. *For all $i \geq j \geq l \in [k]$, $D_{i,l}$ is a convex combination of (and hence is always in between) $D_{i,j}$ and $D_{j,l}$.*

Proof. The lemma follows directly from the identity $D_{i,l} = \frac{1}{i-l}\big((i-j)D_{i,j} + (j-l)D_{j,l}\big)$. $\square$

We next consider how the optimum bundle changes for a given demand i, as v decreases from $\bar{V}$ to 0. For high enough v, the optimum bundle for type (v,i) is i units. As v decreases, the optimal bundle is going to switch at the threshold $\max_{j<i}\{D_{i,j}\}$ to some j in the arg max. Similarly, as v decreases further, by Lemma 7 the optimal bundle is going to switch at $\max_{l<j}\{D_{j,l}\}$ and so on. These sequences for different i are not independent, and we can capture each such sequence of optimum bundles by a single vector $\sigma \in \mathbb{Z}^k$ such that the ith coordinate $\sigma(i) \in \arg\max_{j<i}\{D_{i,j}\}$. Given such a σ, for each i, the sequence of optimal bundles for types with demand i is given by the directed path $\mathcal{P}_\sigma(i)$, defined as the (unique) longest path starting from i in the directed graph on $\{1,\dots,k\}$ with edges $(i,\sigma(i))$. (The path ends when $\sigma(i) = 0$ for some i.)

For example, consider the case that $p_5 = 8$, $p_4 = 6$, $p_3 = 2$, $p_2 = 1$ and $p_1 = 1/2$. Then, the maximum of the $D_{5,i}$s is $D_{5,3} = 3$, the maximum of the $D_{3,i}$s is $D_{3,2}$, and $D_{2,1} > 0$. Therefore, $\mathcal{P}_\sigma(5)$ is the directed path $5 \to 3 \to 2 \to 1 \to 0$. See Fig. 2.

Given a price vector $\mathbf{p}$, there is a closed form formula for the revenue function that depends on the resulting σ. It is going to be more useful, however, to consider the inverse of this map from σ to $\mathbf{p}$: given any $\sigma \in \mathbb{Z}^k$ such that $\sigma(i) \in \{1,\dots,i-1\}$, we define Δ_σ to be all the price vectors where the sequence of optimal bundles (as described above) is given by $\mathcal{P}_\sigma(i)$. Formally,

$$\Delta_\sigma \triangleq \left\{\mathbf{p} : \forall i, \sigma(i) \in \arg\max_{j<i}\{D_{i,j}\}\right\}. \tag{14}$$

Revenue function formula We are now ready to give a closed form formula for the revenue function within each Δ_σ. Recall that F_i denotes the distribution of values conditioned on demand i, and that q_i is the probability that the demand is i. We also use $\sigma^2(i)$ to denote $\sigma(\sigma(i))$. The following function $\mathrm{Rev}_\sigma(\mathbf{p})$, corresponding to some σ, captures the revenue of a price vector $\mathbf{p}$ in Δ_σ.

$$\mathrm{Rev}_\sigma(\mathbf{p}) \triangleq \sum_i q_i \left(p_i\left(1 - F_i(D_{i,\sigma(i)})\right) + \sum_{j\in\mathcal{P}_\sigma(i)} p_{\sigma(j)}\left(F_i(D_{j,\sigma(j)}) - F_i(D_{\sigma(j),\sigma^2(j)})\right)\right) \tag{15}$$

The following lemma states that if $\mathbf{p} \in \Delta_\sigma$, then $\mathrm{Rev}_\sigma(\mathbf{p})$ is indeed the revenue of the price vector $\mathbf{p}$.

Lemma 9. $\mathrm{Rev}(\mathbf{p}) = \mathrm{Rev}_\sigma(\mathbf{p})$ *for all* $\mathbf{p} \in \Delta_\sigma$.

Proof. Suppose $\mathbf{p} \in \Delta_\sigma$. Consider all buyer types with demand i. Among these, all types with value $v > D_{i,\sigma(i)}$ prefer to buy the bundle of i units over any other bundle, by Lemma 7, and because $\mathbf{p} \in \Delta_\sigma$. These contribute $q_i p_i\left(1 - F_i(D_{i,\sigma(i)})\right)$ to the revenue.

Now consider all types with value $v \in [D_{j,\sigma(j)}, D_{\sigma(j),\sigma^2(j)}]$ for some $j \in \mathcal{P}_\sigma(i)$. We need to prove that these prefer a bundle of $\sigma(j)$ units over any other bundle l, so that they contribute to the revenue exactly $q_i p_{\sigma(j)}\left(F_i(D_{j,\sigma(j)}) - F_i(D_{\sigma(j),\sigma^2(j)})\right)$, and the lemma follows. As characterized by Lemma 7, this is implied by the following.

- If $l < \sigma(j)$, then $v \ge D_{\sigma(j),\sigma^2(j)} \ge D_{\sigma(j),l}$. This holds because $\mathbf{p} \in \Delta_\sigma$.
- If $i \ge l > \sigma(j)$, then $v \le D_{j,\sigma(j)} \le D_{l,\sigma(j)}$. We show this in the rest of the proof.

We first prove that $\forall j \in \mathcal{P}_\sigma(i)$, $D_{j,\sigma(j)} \ge D_{\sigma(j),\sigma^2(j)}$. This follows from the fact that $D_{j,\sigma^2(j)}$ is between $D_{j,\sigma(j)}$ and $D_{\sigma(j),\sigma^2(j)}$ (Lemma 8), and that $D_{j,\sigma^2(j)} \le D_{j,\sigma(j)}$ (since $\mathbf{p} \in \Delta_\sigma$). We now prove the following: $\forall j \in \mathcal{P}_\sigma(i)$, and $l \in (\sigma(j), j]$, we have that $D_{l,\sigma(j)} \ge D_{j,\sigma(j)}$. This follows from the fact that if $l \in (\sigma(j), j]$, then $D_{j,\sigma(j)}$ is in between $D_{j,l}$ and $D_{l,\sigma(j)}$ (from Lemma 8), and $D_{j,l} \le D_{j,\sigma(j)}$. Now by a repeated application of the fact $D_{j,\sigma(j)} \ge D_{\sigma(j),\sigma^2(j)}$, we get the same conclusion for all j and l such that $i \ge l > \sigma(j)$. □

Concavity of Rev_σ Given a closed form for the revenue of a price vector, we next show that each Rev_σ is a concave function. We do this by showing that Rev_σ can be written as a positive linear combination of linear functions, and compositions of $v(1 - F_d(v))$ with linear functions. The $v(1 - F_d(v))$ functions are concave by the DMR assumption, and such compositions and positive linear combinations preserve concavity, so Rev_σ is concave as well. Note that $\mathrm{Rev}_\sigma(\mathbf{p})$ being concave, does not imply that $\mathrm{Rev}(\mathbf{p})$ is concave.

Lemma 10. *For all* σ, $\mathrm{Rev}_\sigma(\mathbf{p})$ *is a concave function.*

Proof. Notice that using summation by parts we can write

$$p_i F_i(D_{i,\sigma(i)}) - \sum_{j\in\mathcal{P}_\sigma(i)} p_{\sigma(j)}\left(F_i(D_{j,\sigma(j)}) - F_i(D_{\sigma(j),\sigma^2(j)})\right) = \sum_{j\in\mathcal{P}_\sigma(i)} F_i(D_{j,\sigma(j)})\left(p_j - p_{\sigma(j)}\right)$$

Therefore we can rewrite Rev_σ (defined in (15)) as follows,

$$\begin{aligned}\mathrm{Rev}_\sigma &= \sum_i q_i\left(p_i - \sum_{j\in\mathcal{P}_\sigma(i)} F_i(D_{j,\sigma(j)})\left(p_j - p_{\sigma(j)}\right)\right)\\ &= \sum_i q_i\left(p_i - \sum_{j\in\mathcal{P}_\sigma(i)} F_i(D_{j,\sigma(j)}) D_{j,\sigma(j)}\left(j - \sigma(j)\right)\right),\end{aligned}$$

where the second equality followed from the definition of the thresholds D, Equation (13). We assumed that $v(1 - F_i(v))$ is concave, which implies that $-vF_i(v)$ is concave. $D_{j,\sigma(j)}$ is a linear function of $\mathbf{p}$ for all j. Since composition of linear functions with concave functions is concave, it follows that $-D_{j,\sigma(j)}F_i(D_{j,\sigma(j)})$ is concave. Now Rev_σ is a positive linear combination of concave functions, which makes it concave too. □

Stitching the Rev_σ*s together* Lemma 9 and Lemma 10 imply that Rev is piecewise concave, i.e., inside each Δ_σ it is concave. In general this does not imply that such a function is concave everywhere. One property that would imply that Rev is concave everywhere would be if Rev was equal to $\min_\sigma \mathrm{Rev}_\sigma$, since the minimum of concave functions is concave. Unfortunately, this is not true. In fact, for any $\sigma \neq \sigma'$, there exist prices $\mathbf{p}$ and $\mathbf{p}$ such that $\mathrm{Rev}_\sigma(\mathbf{p}) > \mathrm{Rev}_{\sigma'}(\mathbf{p})$ and

$\mathrm{Rev}_\sigma(\mathbf{p}') < \mathrm{Rev}_{\sigma'}(\mathbf{p}')$. We show a different, and somewhat surprising, property of the Rev_σ that also implies that Rev is concave. We show that at the boundaries between two regions not only do the corresponding Rev_σs agree (which they should, for Rev to be even continuous), but also their gradients agree.

Lemma 11. *For all σ, σ', $\mathbf{p}$ such that $\mathbf{p} \in \Delta_\sigma \cap \Delta_{\sigma'}$, we have that*

$$\mathrm{Rev}_\sigma(\mathbf{p}) = \mathrm{Rev}_{\sigma'}(\mathbf{p}) \text{ and } \nabla \mathrm{Rev}_\sigma(\mathbf{p}) = \nabla \mathrm{Rev}_{\sigma'}(\mathbf{p}).$$

Proof. We first argue that it is sufficient to prove Lemma 11 for the case where σ and σ' disagree in exactly one coordinate, i.e., there is some i^* such that $\sigma(i^*) \neq \sigma'(i^*)$, and $\forall j \neq i^*$, $\sigma(j) = \sigma'(j)$. Suppose we have done that. Now consider any two σ and σ', and a sequence $\sigma = \sigma_1, \sigma_2, \ldots, \sigma_n = \sigma'$ such that for any i, σ_i and σ_{i+1} differ in exactly one coordinate, where σ_i agrees with σ in that coordinate and σ_{i+1} agrees with σ'. The fact that $\mathbf{p} \in \Delta_\sigma \cap \Delta_{\sigma'}$ implies that for all coordinates j such that $\sigma(j) \neq \sigma'(j)$, we have $\sigma(j), \sigma'(j) \in \arg\max_{j'<j}\{D_{j,j'}\}$. Similarly, $\mathbf{p} \in \Delta_{\sigma_i} \cap \Delta_{\sigma_{i+1}}$ requires the same condition, but only for the coordinate that they differ in, and therefore $\mathbf{p} \in \cap_{i=1}^n \Delta_{\sigma_i}$. Since we know Lemma 11 holds when the two σs differ in at most one coordinate, it now follows that Rev and ∇Rev at $\mathbf{p}$ are the same for all σ_i and hence for σ and σ' as well.

Now we prove Lemma 11 when σ and σ' differ at exactly one coordinate, i^*. We consider the portions of the paths $\mathcal{P}_\sigma(i)$ and $\mathcal{P}_{\sigma'}(i)$ that are disjoint, and refer to these disjoint portions as simply $\mathcal{P} \subseteq \mathcal{P}_\sigma(i)$ and $\mathcal{P}' \subseteq \mathcal{P}_{\sigma'}(i)$. Both of these paths start at i^* and end at $\hat{i}$. Note that once the two paths merge, they remain the same for the rest of the way. If the paths do not merge, then we let $\hat{i} = 0$. The critical fact we use is that along these paths the Ds are all the same, which is stated as the following lemma.

Lemma 12. *All $j, j' \in \mathcal{P} \cup \mathcal{P}'$ s.t. $j > j'$ have the same $D_{j,j'}$.*

Proof. We prove the lemma by induction, where we add one node at a time in the following order. We start the base case with $i^*, \sigma(i^*)$ and $\sigma'(i^*)$. At any point let j and j' be the last points on $\mathcal{P}$ and $\mathcal{P}'$ that we have added so far. In the inductive step, if $j > j'$, we add $\sigma(j)$ and otherwise we add $\sigma'(j')$. We stop when all nodes in $\mathcal{P} \cup \mathcal{P}'$ have been added.

For the base case, let $j = \sigma(i^*)$ and $j' = \sigma'(i^*)$. Without loss of generality, assume that $j > j'$. By Lemma 8, we get that $D_{i^*,j'}$ is between $D_{i^*,j}$ and $D_{j,j'}$. Since $D_{i^*,j} = D_{i^*,j'}$, from the definition of i^*, we get $D_{i^*,j} = D_{j,j'} = D_{i^*,j'}$.

For the inductive step, let $j \in \mathcal{P}$ and $j' \in \mathcal{P}'$ be the last points that we have added so far, and again without loss of generality $j > j'$. Let $v = \sigma(j)$. If $v = j'$ we are done. There are two cases: $v > j'$ and $v < j'$. In the former case, we have $D_{j,v} \geq D_{j,j'}$ from the definition of v. From Lemma 8, $D_{j,j'}$ must be in between $D_{j,v}$ and $D_{v,j'}$, therefore $D_{j,j'} \geq D_{v,j'}$. Let $i' \in \mathcal{P}'$ be the predecessor of j', i.e., $\sigma'(i') = j'$. Due to the order in which we added the nodes, it must be that $i' > j$. By definition, $D_{i',j'} \geq D_{i',v}$, and by Lemma 8, $D_{i',j'}$ must be in between $D_{i',v}$ and $D_{v,j'}$, therefore $D_{v,j'} \geq D_{i',j'}$. By the inductive hypothesis, we have that $D_{i',j'} = D_{j,j'}$ and hence they both must be equal to $D_{v,j'}$.

Now consider any $i \neq j'$ that we have already added. It must be that $i < v$, and hence $D_{i,j'}$ must be in between $D_{i,v}$ and $D_{v,j'}$, but from the argument in the previous paragraph and the inductive hypothesis, we have that $D_{i,j'} = D_{v,j'}$, and hence they must be equal to $D_{i,v}$. This completes the induction for this case. The latter case of $v < j'$ is identical. □

Continuing the proof of Lemma 11 To show that Rev_σs agree on the boundary, consider the difference $\mathrm{Rev}_\sigma(\mathbf{p}) - \mathrm{Rev}_{\sigma'}(\mathbf{p})$. For all $i \leq i^*$, or i such that $i^* \notin \mathcal{P}_\sigma(i)$, nothing changes, therefore all those terms cancel out. Moreover, even for i such that $i^* \in \mathcal{P}_\sigma(i)$, the only terms that don't cancel out are $j \in \mathcal{P} \cup \mathcal{P}'$. Therefore, we get:

$$\mathrm{Rev}_\sigma(\mathbf{p}) - \mathrm{Rev}_{\sigma'}(\mathbf{p}) = \sum_{i \geq i^*: i^* \in \mathcal{P}_\sigma(i)} q_i \Big(\sum_{j \in \mathcal{P}} p_{\sigma(j)} \left(F_i(D_{j,\sigma(j)}) - F_i(D_{\sigma(j),\sigma^2(j)})\right)$$
$$- \sum_{j \in \mathcal{P}'} p_{\sigma'(j)} \left(F_i(D_{j,\sigma'(j)}) - F_i(D_{\sigma'(j),(\sigma')^2(j)})\right)\Big),$$

which is zero by Lemma 12.

For the second part of the proof, we'll show that the gradient of $\mathrm{Rev}_\sigma - \mathrm{Rev}_{\sigma'}$ is zero. We only need to consider the partial derivatives w.r.t. p_j for $j \in \mathcal{P} \cup \mathcal{P}'$. Fix a $j \in \mathcal{P}$, and consider the terms in $\frac{\partial(\mathrm{Rev}_\sigma - \mathrm{Rev}_{\sigma'})}{\partial p_j}$ corresponding to some $i \geq i^*$ such that $i^* \in \mathcal{P}_\sigma(i)$, in the outer summation. Let the path $\mathcal{P}_\sigma(i)$ be such that $a \in \mathcal{P}_\sigma(i)$, $b = \sigma(a)$, $j = \sigma(b)$, $c = \sigma(j)$ and $d = \sigma(c)$.

$$i \to \ldots \to i^* \to \ldots \to a \to b \to j \to c \to d \to \ldots$$

Then the terms under consideration are

$$\frac{\partial}{\partial p_j} q_i \left(p_b \left(F_i(D_{a,b}) - F_i(D_{b,j})\right) + p_j \left(F_i(D_{b,j}) - F_i(D_{j,c})\right) + p_c \left(F_i(D_{j,c}) - F_i(D_{c,d})\right)\right) =$$
$$= q_i \left(\frac{p_b}{b-j} f_i(D_{b,j}) - \frac{p_j}{b-j} f_i(D_{b,j}) + F_i(D_{b,j}) - F_i(D_{j,c}) - \frac{p_j}{j-c} f_i(D_{j,c}) + \frac{p_c}{j-c} f_i(D_{j,c}) \right)$$

$$= q_i \left(D_{b,j} f_i(D_{b,j}) - D_{j,c} f_i(D_{j,c}) + F_i(D_{b,j}) - F_i(D_{j,c}) \right).$$

By Lemma 12, $D_{b,j} = D_{j,c}$, and therefore these terms are zero. The cases when $i = i^*$, or $i^* = a, b, j$, or $c, d = 0$, or $j \in \mathcal{P}_{\sigma'}(i)$ are identical. □

We are now ready to prove the main theorem of this section, which is simply arguing how this agreement of gradients implies that Rev is concave everywhere.

Proof of Theorem 2. Consider any two prices $\mathbf{p}_1$ and $\mathbf{p}_2$, and the line segment joining the two. We will argue that Rev is concave along this line segment, which then implies the Theorem. From Lemma 9 and Lemma 10, we have that this line segment is itself divided into many intervals (corresponding to the different Δ_σ), and within each interval, Rev is a concave function. Further, from Lemma 11, we have that these concave functions agree at the intersections of the intervals, and the gradients agree too. Thus Rev is smooth, and the derivative along this line is monotone. This implies that Rev is concave along the line. □

4.2. Optimal prices with two demands

In this section we identify optimal prices with two demands $d \in \{1, 2\}$, assuming that the conditional distributions have DMR and identical highest values.[11] By Theorem 1, optimal mechanisms are identified by the optimal prices $\mathbf{p}^* = (p_1^*, p_2^*)$ where p_i^* is the price of i units, for $i \in \{1, 2\}$.

The proposition below relates optimal prices $\mathbf{p}^*$ to the optimal monopoly prices $\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2)$, where $\hat{p}_1 \in \arg\max R_1(p_1)$ and $\hat{p}_2 \in \arg\max R_2(p_2/2)$. The optimal monopoly prices are the solution to a relaxed problem of

$$\max_{p_1, p_2} q_1 p_1 (1 - F_1(p_1)) + q_2 p_2 (1 - F_2(p_2/2)), \tag{16}$$

in which the incentive constraints between types with different demands are removed. That is, price p_1 is offered for buying 1 unit exclusively to all types $(v, 1)$, and price p_2 is offered for buying 2 units exclusively to all types $(v, 2)$ (a type $(v, 1)$ cannot buy 2 units, and a type $(v, 2)$ cannot buy 1 unit). Notice that a type $(v, 2)$ buys 2 units at price p_2 if $2v \geq p_2$, or $v \geq p_2/2$. The optimal monopoly prices can be calculated straightforwardly given the first order conditions of optimality, $R_1'(\hat{p}_1) = 0$ and $R_2'(\hat{p}_2/2) = 0$. For simplicity we assume that the distributions F_1 and F_2 are strictly regular, that is, $v - \frac{1-F_1(v)}{f_1(v)}$ and $v - \frac{1-F_2(v)}{f_2(v)}$ are increasing in v. This assumption implies that $\hat{\mathbf{p}}$ is unique.

The proposition relates the optimal prices to the optimal monopoly prices by considering three cases, shown in Fig. 3. If $\hat{p}_1 \leq \hat{p}_2 \leq 2\hat{p}_1$, that is $\hat{\mathbf{p}}$ belongs to the middle region in Fig. 3, then we show that the revenue Rev($\hat{\mathbf{p}}$) of posting prices $\hat{\mathbf{p}}$ is equal to the optimal revenue of the relaxed problem, (16). In this case, prices $\hat{p}_1$ and $\hat{p}_2$ are optimal, $\mathbf{p}^* = \hat{\mathbf{p}}$. In the other two cases, $\hat{\mathbf{p}} \neq \mathbf{p}^*$. Nevertheless, $\hat{\mathbf{p}}$ can be used as a guide to identify $\mathbf{p}^*$. If $2\hat{p}_1 < \hat{p}_2$, that is $\hat{\mathbf{p}}$ belongs to the top region in Fig. 3, then so does $\mathbf{p}^*$. In this case, it is optimal to offer a price $p_1^* \in [\hat{p}_1, \hat{p}_2/2]$ for one unit, and $p_2^* = p_1^* + \hat{p}_2/2$ for two units. If $\hat{p}_2 < \hat{p}_1$, that is $\hat{\mathbf{p}}$ belongs to the bottom region in Fig. 3, then so does $\mathbf{p}^*$. In this case, it is optimal to offer a uniform "all you can eat" price which is between the two optimal monopoly prices, that is, $p_1^* = p_2^* \in [\hat{p}_2, \hat{p}_1]$. In each case, the optimal prices are identified uniquely given the appropriate first order conditions. The three cases are summarized in the proposition below.

Proposition 2. *Assume that there are two demands $d \in \{1, 2\}$ and that the conditional distributions are strictly regular and have DMR. Consider prices $\mathbf{p}^*$ defined given the three possible cases for $\hat{\mathbf{p}}$. A mechanism that posts prices $\mathbf{p}^*$ is optimal.*

1. *If $\hat{p}_1 \leq \hat{p}_2 \leq 2\hat{p}_1$, then $\mathbf{p}^* = \hat{\mathbf{p}}$.*
2. *If $2\hat{p}_1 < \hat{p}_2$, then p_1^* is identified by $q_1 R_1'(p_1^*) + q_2 R_2'(p_1^*) = 0$ and $p_2^* = p_1^* + \hat{p}_2/2$. The price p_1^* is in the interval $[\hat{p}_1, \hat{p}_2/2]$.*
3. *If $\hat{p}_2 < \hat{p}_1$, then p_1^* is identified by $q_1 R_1'(p_1^*) + 2q_2 R_2'(p_1^*/2) = 0$ and $p_2^* = p_1^*$. The prices $p_1^* = p_2^*$ are in the interval $[\hat{p}_2, \hat{p}_1]$.*

As an application, consider the following three cases. First, suppose that $F_1 = F_2$. Then $2\hat{p}_1 = \hat{p}_2$ and thus we are in case (1) of the proposition. The price of two units is twice as much as the price of one unit. In other words, the buyer can buy any number of units at a constant per-unit rate $\hat{p}_1 = \hat{p}_2/2$. Second, suppose that the hazard rate of F_1 is dominated by the hazard rate of F_2, that is,

$$\frac{1 - F_1(v)}{f_1(v)} < \frac{1 - F_2(v)}{f_2(v)},$$

for all v. Then, $2\hat{p}_1 < \hat{p}_2$ and thus we are in case (2).[12] Third, suppose that the hazard rates satisfy

[11] The analysis extends straightforwardly to the case of any two demands $\mathcal{D} = \{d_1, d_2\}$.

[12] This is because $\hat{p}_1 = \frac{1-F_1(\hat{p}_1)}{f_1(\hat{p}_1)} < \frac{1-F_2(\hat{p}_1)}{f_2(\hat{p}_1)}$ and $\hat{p}_2/2 = \frac{1-F_2(\hat{p}_2/2)}{f_2(\hat{p}_2/2)}$. Since F_2 is strictly regular, $\hat{p}_1 < \hat{p}_2/2$.

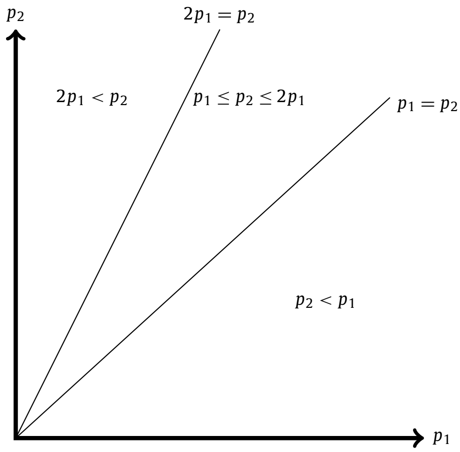

Fig. 3. Three possible cases for prices **p**.

$$\frac{1-F_1(v)}{2f_1(v)} > \frac{1-F_2(v/2)}{f_2(v/2)},$$

for all v. Then, $\hat{p}_2 < \hat{p}_1$ and thus we are in case (3).[13] In case (3), it is optimal to offer a uniform "all you can eat" price.

We defer the proof to Appendix A.6, and here provide an overview. The key step is to show that $\hat{\mathbf{p}}$ and $\mathbf{p}^*$ belong to the same region in Fig. 3. Knowing the region to which $\mathbf{p}^*$ belongs pins down the form of the revenue function at $\mathbf{p}^*$. This is because each region corresponds to a set of price in Δ_σ for some σ, as defined in Equation (14), and thus $\mathrm{Rev} = \mathrm{Rev}_\sigma$. As a result, the optimal prices can be calculated by the first order conditions of optimality of Rev_σ.

To show that $\hat{\mathbf{p}}$ and $\mathbf{p}^*$ belong to the same region, we use Lemma 11 and Theorem 2. By Lemma 11, $\nabla \mathrm{Rev}(\mathbf{p}) = \nabla \mathrm{Rev}_\sigma(\mathbf{p})$ for all prices $\mathbf{p}$ such that $2p_1 = p_2$. If $2\hat{p}_1 < \hat{p}_2$, then the gradient of Rev at a price $2p_1 = p_2$ is pointing to the top region. Since the gradients are equal, the gradient of Rev_σ points to the top region as well. Since Rev is concave, the prices $\mathbf{p}^*$ that maximize Rev must be in the top region. A similar argument applies to the case where $\hat{p}_2 < \hat{p}_1$.

4.3. Computation of optimal prices

We have seen that the revenue as a function of prices of deterministic allocations is concave. In this subsection we discuss how to use this result to find optimal mechanisms in a computationally efficient manner. We do so by applying the Ellipsoid method for convex optimization. The Ellipsoid method can be used to maximize a concave objective function over a convex set, assuming access to two "oracles" that provide answers to certain queries. The first-order oracle gives the value of the objective function and its subgradient at a given point. The separation oracle verifies whether a given point is in the convex set. Further, if the point is not in the set, then the oracle returns a hyperplane that separates the convex set from the point. The running time of the Ellipsoid method depends on the running times of the oracles. See Grötschel et al. (1988) for background.

Theorem 3 *(Ellipsoid for convex optimization. Grötschel et al. (1988)). There is an algorithm such that given (1) a first-order oracle for a concave function $f:\mathbb{R}^n \to \mathbb{R}$, (2) a separation oracle for a convex set $K \subseteq \mathbb{R}^n$, (3) numbers $r, R > 0$ such that $K \subseteq B(0, R)$ and K contains a Euclidean ball of radius r, (4) bounds ℓ_0, u_0 such that $\forall x \in K$, $\ell_0 \le f(x) \le u_0$ and (5) an $\epsilon > 0$, outputs a point $\hat{x} \in K$ such that $f(\hat{x}) \ge f(x) - \epsilon$, for all $x \in K$. The running time of the algorithm is*

$$O\left((n^2 + T_K + T_f)\cdot n^2 \cdot log\left(\frac{R}{r}\cdot\frac{u_0-\ell_0}{\epsilon}\right)^2\right),$$

where T_K and T_f are the running times of the separation oracle for K and a first-order oracle for f, respectively.

We will apply this theorem with the obvious parameters: $f = \mathrm{Rev}$, $n = k$, $K = \{\mathbf{p} : 0 \le p_1 \le \cdots \le p_k \le \bar{v}\}$. $R = \sqrt{k}\bar{v}$, $r = \bar{v}/2k$, $\ell_0 = 0$ and $u_0 = k\bar{v}$.[14] Designing a separation oracle for K is trivial: we can easily check if $\mathbf{p} \in K$, and if $\mathbf{p} \notin K$ the separating hyperplane is given by the violating pair $i < j$ such that $p_i > p_j$. The final piece is getting a *first-order oracle* for Rev, a function that given $\mathbf{p}$ outputs $\mathrm{Rev}(\mathbf{p})$ and a subgradient of Rev at $\mathbf{p}$. Towards this step, notice that our proof of

---

[13] This is because $\hat{p}_2/2 = \frac{1-F_2(\hat{p}_2/2)}{f_2(\hat{p}_2/2)} < \frac{1-F_1(\hat{p}_2)}{2f_1(\hat{p}_2)}$ and $\hat{p}_1 = \frac{1-F_1(\hat{p}_1)}{f_1(\hat{p}_1)}$. Since F_1 is strictly regular, $\hat{p}_2 < \hat{p}_1$.

[14] A ball centered at $(\frac{\bar{v}}{2k}, \frac{3\bar{v}}{2k}, \ldots, \frac{(2k-1)\bar{v}}{2k})$ with radius $\bar{v}/2k$ is contained in K.

Theorem 2 gives us a lot more than concavity. Given a price vector $\mathbf{p}$ we can easily find a closed form for the revenue of $\mathbf{p}$, that is, find $\mathrm{Rev}_\sigma(\mathbf{p})$ as in Equation (15). This allows us to not only calculate the revenue of price vector $\mathbf{p}$, but also the gradient of the revenue function at $\mathbf{p}$. The following corollary is immediate.

Corollary 1. *There is an algorithm such that given $\epsilon > 0$ finds a feasible price vector $\mathbf{p}$, such that $\mathrm{Rev}(\mathbf{p}) \geq \mathrm{Rev}(\hat{\mathbf{p}}) - \epsilon$, for all feasible price vectors $\hat{\mathbf{p}}$, in time*

$$O\left((k^2 + k + T_{\mathrm{Rev}}) \cdot k^2 \cdot \log\left(\frac{k\bar{v}}{\epsilon}\right)^2\right),$$

where T_{Rev} is the running time of the first-order oracle for Rev.

There is one finer point left to address. We have assumed that we can perform exact arithmetic on real numbers. Specifically, we assumed that our separating and first-order oracles give exact answers. Luckily, even if we got approximate answers, say up to ℓ bits, which in turn translate to weak oracles, we can still find an approximately optimal price vector in polynomial time. For more details and precise statements we refer the reader to Grötschel et al. (1988).

5. Dynamic pricing

We have assumed so far that the seller knows the distribution of types. A learning theoretic approach to the problem is to make the same assumptions about the distributions as before, but do not assume that the seller knows these distributions. Instead, the buyers arrive one at a time, and the seller can adaptively change the prices he offers, as she observes more and more of the buyers respond, and converge to an optimal price. This is commonly called the dynamic pricing problem; see den Boer (2015) for a survey. We define the model formally below.

In each round $\tau \in 1, 2, \ldots, T$, for some $T \in \mathbb{Z}_+$, the following takes place. (We state this model and the results for the more general case where the demand $d \in \{d_1, d_2, \cdots, d_k\}$ for arbitrary d_i.)

1. The seller posts a price vector $\mathbf{p}^\tau$.
2. A buyer of type (v^τ, d^τ) is drawn independently from the distribution f.
3. The buyer buys her utility maximizing bundle $x^\tau \in \arg\max_{\{i: d_i \leq d^\tau\}} v^\tau d_i - p_i^\tau$.
4. The seller observes only x^τ.

Assume, for the sake of notational convenience, that $d_0 = 0$ and $p_0^\tau = 0$ for all τ, so $x^\tau = 0$ when the buyer does not buy anything. The goal of the seller is to maximize her expected average revenue

$$\mathbb{E}\left[\frac{1}{T}\sum_{\tau=1}^{T} p_{x^\tau}^\tau\right].$$

We evaluate the performance of a dynamic pricing scheme by its *regret*, which is the difference between the optimal expected revenue, if the seller knew the type distribution, and the average expected revenue of the pricing scheme. We would like the regret to converge to zero, and evaluate an algorithm based on the rate of convergence.

Kleinberg and Leighton (2003) present an algorithm with a $1/\sqrt{T}$ regret for the case of a *single* item, assuming that the revenue function is concave (same assumption as ours).[15] When there are multiple items, without any additional assumptions, the regret scales as $\sqrt{L/T}$ where L is the number of different price vectors to be considered. This requires a discretization of the price space, and the resulting L is exponential in k, the number of different items sold. Besbes and Zeevi (2012) show that if the revenue function is concave, then there is an algorithm with regret $\sqrt{\log T / T^{k+3}}$. The concavity assumption on the revenue function is quite common in such revenue management problems (Talluri and Van Ryzin, 2006). Theorem 2 shows that concavity of the revenue function is implied by the concavity of 1 dimensional revenue functions for each d. The problem of convex bandits (defined formally below) has received much attention lately. The state of the art for this problem is a $\frac{k^{9.5}}{\sqrt{T}}$ regret algorithm by Bubeck et al. (2017). This is a significant improvement over previous results since the dependence on k goes from exponential to polynomial.

The *convex bandits* problem is as follows: in each round $\tau \in 1, 2, \ldots, T$, for some $T \in \mathbb{Z}_+$, the following takes place.

1. The algorithm picks a vector $\mathbf{p}^\tau$ from some convex domain $K \subseteq \mathbb{R}^k$.
2. An adversary picks an arbitrary convex loss function ℓ^τ on domain K and range $[0, 1]$.
3. The algorithm observes only $\ell^\tau(\mathbf{p}^\tau)$.

[15] When we state regret bounds, we ignore the constants. All regret bounds should be interpreted with an $O(1)$ in front.

The goal of the algorithm is to minimize regret, defined as

$$\sum_{\tau=1}^{T} \ell^\tau(\mathbf{p}^\tau) - \min_{\mathbf{p}\in K} \ell^\tau(\mathbf{p}).$$

The reduction from the dynamic pricing problem to the convex bandits problem is quite standard. Let $R^\tau(\mathbf{p})$ denote the revenue obtained from the buyer of type (v^τ, d^τ) when offered a price $\mathbf{p}$. Define the loss function at time τ to be

$$\ell^\tau(\mathbf{p}) = 1 - \frac{R^\tau(\mathbf{p})}{d_k \bar{v}}.$$

Recall that $\bar{v}$ is the highest possible marginal value. The loss has been normalized appropriately so that the range is $[0, 1]$. The regret of the convex bandit problem is the same as that of the dynamic pricing problem, up to a factor of $d_k\bar{v}$.

This is not quite a reduction to the convex bandits problem because this loss function is not convex. What we have from Theorem 2 is that the expectation of the loss function is convex. In personal communication with the authors of the convex bandits paper (Bubeck, 2017), we have confirmed that their result also holds for this case, where the loss function is randomized and is only convex in expectation. Thus we get the following as a corollary of Theorem 2 and Theorem 1 in Bubeck et al. (2017).

Corollary 2. *There is a dynamic pricing scheme where the regret is*

$$\frac{\tilde{O}(k^{9.5}) d_k \bar{v}}{\sqrt{T}}.$$

6. Conclusions and discussions

In this paper, we consider a mechanism design problem with multi unit demands, where the marginal value of each type of the buyer (v, d) is v for the first d units it receives, and 0 afterwards. We prove two main results. First, if the conditional distributions F_d have DMR and identical highest values, then the optimal mechanism is deterministic. Second, if the conditional distributions have DMR, then revenue as a function of vector of prices is concave, and therefore optimal deterministic prices can be computed efficiently. We use concavity to establish two further results. We identify optimal deterministic prices when there are two demands, and provide an approximately optimal dynamic pricing scheme with demand uncertainty. In the subsequent subsections we discuss the DMR condition and briefly discuss possible extensions and further directions.

6.1. The DMR condition

In this section we discuss the DMR condition and its connections with other commonly used regularity conditions.

A simple class of DMR distributions is Uniform $[a, b]$ for any non-negative reals a and b. More generally, any distribution with bounded support and monotone non-decreasing probability density is DMR. The second derivative of the revenue function is $-2f(v) - vf'(v)$, which is negative if $f'(v) \geq 0$. Another standard class of demand distributions that satisfies DMR is a constant elasticity distribution, defined in the example below. The DMR condition is closely related to but different from the regularity condition of Myerson (1981). Whereas regularity requires that the function $\phi(v) = v - \frac{1-F(v)}{f(v)}$ is monotone non-decreasing, DMR requires that $\phi(v) \cdot f(v)$ is monotone non-decreasing. The example below shows that DMR and regularity are incomparable conditions.

Example 5 *(DMR vs. regularity).* Consider the class of constant elasticity distributions with cumulative density $F(v) = 1 - (v/a)^{1/\epsilon}$ for any $a > 0$ and $\epsilon < 0$, supported on $[a, \infty)$. A special case is when $a = 1$ and $\epsilon = -1$, in which case $F(v) = 1 - 1/v$, known as the *equal revenue distribution*. The corresponding revenue function $v(v/a)^{1/\epsilon}$ is concave if $\epsilon \leq -1$. However, the function $\phi(v) = v - \frac{1-F(v)}{f(v)}$ simplifies to $v(1 + \epsilon)$, which is monotone *decreasing* for $\epsilon < 1$. Therefore such a distribution satisfies DMR, but not regular, for $\epsilon < -1$. Notice that the truncation of the distribution on any support $[a, \bar{v}]$ also satisfies DMR. On the other hand, a truncated exponential distribution is regular but does not satisfy DMR. Calculations for this example are straightforward and deferred to Appendix A.

The class of DMR distributions is well-behaved in the sense that it is closed under convex combinations. In particular, the distribution that results from drawing a sample from a DMR distribution with probability α, and from another DMR distribution with probability $1 - \alpha$, is a DMR distribution. The cumulative density of a distribution that samples from F_1 with probability α, and from F_2 otherwise, is $F(v) = \alpha F_1(v) + (1 - \alpha)F_2(v)$. Therefore, the revenue function of the convex combination is the convex combination of the revenue functions of F_1 and F_2, and is concave if F_1 and F_2 are DMR. On the other hand, it is known that regular distributions are not closed under convex combinations (Sivan and Syrgkanis (2013)).

The DMR condition has been extensively used in the mechanism design literature. Che and Gale (1998) use the same condition for a similar problem of selling a single item to a single buyer with budget constraints, rather than demand or capacity constraints. The optimal mechanism there could still be randomized. Fiat et al. (2016) too use the DMR condition. In particular, they show that to derive the optimal mechanism, one needs to *iron* in the value space, rather than the quantile space as in Myerson (1981).[16] DMR is precisely when no ironing is needed in the value space. The same assumption was also made by Kleinberg and Leighton (2003) in the context of dynamic pricing. Generalizations of the DMR condition have been used in the literature on multidimensional mechanism design (e.g., Preston McAfee and McMillan (1988)).

6.2. Extensions and future directions

The focus of this paper is to identify when deterministic mechanisms are optimal, and the structure and computation of optimal deterministic mechanisms. There are many natural questions that arise that are beyond the scope of this paper. We briefly mention these questions and possible extensions, and leave a more detailed analysis for future work.

Approximation In this paper we focus on exact optimality of mechanisms (other than the dynamic pricing extension). It would be interesting to explore conditions under which deterministic mechanisms are approximately optimal. An immediate approach would to formulate approximate notions of our conditions that imply approximate optimality of deterministic mechanisms. Other papers, for instance Cole and Roughgarden (2014), have developed approximate notions of regularity (and monotone hazard rate) that can perhaps be used to define approximate DMR. Fiat et al. (2016) develop an ironing technique to characterize optimal mechanisms in their setting, which may be useful in our setting to identify optimal mechanisms, or to prove approximate optimality of deterministic mechanisms given approximate DMR.

We make two assumptions regarding the support of the distributions, that perhaps can be relaxed to get approximate optimality. First, that the distributions have bounded support. It would be perhaps possible to replace the assumption of identical highest values with a "diminishing tails" assumption, for instance if $vf(v)$ goes to zero as v goes to infinity. The problem with directly extending our current approach is that it is not clear how Lemma 5 should be extended to allow for unbounded supports. Second, that the distributions have identical highest values. Example 4 show that it is not possible to remove this assumption. But it may be possible to develop an approximation guarantee based on how close the highest values are in the supports of the distributions.

Multi-unit multi-agent auctions There are two possible approaches that one can take in studying a multi-agent extension of our model. First, one can attempt to characterize optimal auctions. This is done by Malakhov and Vohra (2009) (with a regularity assumption and assuming that buyers cannot report a higher demand). A second approach, closer to the one taken in this paper, is to identify conditions under which a "simple" class of auctions is optimal. A natural extension of our work would be to identify conditions under which deterministic mechanisms are optimal, although there may be other possible generalization of deterministic mechanisms to multiple agents.

Appendix A. Deferred proofs

A.1. Calculations from Example 1

Lemma 13. *In Example 1, the optimal profit from deterministic mechanisms is* $\frac{7}{3}$.

Proof. A deterministic mechanism posts prices p_1, p_2 and p_3 for buying 1, 2 or 3 units respectively; an agent picks her favorite option. Clearly, $p_1 \leq p_2 \leq p_3$.

If $p_1 > 3$ (which implies that $p_2, p_3 > 3$) t_3 is the only agent who could purchase something. The optimal choice of p_1 is then 6, which yields revenue $\frac{6}{3} = 2$.

If $p_1 \leq 3$, then $p_1 = p_2$. Assume otherwise that $p_1 < p_2$. If any of t_1, t_2 choose to buy 1 unit for a price of p_1, then they would also buy 2 units for a price of $2p_1$ (which yields strictly larger revenue), and t_3 would also have non-negative utility for this option. Therefore, increasing p_1 to equal p_2 will weakly increase revenue.

Assume therefore that $p_1 \leq 3$, then $p_1 = p_2$. Since $p_1 = p_2$, t_3 will always purchase 2 units and contribute $\frac{p_2}{3}$ to the expected revenue. If $p_1 = p_2 > 2$, then t_2 does not purchase anything and t_1 could only purchase 3 units (if $p_3 \leq 3$). The revenue in this case is $\frac{p_2}{3} + \frac{p_3}{3}$ if $p_3 \leq 3$, for a maximum of 2 (by setting $p_2 = p_3 = 3$), or $\frac{p_2}{3}$ if $p_3 > 3$, for a maximum of 2 (by setting $p_2 = p_3 = 6$). Finally, if $p_2 \leq 2$, then t_2 buys 2 units for a price of p_2. If $p_3 > p_2 + 1$, t_3 prefers 2 units to 3 units and the expected revenue is $p_2 \cdot 1$ (maximized at $p_2 = 2$). Otherwise, t_3 buys 3 units for a price of p_3; setting $p_2 = 2$ and $p_3 = 3$ gives expected revenue $p_2 \frac{2}{3} + \frac{p_3}{3} = \frac{7}{3}$. □

[16] Ironing is a technique introduced by Myerson (1981) where the virtual value function is transformed so that it becomes monotone. This corresponds to transforming the corresponding revenue function into a concave function.

A.2. Strong individual rationality

According to definition, a mechanism (A, p) is IR if

$$\mathbb{E}\left[v \min\{A(v,d), d\} - p(v,d)\right] \geq 0.$$

A stronger individual rationality constraint would be to require the utility of each type to be non-negative for any random choice of the mechanism. We show that any IR mechanism can be converted to one that satisfies this stronger IR requirement. To define the IR notion, we make the random choices of the mechanism regarding payments explicit. A mechanism (A, P) is identified by an allocation rule A and a payment rule $P : \mathcal{V} \times \mathcal{D} \to \Delta(\mathbb{R})$. A mechanism is *strongly IR* if $v \min\{A(v,d), d\} - P(v,d) \geq 0$ for all types (v,d) and all random choice of the mechanism. Incentive compatibility is defined as before.

Lemma 14. *For every IC and IR mechanism (A, P), there exists $\tilde{P}$ such that the mechanism $(A, \tilde{P})$ is IC and strongly IR and $\mathbb{E}[P(v,d)] = \mathbb{E}[\tilde{P}(v,d)]$ for all (v,d), where the expectation is taken over the randomization of the mechanism.*

Proof. Consider an IC and IR mechanism (A, P), with expected payment $p(v,d) = \mathbb{E}[P(v,d)]$. First note that we can assume that for each type (v,d), the randomized allocation $A(v,d)$ does not assign a number of units more than d. If this is not true, replace any assignment of more than d units with the assignment of d units. Note that this change does not change the utility of truthful reporting, and cannot improve utility of non-truthful reporting. Therefore the resulting mechanism is IC and IR. Now consider a type (v,d). For each *realization* of allocation $A(v,d)$, define a payment $\tilde{P}(v,d)$ as follows

$$\tilde{P}(v,d) = \frac{p(v,d)A(v,d)}{\mathbb{E}[A(v,d)]}.$$

The payment rule $\tilde{P}$ is randomized since A is randomized. Note that the expected payment of the type stays the same,

$$\mathbb{E}\left[\tilde{P}(v,d)\right] = p(v,d)\frac{\mathbb{E}[A(v,d)]}{\mathbb{E}[A(v,d)]} = p(v,d).$$

As a result, the modified mechanism remains IC. In addition, the ex-post utility of the type from the realized allocation of $A(v,d)$ units is

$$vA(v,d) - \frac{p(v,d)A(v,d)}{\mathbb{E}[A(v,d)]},$$

which is non-negative if and only if

$$v\,\mathbb{E}[A(v,d)] - p(v,d) \geq 0,$$

which hold by IR. □

A.3. Calculations for Example 3 and Example 4

We here provide calculations for Example 3 and Example 4. The idea is straightforward. If a distribution is perturbed slightly, the optimal revenue from randomized mechanisms and the optimal revenue from deterministic mechanisms do not change by much. The distributions in Example 3 and Example 4 are small perturbations of Example 1. Since there is a gap between the optimal revenue and the optimal revenue from deterministic mechanisms Example 1, such a gap will persist for small enough perturbations.

Madarász and Prat (2017) prove the small perturbations argument. We here provide the calculations regardless since the arguments for deterministic mechanisms does not appear in Madarász and Prat (2017). Instead of attempting a general approach (which is probably possible), we argue each case individually.

Calculations for Example 3 Consider the randomized mechanism of Example 1 and discount the prices by 3ϵ and 2ϵ for options (a) and (b) respectively. That is, charge $3 - 3\epsilon$ for 3 units, and charge $2 - 2\epsilon$ for a probability of $3/4$ to receive 2 units. This results in all types (v,d) purchasing the option the unique type with demand d purchased in Example 1. By picking σ small enough, the revenue of this (randomized) mechanism on this new distribution is at least $\frac{7.5}{3} - 4\epsilon$.

It remains to show that no deterministic mechanism for this distribution has revenue much larger than the optimal deterministic mechanism in Example 1. Suppose that there was a deterministic mechanism M that gave revenue at least $\frac{7.1}{3}$ for the new distribution. We will construct a deterministic mechanism $\tilde{M}$ that gives revenue strictly larger than $\frac{7}{3}$ for the distribution of Example 1, leading to a contradiction. First, note that we can safely ignore all types (v,d) with $|v - \mu| > \delta$ for any $\delta > 0$, where μ is the mean of the normal distribution corresponding to d, by picking σ small enough; even if those types paid their value, the contribution to the overall revenue would be insignificant. Now, consider a mechanism $\hat{M}$

that assigns all types (v, d) the most expensive option purchased by any type (v', d), with $|v' - \mu| \le \delta$ (this mechanism is not truthful). The revenue of $\hat{M}$ is at least the revenue of M. Furthermore, the extra utility of any type in mechanism $\hat{M}$ from deviating to another type is at most 2δ. Now, notice that mechanism $\hat{M}$ has the same revenue in the distribution of Example 1 as in the distribution of Example 3.

To complete the proof, consider a mechanism $\tilde{M}$ that reduces all prices of $\hat{M}$ by a factor of $(1-\epsilon)$, for some $\epsilon > 0$, and lets buyers choose whatever option they want. $\tilde{M}$ is truthful and deterministic by definition. We claim that its revenue is at least $\frac{7.1}{3} - \epsilon/(2\delta)$, which is a contradiction, since there are choices of ϵ and δ to make this number strictly larger than $\frac{7}{3}$. To see this most clearly, consider a type (v, d) that is gets allocation x for a price p in $\hat{M}$, but chooses some allocation x' for a price $(1-\epsilon)p'$ in $\tilde{M}$. As argued before, the utility of type (v, d) from allocation x and payment $(1-\epsilon)p$ plus 2δ is at least the utility of type (v, d) for allocation x' for price p'. Since $\tilde{M}$ is truthful, we have that the utility of type (v, d) for allocation x' and payment $(1-\epsilon)$ is at least the utility of type (v, d) for allocation x and payment $(1-\epsilon)p$. Adding up the two equations gives $p' \ge p - \frac{\epsilon}{2\delta}$, which gives the desired result when adding over all types.

Calculations for Example 4 The calculations for Example 4 are similar to those of Example 3. First, the optimal revenue is at least $\frac{7.5}{3} - 4\epsilon$. Second, the revenue of any deterministic mechanism is at most $\frac{7.1}{3}$. Therefore, for ϵ small enough, deterministic mechanisms are not optimal.

Suppose that there was a deterministic mechanism M that gave revenue at least $\frac{7.1}{3}$ for the new distribution. We will construct a deterministic mechanism $\tilde{M}$ that gives revenue strictly larger than $\frac{7}{3}$ for the distribution of Example 1, leading to a contradiction. Now, consider a mechanism $\hat{M}$ that assigns all types (v, d) the most expensive option purchased by any type (v', d) (this mechanism is not truthful). The revenue of $\hat{M}$ is at least the revenue of M. Furthermore, the extra utility of any type in mechanism $\hat{M}$ from deviating to another type is at most 2ϵ. Now, notice that mechanism $\hat{M}$ has the same revenue in the distribution of Example 1 as in the distribution of Example 3.

To complete the proof, consider a mechanism $\tilde{M}$ that reduces all prices of $\hat{M}$ by a factor of $(1-\delta)$, for some $\delta > 0$, and lets buyers choose whatever option they want. $\tilde{M}$ is truthful and deterministic by definition. We claim that its revenue is at least $\frac{7.1}{3} - 2\epsilon\delta$, which is a contradiction, since there are choices of ϵ and δ to make this number strictly larger than $\frac{7}{3}$. To see this, consider a type (v, d) that is gets allocation x for a price p in $\hat{M}$, but chooses some allocation x' for a price $(1-\delta)p'$ in $\tilde{M}$. As argued before, the utility of type (v, d) from allocation x and payment $(1-\delta)p$ plus 2ϵ is at least the utility of type (v, d) for allocation x' for price p'. Since $\tilde{M}$ is truthful, we have that the utility of type (v, d) for allocation x' and payment $(1-\delta)$ is at least the utility of type (v, d) for allocation x and payment $(1-\delta)p$. Adding up the two equations gives $p' \ge p - 2\epsilon\delta$, which gives the desired result when adding over all types.

A.4. Proof of Lemma 2

Lemma 2. *A mechanism (w, p) with indirect utility U defined in (3) is IC if and only if*

1. *U_d is convex and $U_d'(v) = d w_d(v)$ if U_d' exists, and*
2. *$U_d(v) \ge U_{d-1}(v)$, and*
3. *$U_d(v) \ge U_{d+1}(v\frac{d}{d+1})$.*

Proof. The equivalence of property 1 of the lemma with property 1 of Theorem 1 is standard and is omitted.

Property 2 of the lemma is equivalent to property 2 of Theorem 1. By definition, $u(v, d \to v, d) = U_d(v)$, and $u(v, d \to v, d-1) = v(d-1)w_{d-1}(v) - p(v, d-1) = U_{d-1}(v)$. Thus, $u(v, d \to v, d) \ge u(v, d \to v, d-1)$ is equivalent to $U_d(v) \ge U_{d-1}(v)$.

Property 3 of the lemma follows from properties 1 and 3 of Theorem 1. This is because by the payment identify (4),

$$\begin{aligned} u(v, d \to v\frac{d}{d+1}, d+1) &= v d w_{d+1}(v\frac{d}{d+1}) - p(v\frac{d}{d+1}, d+1) \\ &= (d+1)\int_{\underline{v}_{d+1}}^{v\frac{d}{d+1}} w_{d+1}(z)dz + U_d(\underline{v}_d) \\ &= U_{d+1}(v\frac{d}{d+1}). \end{aligned}$$

Thus $u(v, d \to v, d) \ge u(v, d \to v\frac{d}{d+1}, d+1)$ is equivalent to $U_d(v) \ge U_{d+1}(v\frac{d}{d+1})$.

The second statement of the lemma regarding IR is identical to the second statement of Proposition 1. $\square$

A.5. Proof of Lemma 3

Lemma 3. *If a mechanism (w, p) is IC, then the expected revenue is*

$$\mathbb{E}\left[p(v,d)\right] = \sum_d \Big(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) R''_d(v) dv \Big) q_d.$$

Proof. Note for future reference that using integration by parts, we can write

$$\begin{aligned}\int_{\underline{v}_d}^{\bar{v}_d} v f_d(v) U'_d(v) dv &= v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} - \int_{\underline{v}_d}^{\bar{v}_d} \frac{d}{dv}(v f_d(v)) U_d(v) dv \\ &= v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} - \int_{\underline{v}_d}^{\bar{v}_d} \frac{d}{dv}(v f_d(v)) U_d(v) dv. \end{aligned} \tag{A.1}$$

The expected revenue is

$$\sum_d \Big(\int_{\underline{v}_d}^{\bar{v}_d} p(v,d) f_d(v) dv \Big) q_d = \sum_d \Big(\int_{\underline{v}_d}^{\bar{v}_d} (v d w_d(v) - U_d(v)) f_d(v) dv \Big) q_d,$$

substituting $U'_d(v) = d w_d(v)$ from Lemma 2,

$$= \sum_d \Big(\int_{\underline{v}_d}^{\bar{v}_d} (v U'_d(v) f_d(v) - U_d(v) f_d(v)) dv \Big) q_d,$$

using integration by parts (A.1),

$$= \sum_d \Big(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) (-\frac{d}{dv}(v f_d(v)) - f_d(v)) dv \Big) q_d,$$

since $R_d(v) = v(1 - F_d(v))$, we have $R''_d(v) = -\frac{d}{dv}(v f_d(v)) - f_d(v)$ and thus the expected revenue is

$$= \sum_d \Big(v f_d(v) U_d(v)|_{\underline{v}_d}^{\bar{v}_d} + \int_{\underline{v}_d}^{\bar{v}_d} U_d(v) R''(v) dv \Big) q_d. \quad \square$$

A.6. Proof of Proposition 2

Proposition 2. *Assume that there are two demands $d \in \{1, 2\}$ and that the conditional distributions are strictly regular and have DMR. Consider prices $\mathbf{p}^*$ defined given the three possible cases for $\hat{\mathbf{p}}$. A mechanism that posts prices $\mathbf{p}^*$ is optimal.*

1. *If $\hat{p}_1 \le \hat{p}_2 \le 2\hat{p}_1$, then $\mathbf{p}^* = \hat{\mathbf{p}}$.*
2. *If $2\hat{p}_1 < \hat{p}_2$, then p_1^* is identified by $q_1 R'_1(p_1^*) + q_2 R'_2(p_1^*) = 0$ and $p_2^* = p_1^* + \hat{p}_2/2$. The price p_1^* is in the interval $[\hat{p}_1, \hat{p}_2/2]$.*
3. *If $\hat{p}_2 < \hat{p}_1$, then p_1^* is identified by $q_1 R'_1(p_1^*) + 2q_2 R'_2(p_1^*/2) = 0$ and $p_2^* = p_1^*$. The prices $p_1^* = p_2^*$ are in the interval $[\hat{p}_2, \hat{p}_1]$.*

Proof. For future reference, consider prices p_1 and p_2 such that $p_1 \le p_2 \le 2p_1$. Since the price of two units is weakly higher than one unit, a type $(v, 1)$ buys one unit if $v \ge p_1$ and zero units otherwise. For a type $(v, 2)$, consider two cases. If the utility of buying two units is non-negative, $2v - p_2 \ge 0$, then $2v - p_2 \ge (v + p_2/2) - p_2 \ge v - p_1$, and thus the type buys two units. If the utility of buying two units is negative, $2v - p_2 < 0$, then $v < p_2/2 \le p_1$, and therefore the type buys zero units. To summarize, if $p_1 \le p_2 \le 2p_1$, then a type $(v, 1)$ buys one unit if $v \ge p_1$ and zero units otherwise, and a type $(v, 2)$ buys two units if $v \ge p_2/2$ and zero units otherwise.

Define $\sigma(1) = \sigma(2) = 0$. Using the notation of Section 4.1 (see (14)), we have

$$\Delta_\sigma = \{\mathbf{p} : p_1 \le p_2 \le 2p_1\}.$$

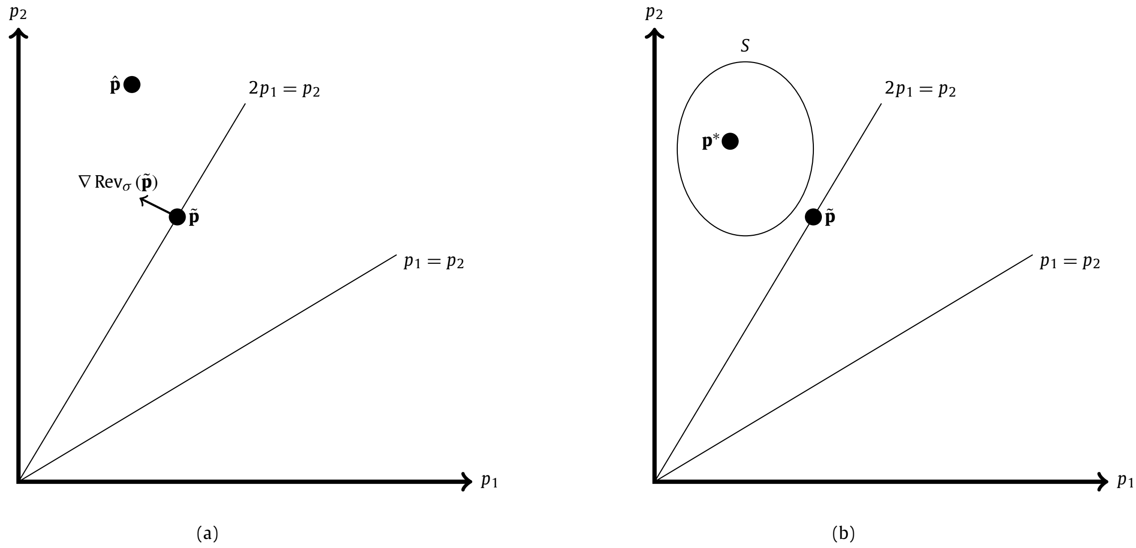

Fig. A.4. (a) The gradient $\nabla \mathrm{Rev}_\sigma(\tilde{\mathbf{p}})$ is orthogonal to the line $2p_1 = p_2$ and points to the left. (b) The set S' of prices $\mathbf{p}$ such that $\mathrm{Rev}(\mathbf{p}) \geq \mathrm{Rev}(\tilde{\mathbf{p}}) + \epsilon$ is non-empty, convex, and separated from the set of all points $\mathbf{p}'$ such that $2p_1' \geq p_2'$ via the line $2p_1 = p_2$.

By Lemma 9, we have

$$\mathrm{Rev}(\mathbf{p}) = \mathrm{Rev}_\sigma(\mathbf{p}) = q_1 R_1(p_1) + q_2 2R_2(p_2/2)$$

for all $p \in \Delta_\sigma$. Notice that given DMR, Rev_σ is concave.

First consider the case where $\hat{p}_1 \leq \hat{p}_2 \leq 2\hat{p}_1$. Thus, $\mathrm{Rev}(\hat{\mathbf{p}}) = \mathrm{Rev}_\sigma(\hat{\mathbf{p}})$ is equal to the optimal revenue in the relaxed problem. Therefore, $\mathbf{p}^* = \hat{\mathbf{p}}$ is optimal.

Now consider the case where $2\hat{p}_1 < \hat{p}_2$. We argue that the optimal prices $\mathbf{p}^*$ satisfy $2p_1^* < p_2^*$. Consider prices $\tilde{\mathbf{p}}$ that maximize $\mathrm{Rev}(\mathbf{p})$ among all prices $\mathbf{p}$ such that $2p_1 = p_2$. Notice that $\mathrm{Rev}(\mathbf{p}) = \mathrm{Rev}_\sigma(\mathbf{p})$ for all such prices. As shown in Fig. A.4, (a), the first order condition implies that the gradient of Rev_σ at $\tilde{\mathbf{p}}$ must be orthogonal to the line $2p_1 = p_2$, that is, $\nabla \mathrm{Rev}_\sigma(\tilde{\mathbf{p}}) = (-2\delta, \delta)$ for some δ. Otherwise a point to the left or right of $\tilde{\mathbf{p}}$ on the line $2p_1 = p_2$ will have higher revenue. Further, the gradient must point to the left, $\delta > 0$, since the global maximizer of Rev_σ, $\hat{\mathbf{p}}$, is to the left. By Lemma 11, $\nabla \mathrm{Rev}(\tilde{\mathbf{p}}) = \nabla \mathrm{Rev}_\sigma(\tilde{\mathbf{p}})$, which implies that $\mathrm{Rev}(\mathbf{p}) = \mathrm{Rev}(\tilde{\mathbf{p}}) + \epsilon$ for some $\mathbf{p}$ that is to the left of the line and some $\epsilon > 0$. Now consider the set $S = \{\mathbf{p} | \mathrm{Rev}(\mathbf{p}) \geq \mathrm{Rev}(\tilde{\mathbf{p}}) + \epsilon\}$, shown in Fig. A.4, (b). The set S is non-empty and convex since Rev is concave. The set S does not include any point on the line $2p_1 = p_2$ since all points in the set have higher revenue than the maximum revenue on the line. Thus, the line $2p_1 = p_2$ is a hyperplane that separates the convex set S from all prices to the right of the hyperplane, and thus $2p_1 < p_2$ for all $\mathbf{p} \in S$. Finally, the set includes $\mathbf{p}^*$. We conclude that $2p_1^* < p_2^*$.

Given that $2p_1^* < p_2^*$, the optimal prices can be calculated using the first order conditions of optimality. Consider any price $\mathbf{p}$ such that $2p_1 < p_2$. A type $(v, 1)$ buys one unit if $v \geq p_1$ and zero units otherwise, and a type $(v, 2)$ buys two units if $v \geq p_2 - p_1$, one unit if $p_1 \leq v \leq p_2 - p_1$, and zero units otherwise. Therefore, the revenue is

$$\mathrm{Rev}(\mathbf{p}) = q_1 R_1(p_1) + q_2\Big(p_1(F_2(p_2 - p_1) - F_2(p_1)) + p_2(1 - F_2(p_2 - p_1))\Big).$$

The first order condition for optimality of p_2^* is

$$1 - F_2(p_2^* - p_1^*) - (p_2^* - p_1^*) f_2(p_2^* - p_1^*) = 0. \tag{A.2}$$

That is, $R_2'(p_2^* - p_1^*) = 0$ and thus $p_2^* - p_1^* = \hat{p}_2/2$. The first order condition for optimality of p_1^* is

$$q_1 R_1'(p_1^*) + q_2\Big(F_2(p_2^* - p_1^*) - F_2(p_1^*) + p_1^*(-f_2(p_2^* - p_1^*) - f_2(p_1^*)) + p_2^* f_2(p_2^* - p_1^*)\Big) = 0.$$

Using (A.2), the condition simplifies to

$$q_1 R_1'(p_1^*) + q_2 R_2'(p_1^*) = 0. \tag{A.3}$$

Notice that by DMR, $q_1 R_1'(p) + q_2 R_2'(p)$ is decreasing in p. In addition, since $2\hat{p}_1 < \hat{p}_2$, we have $q_1 R_1'(\hat{p}_1) + q_2 R_2'(\hat{p}_1) \geq 0$ and $q_1 R_1'(\hat{p}_2/2) + q_2 R_2'(\hat{p}_2/2) \leq 0$. Therefore, $p_1^* \in [\hat{p}_1, \hat{p}_2/2]$.

To complete the proof, consider the case where $\hat{p}_2 < \hat{p}_1$. An argument similar to the one presented above implies that optimal prices $\mathbf{p}^*$ satisfy $p_2^* \leq p_1^*$. For any $\mathbf{p}$ such that $p_2 \leq p_1$, revenue is

$$q_1 R_1(p_2) + 2q_2 R_2(p_2/2).$$

The first order condition for optimality of $\mathbf{p}^*$ is

$$q_1 R_1'(p_2^*) + 2q_2 R_2'(p_2^*/2) = 0.$$

Notice that by DMR, $q_1 R_1'(p) + 2q_2 R_2'(p/2)$ is decreasing in p. In addition, since $\hat{p}_2 < \hat{p}_1$, we have $q_1 R_1'(\hat{p}_2) + 2q_2 R_2'(\hat{p}_2/2) \geq 0$ and $q_1 R_1'(\hat{p}_1) + 2q_2 R_2'(\hat{p}_1/2) \leq 0$. Therefore, $p_1^* = p_2^* \in [\hat{p}_2, \hat{p}_1]$. $\square$

A.7. Calculations from Example 5

We first show that the constant elasticity distribution with cumulative density $F(v) = 1 - (v/a)^{1/\epsilon}$ is DMR. Recall that DMR is equivalent to concavity of the revenue function. To verify concavity, we calculate the second derivate of the revenue function and show that it is negative.

$$\begin{aligned}
R'(v) &= (\frac{v}{a})^{1/\epsilon} + \frac{v}{a\epsilon}(\frac{v}{a})^{1/\epsilon-1}.\\
R''(v) &= (\frac{v}{a})^{1/\epsilon-1}\frac{2}{a\epsilon} + (\frac{v}{a})^{1/\epsilon-2}\frac{v}{a^2\epsilon}(1/\epsilon - 1)\\
&= (\frac{v}{a})^{1/\epsilon-2}(\frac{2v}{a^2\epsilon} + \frac{v}{a^2\epsilon}(1/\epsilon - 1))\\
&= (\frac{v}{a})^{1/\epsilon-2}\frac{v}{a^2\epsilon}(1 + 1/\epsilon) \leq 0.
\end{aligned}$$

Now consider regularity. Note that the probability density function $f(v) = \frac{-1}{\epsilon a}(v/a)^{1/\epsilon-1}$. Recall that a distribution is regular if the function $\phi(v)$ is monotone non-decreasing in v.

$$\begin{aligned}
\phi(v) &= v - \frac{1 - F(v)}{f(v)}\\
&= v - \frac{(v/a)^{1/\epsilon}}{\frac{-1}{a\epsilon}(v/a)^{1/\epsilon-1}}\\
&= v - \frac{v/a}{-1/(a\epsilon)} = v(1 + \epsilon),
\end{aligned}$$

which is monotone *decreasing* since by assumption $\epsilon < -1$.

We finally argue that the exponential distribution, defined as $F(v) = 1 - e^{-v}$ is not DMR but is regular. The revenue function is $R(v) = ve^{-v}$, its first derivative is $R'(v) = (1 - v)e^{-v}$, and its second derivative is $(v - 2)e^{-v}$, which is positive for $v \geq 2$, violating concavity. However, as commonly known, this distribution is regular since $\phi(v) = v - \frac{1-F(v)}{f(v)} = v - \frac{e^{-v}}{e^{-v}} = v - 1$ is monotone non-decreasing in v.

References

Agarwal, Alekh, Foster, Dean P., Hsu, Daniel J., Kakade, Sham M., Rakhlin, Alexander, 2011. Stochastic convex optimization with bandit feedback. In: Advances in Neural Information Processing Systems, pp. 1035–1043.
Ausubel, Lawrence M., Cramton, Peter, Pycia, Marek, Rostek, Marzena, Weretka, Marek, 2014. Demand reduction and inefficiency in multi-unit auctions. Rev. Econ. Stud. 81 (4), 1366–1400.
Besbes, Omar, Zeevi, Assaf, 2012. Blind network revenue management. Oper. Res. 60 (6), 1537–1550.
Börgers, Tilman, 2015. An Introduction to the Theory of Mechanism Design. Oxford University Press, USA.
Briest, Patrick, Chawla, Shuchi, Kleinberg, Robert, Weinberg, S. Matthew, 2015. Pricing lotteries. J. Econ. Theory 156, 144–174.
Bubeck, Sébastien, 2017. Personal communication.
Bubeck, Sébastien, Lee, Yin Tat, Eldan, Ronen, 2017. Kernel-based methods for bandit convex optimization. In: Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing. ACM, pp. 72–85.
Che, Yeon-Koo, Gale, Ian, 1998. Standard auctions with financially constrained bidders. Rev. Econ. Stud. 65 (1), 1–21.
Cole, Richard, Roughgarden, Tim, 2014. The sample complexity of revenue maximization. In: Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing. ACM, pp. 243–252.
Daskalakis, Constantinos, Deckelbaum, Alan, Tzamos, Christos, 2017. Strong duality for a multiple-good monopolist. Econometrica 85 (3), 735–767.
den Boer, Arnoud V., 2015. Dynamic pricing and learning: historical origins, current research, and new directions. Surv. Oper. Res. Manag. Sci. 20 (1), 1–18.
Fiat, Amos, Goldner, Kira, Karlin, Anna R., Koutsoupias, Elias, 2016. The fedex problem. In: Proceedings of the 2016 ACM Conference on Economics and Computation. ACM, pp. 21–22.
Grötschel, Martin, Lovász, Lászlo, Schrijver, Alexander, 1988. Geometric Algorithms and Combinatorial Optimization. Algorithms and Combinatorics, vol. 2. Springer.
Hart, Sergiu, Nisan, Noam, 2013. The menu-size complexity of auctions. In: Proceedings of the Fourteenth ACM Conference on Electronic Commerce. EC '13, New York, NY, USA. ACM. ISBN 978-1-4503-1962-1, pp. 565–566. https://doi.org/10.1145/2492002.2482544.
Hart, Sergiu, Nisan, Noam, 2017. Approximate revenue maximization with multiple items. J. Econ. Theory 172, 313–347.
Hart, Sergiu, Reny, Philip J., 2015. Maximal revenue with multiple goods: nonmonotonicity and other observations. Theor. Econ. 10 (3), 893–922.

Khachiyan, Leonid G., 1980. Polynomial algorithms in linear programming. USSR Comput. Math. Math. Phys. 20 (1), 53–72.
Kleinberg, Robert, Leighton, Tom, 2003. The value of knowing a demand curve: bounds on regret for online posted-price auctions. In: Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on. IEEE, pp. 594–605.
Lee, Yin Tat, Sidford, Aaron, Wong, Sam Chiu-wai, 2015. A faster cutting plane method and its implications for combinatorial and convex optimization. In: FOCS. IEEE, pp. 1049–1065. http://arxiv.org/abs/1508.04874v1. Full version on arXiv.
Madarász, Kristóf, Prat, Andrea, 2017. Sellers with misspecified models. Rev. Econ. Stud. 84 (2), 790–815.
Malakhov, Alexey, Vohra, Rakesh V., 2009. An optimal auction for capacity constrained bidders: a network perspective. Econ. Theory 39 (1), 113–128.
Manelli, Alejandro M., Vincent, Daniel R., 2006. Bundling as an optimal selling mechanism for a multiple-good monopolist. J. Econ. Theory 127 (1), 1–35.
Maskin, Eric, Riley, J., Hahn, F., 1989. Optimal multi-unit auctions. In: The Economics of Missing Markets, Information, and Games.
McAfee, R. Preston, McMillan, John, 1988. Multidimensional incentive compatibility and mechanism design. J. Econ. Theory 46 (2), 335–354.
Myerson, Roger B., 1981. Optimal auction design. Math. Oper. Res. 6 (1), 58–73.
Pavlov, Gregory, 2011. Optimal mechanism for selling two goods. B.E. J. Theor. Econ. 11 (1).
Riley, John, Zeckhauser, Richard, 1983. Optimal selling strategies: when to haggle, when to hold firm. Q. J. Econ. 98 (2), 267–289.
Sivan, Balasubramanian, Syrgkanis, Vasilis, 2013. Vickrey auctions for irregular distributions. In: International Conference on Web and Internet Economics. Springer, pp. 422–435.
Talluri, Kalyan T., Van Ryzin, Garrett J., 2006. The Theory and Practice of Revenue Management, Vol. 68. Springer Science & Business Media.
Thanassoulis, John, 2004. Haggling over substitutes. J. Econ. Theory 117 (2), 217–245.
Vaidya, Pravin M., 1989. A new algorithm for minimizing convex functions over convex sets. In: FOCS. IEEE, pp. 338–343.