Abstract

A seller can produce multiple units of a single good. The buyer has constant marginal value for each unit it 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 a natural regularity condition on the distribution, the optimal selling mechanism is deterministic. It is a price schedule that specifies the payment based on the number of units purchased. Further, under the same condition, 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 the problem in a mechanism design setting that allows for multiple dimensions of heterogeneity, and provide a simple condition 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.

*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.
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. The type of the buyer, consisting of its value per unit and its demand, is drawn at random from a distribution, and is known privately to the buyer. The buyer’s utility is its value for the units received minus payment to the seller. What selling mechanism maximizes expected revenue?

A mechanism is a menu of lotteries. Each lottery in the menu is a pair: a price and a distribution over the number of units. That is, once the buyer purchases the lottery by paying its price, it receives a random number of units from the distribution. The buyer purchases a lottery from the menu to maximize its 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 \(\frac{7}{3}\). 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 [33] shows that the optimal mechanism is deterministic if all types have the same demand, randomized mechanisms become optimal even for slight generalizations (Thanassoulis [38], Pavlov [34], Hart

---

1 A mathematically equivalent setting is that there is a population of buyers, each of whom has a linear valuation up to a demand.

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, and this is without loss of generality.
and Reny [25]). Optimal mechanisms are often structurally and computationally complex and may not even have a finite description (Daskalakis et al. [19], Hart and Nisan [24]). 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 first contribution is to identify a natural condition that guarantees that the optimal mechanism is deterministic. The 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. We also give a detailed description of the optimal prices when there are only two distinct demands in the distribution.

Our second contribution is to show that assuming DMR, the revenue as a function of the price vector is concave. This implies that optimal prices can be found efficiently using the ellipsoid or other cutting plane methods (Khachiyan [27], Vaidya [39], Lee et al. [29]). Note that 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 $p = (p_1, p_2, p_3)$, where $p_i$ is the price of $i$ units. Consider price vectors $p^1 = (2, 2, 3)$, $p^2 = (3, 3, 3)$, and their convex combination $p^3 = \frac{1}{2}p^1 + \frac{1}{2}p^2 = (2.5, 2.5, 3)$. The revenue of $p^1$ is $\frac{7}{3}$, the revenue of $p^2$ is 2, and the revenue of $p^3$ is $\frac{5.5}{3}$. This violates concavity since the convex combination of revenues of $p^1$ and $p^2$ with weights $\frac{1}{2}$ each is $\frac{6.5}{3}$, strictly greater than the revenue of $p^3$. Hence revenue as a function of the price vector is not concave.

The DMR Condition. Che and Gale [18] 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. [20] 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 [33].
DMR is precisely when no ironing is needed in the value space. The same assumption was also
made by Kleinberg and Leighton [28] in the context of dynamic pricing.

A simple class of DMR distributions is Uniform[$a, b$] for any non-negative reals $a$ and $b$.
More generally, any distribution with finite support and monotone non-decreasing probability
density is DMR. Another standard class of demand distributions that satisfies DMR is the
class of constant elasticity distributions. The DMR condition is different from the regularity
condition of Myerson [33]. The class of DMR distributions is well-behaved in the sense
that it is closed under convex combinations. On the other hand, it is known that regular
distributions are not closed under convex combinations (Sivan and Syrgkanis [36]). We show
that the DMR condition cannot be relaxed, by giving a distribution that satisfies monotone
hazard rate, a condition stronger than regularity, for which a deterministic pricing is not
optimal. See Section 2.1 for detailed discussions of properties of DMR distributions.

Our Approach. Let us first discuss our approach to showing optimality of deterministic
mechanisms. 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, without reducing revenue. In
order to do so, we first argue that without loss of generality, any type $(v, d)$ is assigned a
lottery over $d$ units or no allocation (that is, there is no chance of receiving $d' \neq d$ units).
Then we show that the randomized allocation to types with the highest values given each
demand can be converted to a deterministic allocation, without reducing revenue. Our first
step holds generally and does not require the DMR condition. 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. Removal of allocations can
only decrease (or keep fixed) the utility function of the mechanism. However, since the
highest type of each demand was assigned a deterministic allocation, the utility of such a

---

4 Ironing is a technique introduced by Myerson [33] where the virtual value function is transformed so that
it becomes monotone. This corresponds to transforming the corresponding revenue function into a concave
function.

5 As the name suggests, the elasticity of demand for such a distribution is constant over the support. Such
distributions are commonly used in Industrial Organization since they can be easily estimated by measuring
elasticity anywhere on the support [40, 5]. See Example 3 for the definition.

6 The monotone hazard rate condition requires the function $\frac{1-F(v)}{f(v)}$ to be monotone non-increasing.
type remains unchanged. A technical lemma shows that under the DMR condition, lowering utility of all types while keeping the utility of highest types fixed increases revenue.

In order to show that the revenue function is concave, we first give a closed form formula for the revenue function region-wise. We divide the price space into different regions such that a region determines the order in which a type with a certain demand buys a number of items smaller than its demand. For instance, a region might determine that for all the types with demand 10, as their value decreases from $\infty$ down to 0, the number of items they buy goes from 10 to 7 to 3 to 0; the exact transition points of course depend on the prices. We then provide the closed-form formula for revenue in each region and show that it is concave in that region. This in general does not imply that the revenue function is concave everywhere. One might surmise that the revenue function is the minimum of each of these functions, which would show that it is concave everywhere, but that is unfortunately not true. In fact, there is a partial order over these functions such that some of them are always higher than the others. We show a somewhat surprising property, that at the boundaries of the regions where they intersect, not only do the different functions agree (which they should, for the revenue function to be even continuous), but also their gradients agree. Showing this involves arguing that the equalities that hold at an intersection imply a whole set of other equalities such that disparate terms in the two gradients cancel out.

**Dynamic Pricing Application.** As a corollary, we obtain that under the DMR assumption, there is an efficient dynamic pricing scheme, defined as follows. Consider a repeated setting where in each round $\tau \in \{1, 2, \ldots, T\}$, the seller posts a price vector $p^{\tau}$, a buyer is drawn from a fixed distribution, and buys her utility maximizing bundle. The seller does not know the distribution of buyer types, and has to only use the purchase information in previous rounds to set the price. The goal is to approach the optimal revenue as $T$ goes to infinity. Given that the distribution satisfies the DMR assumption, our result on the concavity of the revenue curve implies that this is a special case of the “convex bandits” problem (Agarwal et al. [1], Bubeck et al. [9]). The results of Bubeck et al. [9] imply that there exists a dynamic pricing scheme such that the average revenue per round converges to the optimal revenue at the rate of $\frac{n^{0.5}}{\sqrt{T}}$, where $n$ is the number of units. These bounds are quite strong, since the best known bounds for the dynamic pricing problem in general scale exponentially in $n$; the concavity of the revenue function is an assumption often made to escape this curse of dimensionality (Besbes and Zeevi [6], Talluri and Van Ryzin [37]). We show that this assumption can be weakened to an assumption about the concavity of only the 1 dimensional revenue functions for each $d$. The same assumption was made by Kleinberg and Leighton [28] to get a $1/\sqrt{T}$ regret for the case of a single item.
1.1. Related work

Malakhov and Vohra [31] consider a more general version of our problem, in an auction design setting with multiple buyers, but make two assumptions: (1) that the buyers cannot report a higher demand, and (2) that the distribution satisfies the following: the Myerson virtual value\(^7\) is monotone in both the value and the demand. This essentially results in the problem separating out into one dimensional problems. The non-triviality in the 2 dimensional problem comes because buyers can misreport their demands. The first assumption disallows reporting a higher demand. The second assumption makes reporting a lower demand never profitable, without having to do anything extra. When specialized to the case of a single buyer, it implies that a deterministic pricing is optimal, since the same is true for the 1 dimensional case.

Fiat et al. [20] solve a related single buyer problem, 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” \(d\). The seller offers a service, such as delivering a package, at various points of time, and the buyer’s valuation is \(v\) for any time that is earlier than its deadline \(d\). In their model, a higher \(d\) corresponds to an inferior product, as opposed to our model where higher \(d\) is superior. The other difference is that 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.

We do allow the buyers to report higher demands. Consider the case that there are just 2 different demands (or deadlines) in the distribution, with \(d_1 < d_2\), and the question, when is it optimal to offer each level of service at the monopoly reserve price (say, \(r_1\) and \(r_2\) resp.) for the corresponding marginal distributions over values. The answer for the FedEx problem is, when \(r_1 \geq r_2\), which just says that \(d_1\) should cost more than \(d_2\). In our case, the answer is that \(r_1 \leq r_2\) and \(r_1 \geq \frac{d_1}{d_2}r_2\). Clearly \(d_1\) units should cost less, but not too low either, since in that case some buyers with demand \(d_2\) will actually prefer \(d_1\) units. This points to the added difficulty in our problem: we need to worry about a buyer opting for a bundle that could be of any size, but in the FedEx problem a buyer would never consider later time slots. In addition, the new IC constraints we need to consider are of the form where \((v, d)\) reports \((\frac{d}{d+1}v, d+1)\). These are “diagonal” IC constraints, as compared to the “vertical” ones in the FedEx problem, where a buyer of type \((v, d)\) reports \((v, d - 1)\). These are harder to handle.

\(^7\)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 \(f_d\). \(\phi(v, d) = v - \frac{1 - F_d(v)}{f_d(v)}\).
and the techniques used in the FedEx problem, such as constructing an optimal dual, seem difficult to extend to this case.

Closed form characterizations of optimal multi-product mechanisms are rare. Daskalakis et al. [19] and Haghpanah and Hartline [22] identify conditions under which “selling only the grand bundle” is optimal. Giannakopoulos and Koutsoupias [21] identify a (deterministic) optimal auction for an additive buyer whose valuations are i.i.d. from $U[0,1]$, for up to 6 items. Manelli and Vincent [32] identify conditions under which the optimal bundling mechanism is deterministic.

The lack of characterizations of optimal mechanisms in general settings has been addressed by seeking computational results instead. See Hartline [26] for a thorough overview of this line of work. Alaei et al. [3] and Cai et al. [11, 12, 13, 14] showed that for finite (multi-dimensional) type spaces, the mechanism design problem can be reduced to a related algorithm design problem, thus essentially resolving the computational question for this case. Most of these assume a finite support and the computation time is polynomial in the size of the support. This is different from our model which assumes a continuous distribution.

Another approach to cope with the complexity of optimal mechanisms has been to show that simple mechanisms approximate optimal ones. In this line of work, two classes of valuations have been widely studied, unit demand valuations (Chawla et al. [17], Briest et al. [7], Alaei [2]), and additive valuations (Hart and Nisan [23], Li and Yao [30], Babaioff et al. [4], Yao [41]). A unified approach to both has been presented in Cai et al. [15], and these approaches have been extended to more general valuations in Rubinstein and Weinberg [35], Chawla and Miller [16], Cai and Zhao [10]. Most of these require independence of values for different items. Our model differs in that we allow arbitrary correlations between the marginal value $v$ and the demand $d$.

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 cost normalized to zero. The type $(v, d)$ of the buyer specifies her (per unit) value $v \in \mathcal{V} = [0, V]$ and her demand $d \in \mathcal{D} = \{1, 2, \cdots, k\}$.\footnote{The assumption that $\mathcal{V} = [0, V]$ is for simplicity and can be relaxed. Furthermore, the assumption that $\mathcal{D} = \{1, 2, \cdots, k\}$ can be relaxed to $\mathcal{D} = \{d_1, d_2, \cdots, d_k\}$ for arbitrary $d_i$, without affecting the results.} 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. The type is randomly drawn from a distribution over $\mathcal{V} \times \mathcal{D}$ with density $f$, and is privately known to the buyer.
By the revelation principle, we restrict attention to direct mechanisms. A (direct) mechanism \((A, P)\) is a pair of functions, an allocation rule \(A : V \times D \to \Delta(\mathbb{Z}_+)^9\) and a payment rule \(P : V \times D \to \Delta(\mathbb{R})\). Note that the mechanism may be randomized, that is, both the allocation \(A(v, d)\) and the payment \(P(v, d)\) are random variables.

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 \((v', d')\),

\[
\mathbb{E}[v \min\{A(v, d), d\} - P(v, d)] \geq \mathbb{E}[v \min\{A(v', d'), d\} - P(v', d')],
\]

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. That is, for all \((v, d)\),

\[
\mathbb{E}[v \min\{A(v, d), d\} - P(v, d)] \geq 0.
\]

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

\[
\mathbb{E}[P(v, d)]
\]

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

A stronger notion of individual rationality is ex-post individual rationality, which requires that the utility of a type is positive for any randomization of the mechanism. However, in the lemma below we show that any IR mechanism can be converted to an ex-post individually rational mechanism which guarantees that \(v \min\{A(v, d), d\} - P(v, d) \geq 0\), for all \((v, d)\) and all realizations of \(A(v, d)\) and \(P(v, d)\). The argument is standard and is deferred. (All the missing proofs in the rest of the paper are in Appendix A.)

**Lemma 1.** For every IC and IR mechanism, there exists an IC and ex-post IR mechanism with the same expected payment for each type.

By linearity of expectation, we may assume without loss of generality that the payment rule is deterministic. That is, given an IC and IR mechanism \((A, P)\), define \(p(v, d) = \mathbb{E}[P(v, d)]\). The mechanism \((A, p)\) is IC, IR, has a deterministic payment rule, and has the same revenue as \((A, P)\). Thus throughout the rest of the paper we denote a mechanism by \((A, p)\) where \(p\) is the deterministic payment rule.

---

9\(\Delta(\mathbb{Z}_+)\) is the set of probability distributions over \(\mathbb{Z}_+\), and \(\Delta(\mathbb{R})\) is the set of distributions over \(\mathbb{R}\).
Let $f_d$ and $F_d$ denote the probability density and the cumulative density functions of the marginal distribution of values conditioned on demand $d$. Our main theorems use the following assumption on the distribution of types.

**Definition 1.** The distribution over types has decreasing marginal revenue (DMR) if $v \cdot (1 - F_d(v))$ is concave in $v$ for all $d$.

To explain the 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

$$R_d(v) := v \cdot (1 - F_d(v)).$$ (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.\(^\text{10}\) We provide a more detailed comparison between DMR and regularity in Section 2.1.

Our first theorem shows that under DMR, 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 $i \in \mathbb{Z}_+$. We say that the optimal mechanism is deterministic if there exists an optimal mechanism that is deterministic.

**Theorem 1.** If the distribution of types has DMR, 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 $i \in [k]$, 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 $p$ denote the vector of unit prices

\(^{10}\)The more commonly used, but equivalent, definition is that $v - \frac{1 - F_d(v)}{f_d(v)}$ is monotone non-decreasing.
(p_1, \ldots, p_k). We assume without loss of generality that p_1 \leq p_2 \leq \cdots \leq p_k. We denote by \text{Rev}(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 p. Since maximizing a concave function can be done in polynomial time, the optimal mechanism can be found efficiently.

**Theorem 2.** If the distribution of types has DMR, then \text{Rev}(p) is concave, and as a result the optimal mechanism can be computed in polynomial time in k.

**Dynamic pricing.** Consider the following dynamic variant of the problem. In each round \tau \in 1, 2, \ldots, T, for some T \in \mathbb{Z}_+, the following takes place.

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

Assume, for the sake of notational convenience, that d_0 = 0 and p_0^\tau = 0 for all \tau, so x^\tau = 0 when the buyer doesn’t buy anything. The goal of the seller is to maximize her average (or equivalently, total) revenue

\[ \frac{1}{T} \sum_{\tau=1}^{T} p^\tau_{x^\tau}. \]

We evaluate the performance of a dynamic pricing scheme by its regret, which is the difference between the optimal expected revenue and the average expected revenue of the pricing scheme.

This problem is almost a special case of the convex bandits problem studied in Bubeck et al. [9]: 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 p^\tau from some convex domain K.
2. An adversary picks an arbitrary convex loss function \ell^\tau on domain K and range [0, 1].
3. The algorithm observes only \ell^\tau(p^\tau).

---

\[11\text{If } p_i > p_{i+1}, \text{ then no type would choose to receive } i \text{ units. Consider an alternative mechanism with prices } p' \text{ such that } p'_j = p_j \text{ for all } j \neq i, \text{ and } p'_i = p_{i+1}. \text{ The new mechanism satisfies the property that } p'_i \leq p'_{i+1}, \text{ and the two mechanisms have the same revenue.} \]
The goal of the algorithm is to minimize regret, defined as

\[ \sum_{\tau=1}^{T} \ell^\tau(p^\tau) - \min_{p \in \mathcal{K}} \ell^\tau(p). \]

Consider a reduction from the dynamic pricing problem to the convex bandits problem where we define the loss function at time \( \tau \) to be the negative revenue obtained from the buyer of type \((v^\tau, d^\tau)\) (and normalized appropriately so that the range is \([0, 1]\)). 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 [8], 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. [9]. Recall that \( \bar{V} \) is the highest possible value.

**Corollary 1.** There is a dynamic pricing scheme where the regret is

\[ \tilde{O}(n^{0.5})d_k \bar{V} \sqrt{T}. \]

2.1. The DMR Condition

A simple class of DMR distributions is Uniform\([a, b]\) for any non-negative reals \(a\) and \(b\). More generally, any distribution with finite 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 different from the regularity condition of Myerson [33], which requires that the function \(\phi(v) = v - \frac{1 - F(v)}{f(v)}\) is monotone non-decreasing in \(v\). The example below shows that DMR and regularity are incomparable conditions.

**Example 3** (DMR vs. regularity). Consider the class of constant elasticity distributions with cumulative density \(F(v) = 1 - (v/a)^{1/\epsilon}\) for any \(a \geq 0\) 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 is DMR, but not regular, for \(\epsilon < -1\). On the other hand, the exponential distribution is regular but not 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 \( \alpha \), 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 \( \alpha \), 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 combinations of the revenue functions of \( F_1 \) and \( F_2 \), and is concave if \( F_1 \) and \( F_2 \) are DMR.

We show that the DMR condition cannot be relaxed, by giving a distribution with monotone hazard rate (MHR), a condition stronger than regularity, for which a deterministic pricing is not optimal. The monotone hazard rate condition requires the function \( \frac{1 - F(v)}{f(v)} \) to be monotone non-increasing.

**Example 4 (MHR distributions where deterministic pricing is not optimal).** The marginal distributions of Example 1 for \( d = 1, 2 \) and 3 are point masses at 6, 1 and 1 respectively. Replace them with normal distributions \( N(1 - \epsilon, \sigma) \), \( N(1 - \epsilon, \sigma) \) and \( N(6 - \epsilon, \sigma) \), truncated at 0 and \( V \), for some \( V > 6 \), and some \( \epsilon > 0 \). Truncated normal distributions satisfy the monotone hazard rate condition. For any \( \delta > 0 \), we can choose \( \sigma \) and \( \epsilon \) small enough, such that the revenue of the optimal deterministic and randomized mechanisms from Example 1 change by less than \( \delta \). Furthermore, running these mechanisms on the new distributions yield essentially the same revenue.

### 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 2.** 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 \( a, d, d' \in \mathbb{Z}_+ \),

\[
\min \left\{ d, d' \right\} \frac{\min \left\{ a, d' \right\}}{d'} = \min \left\{ \frac{\min \left\{ d, d' \right\}}{d} \times a, \frac{\min \left\{ d, d' \right\}}{d'} \times d' \right\} \\
\leq \min \left\{ a, \min \left\{ d, d' \right\} \right\} \\
\leq \min \left\{ a, d \right\},
\]

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

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

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

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 truth-telling is equal in the two mechanisms. As a result, the mechanism \((A, p)\) is IC and IR.

Note that Lemma 2 does not imply that every optimal mechanism must allocate only the demanded number of units. We use Lemma 2 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 
\[ u(v, d) = v \min \{d, d'\} w_{d'}(v') - p(v', d'). \]

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:

\[ p(v, d) = v d w_d(v) - d \int_0^v w_d(z) \, dz + p(0, d). \]  

(3)

The second set of local constraints are the “vertical” 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 \((v d/(d + 1), d + 1)\). To state our claim regarding simplification of IC constraints formally, let \(u(v, d \rightarrow v', d')\) 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 (3), and
2. \(u(v, d \rightarrow v, d) \geq u(v, d \rightarrow v, d - 1)\), and
3. \(u(v, d \rightarrow v, d) \geq u(v, d \rightarrow v d/(d+1), d + 1)\).

Furthermore, and IC mechanism is IR if \(u(0, d \rightarrow 0, d) \geq 0\) 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

\[ u(v, d \rightarrow v, d) \geq u(v, d \rightarrow v, d - 1) \]
\[ = v(d - 1) w_{d-1}(v) - p(v, d - 1) \]
\[ = u(v, d - 1 \rightarrow v, d - 1). \]
Applying this inequality inductively, we have that

\[ u(v, d \to v, d) \geq u(v, d' \to v, d') \] (4)

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

\[ u(v, d' \to v, d') \geq u(v, d' \to v', d') = vd'w_d(v') - p(v', d') = u(v, d \to v', d') \] (5)

Note that (4) and (5) together imply that \( u(v, d \to v, d) \geq u(v, d \to v', d') \) 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

\[ u(v, d \to v, d) \geq u(v, d \to v, d') \]

\[ = vd'w_{d+1}(v') - p(v', d') = u(v, d \to v', d') \] (6)

From horizontal incentive compatibility we have

\[ u(v, d \to v, d) \geq u(v, d' \to v, d') \]

\[ = u(v, d' \to v', d') \] (7)

Putting (6) and (7) together, we conclude that \( u(v, d \to v, d) \geq u(v, d \to v', d') \) 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 (3) that the utility of type \((v, d)\) from truthtelling is

\[
d \int_0^v w_d(z) \, dz - p(0, d) \geq -p(0, d).
\]

Note that the right hand side is the utility of type \((0, d)\) from truthtelling. We conclude that if the IR constraint is satisfied for a type \((0, d)\), then it is also satisfied for all types \((v, d)\).

We now argue that in the optimal mechanism we must have \(p(0, d) = 0\) for all \(d\). Incentive compatibility requires that types \((0, d)\) and \((0, d')\) pay the same amount for all \(d, d'\), since otherwise the type with higher payment would prefer to report being the other type and pay less (such a type gets no utility from allocation). A mechanism where \(p(0, \cdot) < 0\) cannot be optimal. To see this, construct another mechanism which adds \(|p(0, \cdot)|\) to the payment of all types. The new mechanism is IC since \(|p(0, \cdot)|\) is added to both sides of each IC constraint. The new mechanism is IR since the utility of type \((0, d)\) is zero for all \(d\), and individual rationality for all types \((0, d)\) implies IR by Proposition 1. The new mechanism also has higher revenue. Thus, we assume that \(p(0, d) = 0\) for all \(d\), and the payments identity simplifies to:

\[
p(v, d) = vd w_d(v) - d \int_0^v w_d(z) \, dz.
\] (8)

It is interesting to compare this global-to-local reduction with that used in the FedEx problem. Syntactically, for the FedEx problem just the first 2 constraints above are sufficient, but the semantics are different. 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 indirect utility function \(U\) of a mechanism \((w, p)\) specifies the utility of truthtelling for each type,

\[
U_d(v) := vd w_d(v) - p(v, d)
\] (9)

The following lemma is a reformulation of Proposition 1 in terms of the indirect utility function \(U\). Define \(U'_d(v) = \frac{d}{dv} U_d(v)\).
Lemma 3. A mechanism \((w, p)\) with indirect utility \(U\) defined in (9) is IC if and only if

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

Furthermore, an IC mechanism \((w, p)\) is IR if \(U_d(0) \geq 0\) for all \(d\).

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 4. If a mechanism \((w, p)\) is IC and \(p(0, d) = 0\) for all \(d\), then the expected revenue is

\[\mathbb{E}[p(v, d)] = \sum_d \left( \bar{V} f_d(\bar{V}) U_d(\bar{V}) + \int_0^{\bar{V}} U_d(v) R''_d(v) dv \right) q_d.\]

Using Lemma 3 and Lemma 4, we now restate the optimal mechanism design problem.

\[
\max_U \sum_d \left( \bar{V} f_d(\bar{V}) U_d(\bar{V}) + \int_0^{\bar{V}} U_d(v) R''_d(v) dv \right) q_d \tag{10}
\]

subject to:

- \(U_d(v)\) is convex \(\forall d \in [k]\)
- \(d \geq U'_d(v) \geq 0\) \(\forall d \in [k], \forall v\)
- \(U_d(0) = 0\) \(\forall d \in [k]\)
- \(U_d(v) \geq U_{d-1}(v)\) \(\forall d \in \{2, \ldots, k\}\)
- \(U_d(v) \geq U_{d+1}(v\frac{d}{d+1})\) \(\forall d \in [k-1]\)

The objective of the mathematical program, the expected revenue, is written in terms of the indirect utility function using Lemma 4. Let us discuss the constraints. The first two constraints represent horizontal IC constraints and feasibility of the mechanism and follow from Lemma 4. 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 equivalent to \(p(0, d) = 0\). The last two constraints represent the vertical constraints and follow from Lemma 3.
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.

**Lemma 5.** Consider two IC mechanisms with indirect utilities $U$ and $\hat{U}$, such that $U_d(v) \leq \hat{U}_d(v)$ for all types, and $U_d(\bar{V}) = \hat{U}_d(\bar{V}), U_d(0) = \hat{U}_d(0) = 0$ for all $d$. If the distribution of types has DMR, then the revenue of the mechanism with indirect utility $U$ is at least as high as the revenue of the mechanism with indirect utility $\hat{U}$.

**Proof.** The proof follows directly from the expression of revenue in Lemma 4. The DMR property, which requires concavity of $R_d$, is equivalent to $R''_d(v) \leq 0$. Together with the assumption that $U_d(v) \leq \hat{U}_d(v)$, we have $U_d(v)R''_d(v) \geq \hat{U}_d(v)R''_d(v)$. Thus revenue of the mechanism with indirect utility $U$ is

$$\sum_d \left( \bar{V} f_d(\bar{V}) U_d(\bar{V}) + \int_0^{\bar{V}} U_d(v) R''_d(v) dv \right) q_d$$

$$\geq \sum_d \left( \bar{V} f_d(\bar{V}) \hat{U}_d(\bar{V}) + \int_0^{\bar{V}} \hat{U}_d(v) R''_d(v) dv \right) q_d,$$

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

The main theorem follows immediately from the following two lemmas.

The first lemma states that we can improve the revenue of any mechanism by assigning a deterministic allocation of $d$ units to a type $(\bar{V},d)$. 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 \((V, d)\) 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 6.** For any IC and IR mechanism, there exists an IC and IR mechanism, with revenue at least as large, wherein a type \((V, d)\) 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 \(w_d(V) = 1\) and \(p_d(V) = Vd - \hat{U}_d(V)\). 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 \((\bar{w}, \bar{p})\) is no lower than the revenue of \((\hat{w}, \hat{p})\). Indeed, note that

\[
p_d(V) = Vd - \hat{U}_d(V) \\
= Vd - (Vd\hat{w}_d(V) - \hat{p}_d(V)) \\
= Vd(1 - \hat{w}_d(V)) + \hat{p}_d(V)).
\]  

(11)

Thus \(p_d(V) \geq \hat{p}_d(V)\) 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 truth-telling in the two mechanisms is the same, that is

\[
U_d(v) = \hat{U}_d(v).
\]

(12)

Indeed, for any type \((V, d)\) we have \(U_d(V) = Vd - (Vd - \hat{U}_d(V)) = \hat{U}_d(V)\). And for or any type \((v, d)\) where \(v < V\), 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

\[
u(v, d \rightarrow \bar{V}, d') = v \min(d, d') - p_{d'}(\bar{V}) \\
= v \min(d, d') - \bar{V}d'(1 - \hat{w}_{d'}(\bar{V})) - \hat{p}_{d'}(\bar{V}),
\]
where the second equality followed form (11). Since $1 - \hat{w}_{d'}(\bar{V}) \geq 0$, we conclude that

$$u(v, d \rightarrow \bar{V}, d') \leq v \min(d, d') - v \min(d, d')(1 - \hat{w}_{d'}(\bar{V})) - \hat{p}_{d'}(\bar{V})$$

$$= v \min(d, d')\hat{w}_{d'}(\bar{V}) - \hat{p}_{d'}(\bar{V}).$$

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 $\hat{u}(v, d \rightarrow v, d)$. Therefore, we conclude that

$$u(v, d \rightarrow \bar{V}, d') \leq \hat{u}(v, d \rightarrow v, d) = \hat{U}_d(v) = U_d(v),$$

where the last equation followed from (12). Thus the mechanism is incentive compatible, and the lemma follows.

The next lemma shows that for any mechanism in which all types $(\bar{V}, d)$ deterministically receive $d$ units, there exists a deterministic mechanism with revenue at least as large. 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 $(\bar{V}, d)$ stays the same since it still has access to its most desirable assignment (the deterministic assignment). Lemma 5 can then be used to argue that the revenue of a deterministic mechanism is weakly higher.

**Lemma 7.** Consider any mechanism where any type with highest value $(\bar{V}, d)$ deterministically receives $d$ units. If the distribution of types has DMR, then there exists a deterministic mechanism with revenue at least as large.

**Proof.** Fix any type with highest value $(\bar{V}, d)$ 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. The alternative menu contains $k$ choices of deterministic allocations of 1 to $k$ 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 $(\bar{V}, d)$ 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 5, 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 6, there exists a mechanism with revenue at least as large, where any type \((\bar{V}, d)\) deterministically receives \(d\) units. By Lemma 7, there exists a deterministic mechanism with revenue at least as large. Notice that an optimal deterministic mechanism, that is a vector of prices \(p\) that maximizes \(\text{Rev}(p)\), exists. The optimal deterministic mechanism is IC and IR, and has no less revenue than any other mechanism, and is thus optimal.

4. Concavity of the Revenue Function

In this section we prove Theorem 2. Recall that the demands in the support of the distribution are 1, 2, \ldots, \(k\), and that for all \(i \in [k]\), \(p_i\) denotes the price for the bundle of \(i\) units, and \(p = (p_1, \ldots, p_k)\) denotes the vector of all prices. Without loss of generality, we may assume that the domain of \(p\) is such that

\[0 \leq p_1 \leq p_2 \leq \cdots \leq p_k.\]

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

**Theorem 2.** If the distribution of types has DMR, then \(\text{Rev}(p)\) is concave, and as a result the optimal mechanism can be computed in polynomial time in \(k\).

**Characterizing optimal bundles.** Given a price vector \(p\), 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 \(j \neq l \in [k]\). The following quantity turns out to be the threshold at which the preference changes.

\[\forall j, l \in [k] : j > l, \quad D_{j, l} \triangleq \frac{p_j - p_l}{j - l}.\]

For convenience, we also define \(D_{j, 0} \triangleq p_j/j\) for all \(j \in [k]\). 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 8.** For all \(i \geq j > l \in [k]\), 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 \(v j - p_j > v l - p_l\). 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 $i \geq j \geq l \in [k]$. 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 Figure 1.

![Figure 1](attachment:image.png)

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

**Lemma 9.** 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 $D_{i,l} = \frac{1}{i-l} ((i-j)D_{i,j} + (j-l)D_{j,l})$. \qed

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, 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 $i^{th}$ coordinate $\sigma(i) \in \arg\max_{j<i}\{D_{i,j}\}$. Given such a $\sigma$, for each $i$, the sequence of optimal bundles for types with demand $i$ is given by the directed path $P_\sigma(i)$, defined as the (unique) longest path starting from $i$ in the directed graph on $[k]$ 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, $P_\sigma(5)$ is the directed path $5 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 0$. See Figure 2.

![Figure 2](attachment:image.png)

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

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

**Revenue function formula.** We are now ready to give a closed form formula for the revenue function within each $\Delta_\sigma$. Recall that $F_i$ denotes 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 $\text{Rev}_\sigma(\mathbf{p})$, corresponding to some $\sigma$, captures the revenue of a price vector $\mathbf{p}$ in $\Delta_\sigma$.

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

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

**Lemma 10.** $\text{Rev}(\mathbf{p}) = \text{Rev}_\sigma(\mathbf{p})$ 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 8, 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 8, this is implied by the following.

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

We first prove that $\forall j \in \mathcal{P}_\sigma(i), D_{j,\sigma(j)} \geq D_{\sigma(j),\sigma^2(j)}$. 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 9), and that $D_{j,\sigma^2(j)} \leq D_{j,\sigma(j)}$ (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 $D_{l,\sigma(j)} \geq D_{j,\sigma(j)}$. 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 9), and $D_{j,l} \leq D_{j,\sigma(j)}$. Now by a repeated application of the fact $D_{j,\sigma(j)} \geq D_{\sigma(j),\sigma^2(j)}$, we get the same conclusion for all $j$ and $l$ such that $i \geq l > \sigma(j)$. \qed
Concavity of $\text{Rev}_\sigma$. Given a closed form for the revenue of a price vector, we next show that each $\text{Rev}_\sigma$ is a concave function. We do this by showing that $\text{Rev}_\sigma$ 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 $\text{Rev}_\sigma$ is concave as well. Note that $\text{Rev}_\sigma(p)$ being concave, does not imply that $\text{Rev}(p)$ is concave.

Lemma 11. For all $\sigma$, $\text{Rev}_\sigma(p)$ is a concave function.

Proof. We can rewrite $\text{Rev}_\sigma$ as follows, using the definition of $D_{j,l}$.

$$\text{Rev}_\sigma = \sum_i q_i \left( p_i - \sum_{j \in P(\sigma)} F_i(D_{j,\sigma(j)}(j - \sigma(j))) \right)$$

$$= \sum_i q_i \left( p_i - \sum_{j \in P(\sigma)} F_i(D_{j,\sigma(j)}(j - \sigma(j))) \right).$$

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 $p$ for all $j$. Since composition of linear functions with concave functions is concave, it follows that $-F_i(D_{j,\sigma(j)}(j - \sigma(j)))$ is concave. Now $\text{Rev}_\sigma$ is a positive linear combination of concave functions, which makes it concave too.

Stitching the $\text{Rev}_\sigma$s together. Lemma 10 and Lemma 11 imply that Rev is piecewise concave, i.e., inside each $\Delta_\sigma$ 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 $\min_\sigma \text{Rev}_\sigma$, since the minimum of concave functions is concave. Unfortunately, this is not true. In fact, there is a partial order over $\sigma$ that determine when one $\text{Rev}_\sigma$ is always greater than the other. We show a different, and somewhat surprising, property of the $\text{Rev}_\sigma$ that also implies that Rev is concave. We show that at the boundaries between two regions not only do the corresponding $\text{Rev}_\sigma$s agree (which they should, for Rev to be even continuous), but also their gradients agree.

Lemma 12. For all $\sigma, \sigma'$, $p$ such that $p \in \Delta_\sigma \cap \Delta_{\sigma'}$, we have that

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

Proof. We first argue that it is sufficient to prove Lemma 12 for the case where $\sigma$ and $\sigma'$ 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 $\sigma$ and $\sigma'$, and a sequence $\sigma = \sigma_1, \sigma_2, \ldots, \sigma_n = \sigma'$ such that for any $i$, $\sigma_i$ and $\sigma_{i+1}$ differ in exactly one coordinate, where $\sigma_i$ agrees with $\sigma$ in that coordinate and $\sigma_{i+1}$ agrees with $\sigma'$. The fact
that $p \in \Delta_\sigma \cap \Delta_{\sigma'}$ implies that for all coordinates $j$ such that $\sigma(j) \neq \sigma'(j)$, both $\sigma(j)$ and $\sigma'(j) \in \arg\max_{j < j'} \{D_{j,j'}\}$. Similarly, $p \in \Delta_{\sigma_i} \cap \Delta_{\sigma_{i+1}}$ requires the same condition, but only for the coordinate that they differ in, and therefore $p \in \cap_{i=1}^n \Delta_{\sigma_i}$. Since we know Lemma 12 holds when the two $\sigma$s differ in at most one coordinate, it now follows that Rev and Rev at $p$ are the same for all $\sigma_i$ and hence for $\sigma$ and $\sigma'$ as well.

Now we prove Lemma 12 when $\sigma$ and $\sigma'$ differ at exactly one coordinate, $i^*$. We consider the portions of the paths $P_\sigma(i)$ and $P_{\sigma'}(i)$ that are disjoint, and refer to these disjoint portions as simply $P \subseteq P_\sigma(i)$ and $P' \subseteq P_{\sigma'}(i)$. 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 $D$s are all the same, which is stated as the following lemma.

**Lemma 13.** All $j, j' \in P \cup P'$ 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 $P$ and $P'$ 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 $P \cup P'$ 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 9, 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 $j \in P$ and $j' \in P'$ 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 9, $D_{j,j'}$ must be in between $D_{j,v}$ and $D_{v,j'}$, therefore $D_{j,j'} \geq D_{v,j'}$. Let $i' \in P'$ 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 9, $D_{i',j'}$ must be in between $D_{i',v}$ and $D_{v,j'}$, therefore $D_{i',j'} \geq D_{v,j'}$. 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. \[\Box\]

Continuing the proof of Lemma 12. To show that Rev$_\sigma$s agree on the boundary, consider the difference Rev$_\sigma(p) - \text{Rev}_{\sigma'}(p)$. For all $i \leq i^*$, or $i$ such that $i^* \notin P_\sigma(i)$, nothing changes, therefore all those terms cancel out. Moreover, even for $i$ such that $i^* \in P_\sigma(i)$, the only
terms that don’t cancel out are \( j \in P \cup P' \). Therefore, we get:

\[
\text{Rev}_\sigma(p) - \text{Rev}_{\sigma'}(p) = \sum_{i \geq i^*} q_i \left( \sum_{j \in P} p_{\sigma(j)} \left( F_i(D_{j,\sigma(j)}) - F_i(D_{\sigma(j),\sigma'(j)}) \right) \right)
\]

\[
- \sum_{j \in P'} p_{\sigma'(j)} \left( F_i(D_{j,\sigma'(j)}) - F_i(D_{\sigma'(j),\sigma''(j)}) \right),
\]

which is zero by Lemma 13.

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

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

\[
= q_i \left( \frac{p_b}{b-j} f_i(D_{b,j}) - \frac{p_j}{b-j} f_i(D_{b,j}) + F_i(D_{b,j}) - F_i(D_{j,c}) \right) - \frac{p_j}{j-c} f_i(D_{j,c}) + \frac{p_c}{j-c} f_i(D_{j,c})
\]

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

By Lemma 13, \( D_{b,j} = D_{j,c} \), and therefore these terms are zero. The cases when \( i = i^* \), or \( i^* = a, b, j \), or \( c, d = 0 \), or \( j \in P_{\sigma'}(i) \) 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.

\textbf{Proof of Theorem 2.} Consider any two prices \( p_1 \) and \( p_2 \), 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 10 and Lemma 11, we have that this line segment is itself divided into many intervals (corresponding to the different \( \Delta_{\sigma} \)), and within each interval, Rev is a concave function. Further, from Lemma 12, 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. \( \Box \)
References


Appendix A. Deferred Proofs

Appendix A.1. Calculations from Example 1

Lemma 14. 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 $p_1 = p_2$. To see this most clearly, notice that 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. Note that since $p_1 = p_2$, $t_3$ will always purchase 2 units and contribute $\frac{6}{3}$ 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 $p_2 \frac{2}{3} + \frac{p_3}{3} = \frac{7}{3}$. □
Appendix A.2. Calculations from Example 3

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.

\[
R'(v) = \left( \frac{v}{a} \right)^{1/\epsilon} + \frac{v}{a \epsilon} \left( \frac{v}{a} \right)^{1/\epsilon - 1}.
\]

\[
R''(v) = \left( \frac{v}{a} \right)^{1/\epsilon - 1} \frac{2}{a \epsilon} + \left( \frac{v}{a} \right)^{1/\epsilon - 2} \frac{v}{a^2 \epsilon} (1/\epsilon - 1)
\]

\[
= \left( \frac{v}{a} \right)^{1/\epsilon - 2} \left( \frac{2v}{a^2 \epsilon} + \frac{v}{a^2 \epsilon} (1/\epsilon - 1) \right)
\]

\[
= \left( \frac{v}{a} \right)^{1/\epsilon - 2} \frac{v}{a^2 \epsilon} (1 + 1/\epsilon) \leq 0.
\]

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 \( v \).

\[
\phi(v) = v - \frac{1 - F(v)}{f(v)}
\]

\[
= v - \frac{1}{a \epsilon} (v/a)^{1/\epsilon - 1}
\]

\[
= v - \frac{v/a}{-1/(a \epsilon)} = v(1 + \epsilon),
\]

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 \( \phi(v) = v - \frac{1 - F(v)}{f(v)} = v - \frac{v}{e - e^{-v}} = v - 1 \) is monotone non-decreasing in \( v \).

Appendix A.3. Proof of Lemma 1

**Lemma 1.** For every IC and IR mechanism, there exists an IC and ex-post IR mechanism with the same expected payment for each type.

**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.

\[\square\]

Appendix A.4. Proof of Lemma 3

**Lemma 3.** A mechanism \((w, p)\) with indirect utility \(U\) defined in (9) is IC if and only if

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

Furthermore, an IC mechanism \((w, p)\) is IR if \(U_d(0) \geq 0\) for all \(d\).

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

Property 2 of the lemma is equivalent to property 2 of Proposition 1. By definition, \(u(v, d \rightarrow v, d) = U_d(v)\), and \(u(v, d \rightarrow v, d - 1) = v(d - 1)w_{d-1}(v) - p(v, d - 1) = U_{d-1}(v)\). Thus, \(u(v, d \rightarrow v, d) \geq u(v, d \rightarrow v, d - 1)\) is equivalent to \(U_d(v) \geq U_{d-1}(v)\).
Property 3 of the lemma follows from properties 1 and 3 of Proposition 1. This is because
\[
    u(v, d \to v \frac{d}{d+1}, d+1) = vdw_{d+1}(v \frac{d}{d+1}) - p(v \frac{d}{d+1}, d+1)
    = vdw_{d+1}(v \frac{d}{d+1}) - (vdw_{d+1}(v \frac{d}{d+1}) - \int_0^{\frac{d}{d+1}} w_{d+1}(z)dz)
    = \int_0^{\frac{d}{d+1}} w_{d+1}(z)dz
\]

Notice that the payment identity (8) implies that
\[
    U_d(v) = vdw_d(v) - \int_0^v w_d(z)dz = \int_0^v w_d(z)dz.
\]

Thus the above expression is
\[
    = U_{d+1}(v \frac{d}{d+1}).
\]

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

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

Appendix A.5. Proof of Lemma 4

Lemma 4. If a mechanism \((w, p)\) is IC and \(p(0, d) = 0\) for all \(d\), then the expected revenue is
\[
    \mathbb{E}[p(v, d)] = \sum_d \left( \bar{V} f_d(\bar{V})U_d(\bar{V}) + \int_0^{\bar{V}} U_d(v)R_d'(v)dv \right)q_d.
\]

Proof. Note for future reference that using integration by parts, we can write
\[
    \int_0^{\bar{V}} v f_d(v)U_d'(v)dv = v f_d(v)U_d(v)|_0^{\bar{V}} - \int_0^{\bar{V}} \frac{d}{dv}(vf_d(v))U_d(v)dv
    = \bar{V} f_d(\bar{V})U_d(\bar{V}) - \int_0^{\bar{V}} \frac{d}{dv}(vf_d(v))U_d(v)dv. \quad (A.1)
\]

The expected revenue is
\[
    \sum_d \left( \int_0^{\bar{V}} p(v, d)f_d(v)dv \right)q_d = \sum_d \left( \int_0^{\bar{V}} (vdw_d(v) - U_d(v))f_d(v)dv \right)q_d.
\]
subsututing $U_d'(v) = dw_d(v)$ from Lemma 3,

$$= \sum_d \left( \int_0^{\bar{v}} (vU_d'(v)f_d(v) - U_d(v)f_d(v))dv \right)q_d,$$

using integration by parts (A.1),

$$= \sum_d \left( \bar{V}f_d(\bar{V})U_d(\bar{V}) + \int_0^{\bar{v}} U_d(v)(-\frac{d}{dv}(vf_d(v)) - f_d(v))dv \right)q_d,$$

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

$$= \sum_d \left( \bar{V}f_d(\bar{V})U_d(\bar{V}) + \int_0^{\bar{v}} U_d(v)R_d'(v)dv \right)q_d.$$

$\Box$

Appendix B. Detailed Characterization for $k = 2$

We first complete the case analysis that shows that the optimal mechanism is deterministic for two demands $d_1, d_2$.

We have the following cases for $v_1$ and $v_2$:

$v_1 \leq v_2$: Thus, for all $v \leq v_1 \leq v_2$, we have that $U_{d_1}(v) = \frac{d_2}{d_1}U_{d_2} \left( \frac{d_1}{d_2}v \right) = U_{d_1} \left( v \frac{d_1}{d_2} \right)$, which implies that $U_{d_1}(v) = 0$, and therefore $U_{d_2}(v) = 0$. For all $v \geq v_1$, we have that $U_{d_1}'(v) = 1$, i.e. $w_{d_1}(v) = 1$; this corresponds to a posted price of $d_1v_1$ for a bundle of $d_1$ units. For $v \in [v_1, v_2]$, we have $U_{d_2}(v) = \frac{d_1}{d_2}U_{d_1}(v) = \frac{d_1}{d_2}(v - v_1)$. This implies that the allocation function $w_{d_2}(v)$ is equal to $\frac{d_1}{d_2}$; for a price of $d_1v_1$, we offer a bundle of $d_2$ units with probability $\frac{d_1}{d_2}$. For $v \geq v_2$, we have a posted price of $d_2v_2 - (v_2 - v_1)d_1$ for a bundle of $d_2$ units. The same allocation rule can be induced by just two menu units (and no randomization): $d_1$ units cost $v_1d_1$ and $d_2$ units cost $v_2d_2 - (v_2 - v_1)d_1$.

$v_2 \leq v_1$ and $v_1d_1 \leq v_2d_2$: As before, for all $v \leq v_2$, $U_{d_2}(v) = U_{d_1}(v) = 0$, and the bundle of $d_2$ units has a posted price of $v_2d_2$. For $v \in [v_2, v_1]$, we have $U_{d_1}(v) = \frac{d_2}{d_1}U_{d_2} \left( \frac{d_1}{d_2}v \right) \leq \frac{d_2}{d_1}U_{d_2}(v_2) = 0$. For $v \geq v_1$, $U_{d_1}'(v) = 1$; this is a posted price of $d_1v_1$ for $d_1$ units.

$v_2 \leq v_1$ and $v_1d_1 > v_2d_2$: Once again, for all $v \leq v_2$, $U_{d_2}(v) = U_{d_1}(v) = 0$, and the bundle of $d_2$ units has a posted price of $d_2v_2$. For $v \in \left[ v_2, \frac{d_2}{d_1}v_2 \right]$, we have $U_{d_1}(v) = \frac{d_2}{d_1}U_{d_2} \left( \frac{d_1}{d_2}v \right) =$
0. For $v \in \left[ \frac{d_2}{d_1} v_2, v_1 \right]$, $U'_e(v) = 1$; offer a bundle of $d_1$ units for a price of $d_1 \frac{d_2}{d_1} v_2 = d_2 v_2$. This corresponds to selling only the $d_2$ bundle for a price of $v_2 d_2$.

We now characterize the optimal thresholds $v_1$ and $v_2$. Let $v_1$ and $v_2$ be the values after which $(.,d_1)$ and $(.,d_2)$ type agents are allocated $d_1$ and $d_2$ units respectively. Then, the optimal mechanism posts a price for $d_1$ units and a price for $d_2$ units that is either: (1) $v_1 d_1$ and $v_2 d_2 - (v_2 - v_1)d_1$, (2) $v_1 d_1$ and $v_2 d_2$, or (3) $v_2 d_2$ for both. This is equivalent to the maximum of:

- $\max v_1 d_1 (2 - F_1(v_1) - F_2(v_1)) + v_2 (d_2 - d_1) (1 - F_2(v_2))$
  subject to $v_2 \geq v_1$.

- $\max v_1 d_1 (1 - F_1(v_1)) + v_2 d_2 (1 - F_2(v_2))$
  subject to $v_1 \geq v_2$ and $v_1 \leq \frac{d_2}{d_1} v_2$.

- $\max v_2 d_2 \left(2 - F_2(v_2) - F_1(v_2 \frac{d_2}{d_1})\right)$.

Let $v_1^*$ and $v_2^*$ be the optimal choices for $v_1$ and $v_2$. Also, let $\hat{v}_1$ and $\hat{v}_2$ be the monopoly pricing solutions, i.e. $\hat{v}_i = \arg\max v d_i (1 - F_i(v))$.

Then, we have the following options for $v_1^*$ and $v_2^*$:

1. $v_1^* = \hat{v}_1$ and $v_2^* = \hat{v}_2$ (unconstrained version of the second bullet)

2. $v_1^* = \arg\max v d_1 (2 - F_1(v) - F_2(v))$ and $v_2^* = \hat{v}_2$ (unconstrained version of the first bullet)

3. $v_1^* = v_2^* = \arg\max v (d_1 (1 - F_1(v)) + d_2 (1 - F_2(v)))$

4. $\frac{d_1}{d_2} v_1^* = v_2^* = \arg\max v d_2 \left(2 - F_2(v) - F_1(v \frac{d_2}{d_1})\right)$

This corresponds to the following: compute $\hat{v}_1$ and $\hat{v}_2$. If $\frac{d_2}{d_1} \hat{v}_2 \geq \hat{v}_1 \geq \hat{v}_2$ we’re done. Otherwise, compute $\arg\max v d_1 (2 - F_1(v) - F_2(v))$. If it is at most $\hat{v}_2$, then pick the best option out of 2,3 and 4. If not, pick the best out of 3 and 4.

- $\max v_1 d_1 (2 - F_1(v_1) - F_2(v_1)) + v_2 (d_2 - d_1) (1 - F_2(v_2))$
  subject to $v_2 \geq v_1$.

- $\max v_1 d_1 (1 - F_1(v_1)) + v_2 d_2 (1 - F_2(v_2))$
  subject to $v_1 \geq v_2$ and $v_1 \leq \frac{d_2}{d_1} v_2$.

- $\max v_2 d_2 \left(2 - F_2(v_2) - F_1(v_2 \frac{d_2}{d_1})\right)$

35
Let \( v_1^* \) and \( v_2^* \) be the optimal choices for \( v_1 \) and \( v_2 \). Also, let \( \hat{v}_1 \) and \( \hat{v}_2 \) be the monopoly pricing solutions, i.e. \( \hat{v}_i = \arg \max v d_i (1 - F_i(v)) \). The following procedure gives the optimal \( v_1^* \) and \( v_2^* \): compute \( \hat{v}_1 \) and \( \hat{v}_2 \), and check whether they satisfy the IC constraints. If they do, then we are done. If they do not, it must be that either \( \hat{v}_1 < \hat{v}_2 \), or \( \hat{v}_1 > \frac{d_2}{d_1} \hat{v}_2 \).

In the former case, compute the best per unit price \( q \), i.e. a price \( q \) such that \( d_1 \) units cost \( q d_1 \) and \( d_2 \) units cost \( q d_2 \). This corresponds to the solution of the first bullet.

In the latter case, compute the best bundle price, i.e. the best price \( p \) that is going to be the same for \( d_1 \) and \( d_2 \). This corresponds to the solution of the third bullet. The best of \( p \) and \( q \) the two is optimal, and given that, \( v_1^* \) and \( v_2^* \) can be easily calculated.

Then, we have the following options for \( v_1^* \) and \( v_2^* \):

1. \( v_1^* = \hat{v}_1 \) and \( v_2^* = \hat{v}_2 \) (unconstrained version of the second bullet)
2. \( v_1^* = \arg \max v d_1 (2 - F_1(v) - F_2(v)) \) and \( v_2^* = \hat{v}_2 \) (unconstrained version of the first bullet)
3. \( v_1^* = v_2^* = \arg \max v (d_1 (1 - F_1(v)) + d_2 (1 - F_2(v))) \)
4. \( \frac{d_1}{d_2} v_1^* = v_2^* = \arg \max v d_2 \left( 2 - F_2(v) - F_1(v) \frac{d_2}{d_1} \right) \)

This corresponds to the following: compute \( \hat{v}_1 \) and \( \hat{v}_2 \). If \( \frac{d_2}{d_1} \hat{v}_2 \geq \hat{v}_1 \geq \hat{v}_2 \) we’re done. Otherwise, compute \( \arg \max v d_1 (2 - F_1(v) - F_2(v)) \). If it is at most \( \hat{v}_2 \), then pick the best option out of 2,3 and 4. If not, pick the best out of 3 and 4.