

Yale University

## EliScholar – A Digital Platform for Scholarly Publishing at Yale

---

Cowles Foundation Discussion Papers

Cowles Foundation

---

7-12-2021

### The Optimality of Upgrade Pricing

Dirk Bergemann  
*Yale University*

Alessandro Bonatti

Andreas Haupt

Alex Smolin

Follow this and additional works at: <https://elischolar.library.yale.edu/cowles-discussion-paper-series>



Part of the [Economics Commons](#)

---

#### Recommended Citation

Bergemann, Dirk; Bonatti, Alessandro; Haupt, Andreas; and Smolin, Alex, "The Optimality of Upgrade Pricing" (2021). *Cowles Foundation Discussion Papers*. 2627.  
<https://elischolar.library.yale.edu/cowles-discussion-paper-series/2627>

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact [elischolar@yale.edu](mailto:elischolar@yale.edu).

# THE OPTIMALITY OF UPGRADE PRICING

By

Dirk Bergemann, Alessandro Bonatti, Andreas Haupt, and Alex Smolin

July 2021

COWLES FOUNDATION DISCUSSION PAPER NO. 2290



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS  
YALE UNIVERSITY  
Box 208281  
New Haven, Connecticut 06520-8281

<http://cowles.yale.edu/>

# The Optimality of Upgrade Pricing

Dirk Bergemann\*   Alessandro Bonatti†   Andreas Haupt‡   Alex Smolin§

July 12, 2021

## Abstract

We consider a multiproduct monopoly pricing model. We provide sufficient conditions under which the optimal mechanism can be implemented via upgrade pricing—a menu of product bundles that are nested in the strong set order. Our approach exploits duality methods to identify conditions on the distribution of consumer types under which (a) each product is purchased by the same set of buyers as under separate monopoly pricing (though the transfers can be different), and (b) these sets are nested.

We exhibit two distinct sets of sufficient conditions. The first set of conditions is given by a weak version of monotonicity of types and virtual values, while maintaining a *regularity* assumption, i.e., that the product-by-product revenue curves are single-peaked. The second set of conditions establishes the optimality of upgrade pricing for type spaces with *monotone marginal rates of substitution (MRS)*—the relative preference ratios for any two products are monotone across types. The monotone MRS condition allows us to relax the earlier regularity assumption.

Under both sets of conditions, we fully characterize the product bundles and prices that form the optimal upgrade pricing menu. Finally, we show that, if the consumer’s types are monotone, the seller can equivalently post a vector of single-item prices: upgrade pricing and separate pricing are equivalent.

KEYWORDS: Revenue Maximization; mechanism design; strong duality; upgrade pricing.

JEL CLASSIFICATION: D82, D42.

---

\*Department of Economics, Yale University, New Haven, CT 06511, [dirk.bergemann@yale.edu](mailto:dirk.bergemann@yale.edu).

†MIT Sloan School of Management, Cambridge, MA 02142, [bonatti@mit.edu](mailto:bonatti@mit.edu).

‡Institute for Data, Systems, and Society, MIT, Cambridge, MA 02142, [haupt@mit.edu](mailto:haupt@mit.edu).

§Toulouse School of Economics, 31000 Toulouse, France, [alexey.v.smolin@gmail.com](mailto:alexey.v.smolin@gmail.com).

# 1 Introduction

## 1.1 Motivation and Results

Pricing multiple goods with market power is a canonical problem in the theory of mechanism design. It is also a challenge of growing importance and complexity for online retailers and service providers, such as Amazon and Netflix. Both in theory and in practice, designing the optimal *mixed bundling* mechanism, (i.e., pricing every subset of products) becomes exceedingly complex in the presence of a large number of goods.

A natural question is then whether simpler pricing schemes are optimal under suitable demand conditions. A simple, commonly used mechanism consists of *upgrade pricing*, whereby the available options are ranked by set inclusion, i.e., some goods are only available as add-ons Ellison (2005). For example, many online streaming services use a tiered subscription model, whereby users can pay to upgrade to a “premium package”—a subscription with a larger selection of the provider’s content relative to the “basic package” Philips (2017).

In this paper, we obtain sufficient conditions under which upgrade pricing maximizes the seller’s revenue. Our approach consists of first identifying conditions under which the consumer’s types can be ordered in terms of their absolute or relative willingness to pay for the seller’s goods, and then ranking the goods themselves by the profitability of selling them to larger sets of consumer types. Our sufficient conditions not only establish the optimality of *some* upgrade pricing menu: they also show that the optimal bundles are deterministic, and they reveal the order in which they are ranked in the menu. That is, we identify all the nested bundles that appear in the seller’s menu, and the profit-maximizing price for each one.

Our results consist of two distinct sets of conditions. The first set of conditions (Theorem 1) illustrates the essence upgrade pricing optimality in what we label as “regular” settings. While these conditions are reminiscent of regularity in one dimension, they are in fact weaker than the monotonicity of the buyer’s multidimensional types and of the (item by item) Myersonian virtual values. What we require is for the consumer’s types to be ranked in such a way that the virtual values for each item are negative over an initial and positive over a final segment. Furthermore, we require any consumer with a positive virtual value for an item to also have a larger value for that item, relative to any type with a negative virtual value. At the optimal prices, the lowest type buying each good is indifferent between buying it and not buying it. Finally, the sets of types buying each item are nested under the *weak* monotonicity property, which implies the optimal allocation can be implemented via upgrade pricing.

The second set of conditions (Theorem 2) describes our best attempt at extending our

approach to non-regular distribution of types. In order to further weaken the regularity requirement, we restrict attention to type spaces for which the relative preference ratios for any two goods are monotone across types. An example of ordered relative preferences is if higher types have a stronger preference for good 2 over good 1. We refer to such a condition as “monotone marginal rates of substitution” (monotone MRS).

The intuition for our two results can be grasped by considering the demand functions for each good separately. Under monotonicity and monotone MRS, the optimal monopoly prices for each of the goods are ranked. In the special case where the Myersonian virtual values for our ordered types

$$\phi_i^k = \theta_i^k - \frac{1 - F_i}{f_i} (\theta_{i+1}^k - \theta_i^k)$$

are also monotone for each item  $k$ , the first set of conditions applies.

When virtual values are not monotone, however, they can cross zero more than once. In that case, the result still holds, but the proof requires the right ironing procedure. Our ironing procedure relaxes the standard approach of Myerson (1981) and the literature up to Haghanah and Hartline (2020). Specifically, we do not iron with the goal of monotone virtual values, which corresponds to a concave revenue curve. Rather we iron towards single-crossing virtual values which leads to a *quasi*concave revenue curve. We then use the structure implied by monotone MRS to derive a dual certificate of optimality.

Under either set of conditions, each good is purchased by the same set of buyers that would buy it if that were the seller’s *only* product. We further show (Theorem 3) that, if the consumer’s types are (not weakly) monotone, the seller can equivalently post the vector of single-item monopoly prices—i.e., bundling is redundant. For example, in the case of two goods sold separately, monotone type spaces mean that no consumer type will buy good 2 without also buying good 1. More generally, the seller benefits from restricting the set of bundles the consumer can purchase through a proper menu of options with the upgrade property. However, examples also show that implementability through separate pricing is neither necessary nor sufficient for the optimality of upgrade pricing.

## 1.2 Related Literature

First and foremost, our paper contributes to the economics literature on product bundling. The profitability of mixed bundling relative to separate pricing was first examined by Adams and Yellen (1976), and further generalized by McAfee et al. (1989). More recently, a number of contributions have studied the optimal selling mechanisms in the case of two or three goods, and derived conditions for the optimality of pure bundling (see, for example, Manelli

and Vincent (2006) and Pavlov (2011)). Daskalakis et al. (2017) use duality methods to characterize the solution of the multiproduct monopolist’s problem, and show how the optimal mechanism may involve a continuum of lotteries over items. Bikhchandani and Mishra (2020) derive conditions under which the optimal mechanism is deterministic when the buyer’s utility is not necessarily additive. Finally, Ghili (2021) establishes conditions for the optimality of pure bundling when buyers’ values are interdependent. Relative to all these papers, we focus on a specific class of simple mechanisms, which includes pure bundling as a special case.

Hart and Nisan (2017) and Babaioff et al. (2014) also study the properties of simpler schemes. The former derives a lower bound on the revenue obtained from separate item pricing. The latter obtains an upper bound on the revenue of the optimal mechanism, relative to the better of pure bundling and separate pricing.

Our formulation of the dual problem follows Cai et al. (2016), who present a general duality approach to Bayesian mechanism design. Cai et al. (2016) formulate virtual valuations in terms of dual variables, state the weak and the strong duality results, and use them to establish lower bounds for relative performance of simple mechanisms. An important contribution by Haghpanah and Hartline (2020) exploits the duality machinery to provide sufficient conditions for the exact optimality of a specific, simple mechanism—pure bundling—consisting of offering a maximal bundle at a posted price. Under their sufficient conditions, the dual variables can be recovered from a single-dimensional problem in which the seller is restricted to bundle all items together.

We follow the approach of Haghpanah and Hartline (2020) by leveraging the duality approach to provide sufficient conditions for the optimality of a particular class of mechanisms. Haghpanah and Hartline (2020) gave a characterization of the optimality of the grand bundle, we provide a characterization for upgrade pricing. As upgrade pricing allows multiple items to be present in the menu, we cannot assign the dual variables in a pre-specified way. Instead, we develop a novel ironing algorithm that generates these variables for any given problem. Under our sufficient conditions, the so-constructed virtual surplus is maximized by an element-wise monotone allocation that can be implemented by upgrade pricing; by complementary slackness, this certifies the optimality of upgrade pricing. Because pure bundling is one instance of upgrade pricing, our conditions differ from those of Haghpanah and Hartline (2020).

Our ironing differs from existing ironing approaches using duality and tackles a more general problem. In comparison to Haghpanah and Hartline (2020), we prove optimality for mechanisms with menu size surpassing two. Fiat et al. (2016) studies a two-parameter model, and uses an ironing approach that leads from the revenue curves to their concave

closure. Devanur et al. (2020) generalizes Fiat et al. (2016) to more general orders on the second parameter. Our approach tackles optimality for an arbitrary finite number of items and varies the ironing procedure. On a technical level, our ironing procedure yields quasi-concave ironed revenue curves, whereas the ironed revenue curves in Haghpanah and Hartline (2020); Fiat et al. (2016); Devanur et al. (2020) are concave.

Our results also feed into a literature specifying optimal finite mechanisms for multi-dimensional types. (Daskalakis et al., 2017, Section 7) for example characterizes the optimal mechanisms for the two-good monopolist problem if the optimal mechanism has a particular structure. While Daskalakis et al. (2017) requires that the region of the type space that is not allocated any item is not adjacent to all regions getting specific constant allocations, upgrade pricing mechanisms consistently break this requirement.

### 1.3 Structure of the Paper

The model is introduced in section 2. The first set of sufficient condition is presented in section 3. In section 4, we present our results for monotone MRS type spaces. In section 5, we discuss the relationship between separate pricing and upgrade pricing. In section 6, we numerically explore the “robust” optimality of upgrade pricing with respect to the distribution of types. We conclude in section 7.

## 2 Model

We consider a standard multiple-good monopoly setting. There is a single seller of  $d \geq 1$  goods and a single buyer. The seller’s marginal costs of production are normalized to zero. The buyer’s utility function is additive across goods. We refer to the vector of marginal utilities  $\theta_i \in \mathbb{R}^d$  as the buyer’s type. Therefore, the utility of buyer type  $\theta_i$  from the consumption vector  $q \in [0, 1]^d$  is given by

$$U(\theta_i, q) = \sum_{k=1}^d \theta_i^k q^k.$$

We also write as a shorthand  $\langle \theta_i, q \rangle := \sum_{k=1}^d \theta_i^k q^k$ . As a convention, we denote types by subscripts and refer to items by superscripts. The buyer’s utility is quasi-linear in transfers and his outside option is also normalized to zero.

The buyer knows her type. From the seller’s perspective, the buyer’s type is distributed over a finite set  $\Theta \subseteq \mathbb{R}_+^d$ , with  $|\Theta| = n$ , according to distribution  $f \in \Delta(\Theta)$ . For any positive integer  $n$ , we adopt the convention that  $[n] := \{1, 2, \dots, n\}$ , and we index types by  $i \in I = [n]$ . We denote  $f_i := f(\theta_i)$  and denote the cumulative distribution sequence by

$$F_i = \sum_{j=1}^i f_j, i \in [n].$$

The seller aims to maximize revenue. By the revelation principle, we can focus on direct mechanisms  $(q, t) = (q_i, t_i)_{i \in \{0\} \cup [n]}$  interpreted as menus with  $n + 1$  items so that item  $i$  delivers consumption vector  $q_i$  at price  $t_i$  and item  $(q_0, t_0) := (0, 0)$  captures an outside option. In a direct mechanism, each type  $\theta_i$  prefers item  $i$  to all other items.

We call a menu *upgrade pricing* if  $\{q_0, q_1, \dots, q_n\}$  can be ordered in the component-wise partial order on  $\mathbb{R}^d$ . Our main goal is to provide conditions under which upgrade pricing maximizes the seller's revenue among all direct mechanisms.

We will make prominent use of (partial) Lagrangian duality-based certificate of optimality, as used by Cai et al. (2016), which we state and prove to fix notation. In what follows, we will associate with  $\lambda_{ji}$  the Lagrange multiplier of the incentive compatibility constraint of type  $\theta_j$  deviating to type  $\theta_i$ ,  $j \in [n], i \in \{0\} \cup [n]$ :

$$\langle q_j, \theta_j \rangle - t_j \geq \langle q_i, \theta_j \rangle - t_i.$$

We note that the incentive constraints corresponding to  $\lambda_{j0}$ ,  $j \in [n]$  are type  $j$ 's individual rationality constraints. As a main tool in our analysis, we define the multi-dimensional *virtual values* associated with Lagrange multipliers  $\lambda \in \mathbb{R}^n \times \mathbb{R}^{n+1}$  as

$$\phi_i^\lambda := \theta_i - \frac{1}{f_i} \sum_{j=1}^n \lambda_{ji} (\theta_j - \theta_i). \quad (1)$$

**Lemma 1.** *A mechanism  $(q_i, t_i)_{i \in \{0\} \cup [n]}$  maximizes revenue if and only if there are multipliers  $\lambda_{ji}$ ,  $j \in [n], i \in \{0\} \cup [n]$  such that*

1.  $\lambda_{ji} \geq 0$  (*Non-Negativity*)
2.  $(q_i)_{i \in [n]}$  optimizes  $\max_{(q_i)_{i \in [n]} \in [0, 1]^n} \sum_{i=1}^n f_i \langle q_i \cdot \phi_i^\lambda \rangle$  (*Virtual Welfare Maximization*)
3.  $f_i = \sum_{j=0}^n \lambda_{ij} - \sum_{j=1}^n \lambda_{ji}$  for all  $i \in [n]$  (*Feasibility of Flow*)
4.  $\lambda_{ji} (\langle q_j, \theta_j \rangle - t_j - \langle q_i, \theta_j \rangle - t_i) = 0$  for all  $j \in [n], i \in \{0\} \cup [n]$  (*Complementary Slackness*)
5. *There are transfers  $t$  such that  $(q, t)$  is implementable (Implementability)*

*Proof of Lemma 1.* The Karush-Kuhn-Tucker conditions, respectively Slater's condition for affine inequality constraints (see (Boyd and Vandenberghe, 2004, p. 227)), allow us to write revenue maximization subject to the incentive compatibility and individual rationality constraints as an unconstrained optimization problem for  $(q_i, t_i)_{i \in [n]}$  subject to complementary



slackness and non-negativity of dual variables. The Lagrangian reads:

$$\begin{aligned}
\mathcal{L} &= \sum_{i=1}^n f_i t_i + \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} (\langle q_j, \theta_j \rangle - t_j - \langle q_i, \theta_j \rangle - t_i) \\
&= \sum_{i=1}^n t_i \left( f_i - \sum_{j=0}^n \lambda_{ij} + \sum_{j=1}^n \lambda_{ji} \right) + \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_j, \theta_j \rangle - \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_i, \theta_j \rangle \\
&= \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_j, \theta_j \rangle - \sum_{j=1}^n \sum_{i=0}^n \lambda_{ji} \langle q_i, \theta_j \rangle \\
&= \sum_{j=1}^n \left( \left( \sum_{i=1}^n \lambda_{ij} - \sum_{i=0}^n \lambda_{ji} \right) \langle q_j, \theta_j \rangle - \sum_{i=0}^n \lambda_{ji} (\langle q_i, \theta_j \rangle - \langle q_j, \theta_j \rangle) \right) \\
&= \sum_{j=1}^n \left( f_j \langle q_j, \theta_j \rangle - \sum_{i=0}^n \lambda_{ji} (\langle q_i, \theta_j \rangle - \langle q_j, \theta_j \rangle) \right) \\
&= \sum_{j=1}^n f_j \langle q_j, \phi_j \rangle.
\end{aligned}$$

Clearly, it is necessary for an optimal mechanism to be implementable. To conclude the proof, we need to show that virtual welfare maximization and feasibility of flow is equivalent to maximizing the Lagrangian. Assume virtual welfare maximization and feasibility of flow. Then the above equalities show that the Lagrangian is maximized and certify the optimality of the mechanism  $(q_i, t_i)_{i \in [n]}$ . Assume on the other hand that the Lagrangian is maximized by  $(q_i, t_i)_{i \in [n]}$ . If feasibility of flow would not hold, then choosing  $t_i$  arbitrarily large or small would lead to a higher value for the Lagrangian, a contradiction. Given that this is zero, the Lagrangian equals virtual welfare, and virtual welfare maximization follows from optimality.  $\square$

We will call  $\lambda_{ji}$  the *flow* from type  $j$  to type  $i$ , as condition 3 resembles flow conservation constraints known from maximum flows and minimum cost flows in discrete mathematics Korte and Vygen (2011).

### 3 Optimal Mechanisms for Regular Distributions

Our first set of sufficient conditions for upgrade pricing optimality consists of a weak monotonicity condition and a regularity condition.

Let  $I \subseteq \mathbb{R}$  be an index set. We refer to a type distribution  $F$  as *weakly monotone* with

respect to cutoffs  $i^1, i^2, \dots, i^d \in [n]$  if for any  $i, j \in [n]$  and  $k \in \{1, 2, \dots, d\}$ ,

$$i \leq i^k \leq j \implies \theta_i^k \leq \theta_j^k.$$

Similarly, a type distribution  $F$  is *regular* with respect to cutoffs  $i^1, i^2, \dots, i^d \in [n]$  if for any  $i, j \in [n]$  and  $k \in \{1, 2, \dots, d\}$ ,

$$i \leq i^k \leq j \implies \phi_i^k \geq 0 \geq \phi_j^k,$$

where  $\phi_i$  denotes the *initial  $d$ -dimensional virtual values*

$$\phi_i := \theta_i - \frac{1 - F_{i-1}}{f_i}(\theta_{i+1} - \theta_i). \quad (2)$$

Initial  $d$ -dimensional virtual values can be seen as multi-dimensional versions of Myerson (1981).

Our regularity condition can be equivalently stated in terms of the *pseudo-revenues*

$$R_i^k := (1 - F_i)\theta_i^k. \quad (3)$$

We call (3) *pseudo-revenue* because, without an assumption that the values for all items  $k$  are monotone, the pseudo-revenue does not correspond to the revenue from sales of item  $k$  at a posted price of  $\theta_i^k$ . In particular, because we have

$$\phi_i^k = \frac{R_i^k - R_{i+1}^k}{f_i},$$

regularity with respect to cutoffs  $i^k$  is equivalent to requiring that  $R_i^k$  is single-peaked with peak  $i^k$ .

Finally, we say that a type distribution  $F$  is *compatibly* weakly monotone and regular if it is both weakly monotone and regular with respect to the same cutoffs. Figure 1 illustrates such a type distribution, and the associated pseudo-revenues.<sup>1</sup>

---

<sup>1</sup>In this example, the type space is  $\Theta = \{(9/128, 27/64), (1/4, 3/2), (1/2, 2), (1, 1)\}$  and the type distribution is  $f = (7/16, 3/16, 1/8, 1/4)$ . The optimal mechanism sells good 2 at a price of 1 and good 1 as an upgrade, also at a price of 1. All types except  $\theta_1$  buy good 2, and only type  $\theta_4$  buys good 1.

**Theorem 1.** *If the type distribution  $F$  is compatibly weakly monotone and regular, then upgrade pricing is optimal. In particular, the following mechanism is optimal:*

$$q_i^k := \begin{cases} 1 & R_i^k \geq R_{i+1}^k, \\ 0 & \text{else.} \end{cases}, \quad i \in [n], k \in [d]. \quad (4)$$

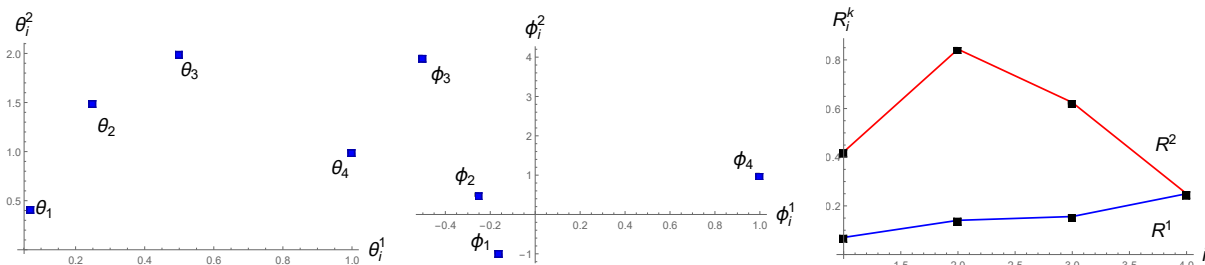


Figure 1: Left: Dimension-wise types. Center: Virtual values. Right: Pseudo-revenues.

*Proof of Theorem 1.* It is direct from the definition (1) that the dual variables

$$\hat{\lambda}_{ij} = \begin{cases} 1 - F_{i-1} & \text{if } j = i + 1 \\ 0 & \text{else,} \end{cases} \quad (5)$$

induce the initial virtual values,  $\phi_i = \phi_i^\lambda$ .

We check the properties of Lemma 1. As we allocate items if and only if the virtual value is non-negative, condition 2 is satisfied. For condition 3, observe that

$$\sum_{j=1}^n \hat{\lambda}_{ij} - \sum_{j=0}^n \hat{\lambda}_{ji} = 1 - F_{i-1} - (1 - F_i) = f_i.$$

The mechanism is implementable in dominant strategies, condition 5, by assumption of compatible weak monotonicity and regularity. Finally, we need to check that complementary slackness, condition 4, holds. Observe that it is sufficient to see that all local downwards incentive compatibility constraints bind. As  $\hat{\lambda}_{ij} > 0 \implies j = i - 1$ , to show complementary slackness, we only need to check that incentive constraints bind for any consecutive types. Such types either get the same allocation and payment in which case complementary slackness is trivially satisfied; or, they are assigned a transfer making them indifferent with the next lower type, for otherwise the transfer could be raised, strictly increasing revenue, which would be a contradiction to revenue maximization.  $\square$

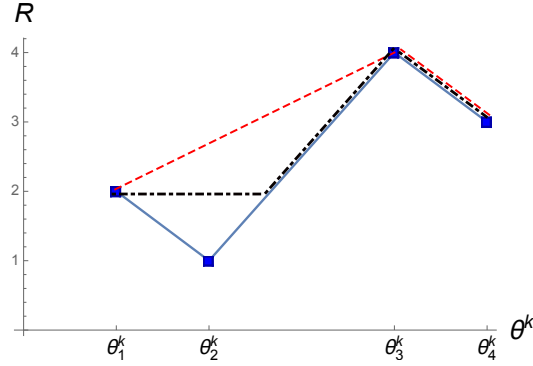


Figure 2: Blue: a base function. Red: a concave closure. Black: a quasi-concave closure.

Our assumptions of regularity and weak monotonicity relax the monotonicity of types and Myersonian virtual values by allowing for permutations above and below the monopoly price. These assumptions nonetheless require that the set of types that buy each object remains an upper selection, and conversely the set of types that do not buy remains a lower selection. The intuition for why this works is similar to the idea that the monopoly price does not depend on the valuations of inframarginal types, just as long as they remain, in fact, inframarginal.

Our next set of conditions imposes similar requirements, strengthened appropriately to allow for ironing of non-regular type distributions.

## 4 Optimal Mechanisms for Non-Regular Distributions

We now establish optimality of an upgrade pricing mechanism in settings without regularity. We say that a type space  $\Theta$  has *monotone marginal rates of substitution* if

$$i \leq j \text{ and } k \leq l \implies \frac{\theta_i^k}{\theta_i^l} \leq \frac{\theta_j^k}{\theta_j^l}.$$

for any  $i, j \in [n]$ ,  $l, k \in [d]$ .

Recall that pseudo-revenue is given by

$$R_i^k = (1 - F_i)\theta_i^k.$$

We call a scalar sequence  $(R_i)_{i \in [n]}$ , *quasi-concave* if there is a cutoff  $i' \in [n]$  such that  $i' \leq i \leq j$  or  $j \leq i \leq i'$  implies  $R_i \geq R_j$ . We call the point-wise smallest quasi-concave sequence that is point-wise larger than  $(R_i)_{i \in [n]}$  its *quasi-concave closure* and denote it by  $(\bar{R}_i)_{i \in [n]}$ . See Figure 2 for an example of a quasi-concave closure. We will make regular use

of the sequence  $(\overline{R_i^k})_{i \in [n]}$ , the quasi-concave closure of the values for item  $k$ ,  $(R_i^k)_{i \in [n]}$ .

We relax the regularity assumption on pseudo-revenues  $R_i^k$ . Instead of assuming regularity, i.e.  $R_i^k$  to be single-peaked with peak  $i^k$ , we assume three properties that are in combination weaker than regularity. We call a type distribution  $F$  *mostly regular* if for some  $i^k \in \arg \max_{i \in [n]} R_i^k$  and any  $i$  such that  $i^k < i \leq i^{k+1}$

1. If  $R_i^l \neq \overline{R_i^l}$ , then either  $R_{i-1}^l \neq \overline{R_{i-1}^l}$  or  $R_{i-1}^{l'} = \overline{R_{i-1}^{l'}}$  for  $l' \in \{k-1, k+1\}$  (no overlap)
2.  $R_{i^k}^l = \overline{R_{i^k}^l}$  for  $l \in \{k-1, k+1\}$  (no ironing on maxima)
3. If  $i^k \leq i < j \leq i^{k+1} \in [n]$  and  $\overline{R_r^k} \neq R_r^k$  for any  $i \leq r \leq j$ , then  $\theta_i^{k+1} \leq \theta_r^{k+1}$  (not too shuffled)

We can think of intervals of types  $i \in [n]$  such that  $R_i^l \neq \overline{R_i^l}$  as *ironing candidate intervals for item  $k$* . With this language, the first property states that candidate ironing intervals of non-regularity for two consecutive goods (in the marginal rates of substitution order) are either disjoint or one contains the other. The second property states that a maximum of the pseudo-revenue, which is the first type in the MRS order to buy item  $k$ , is not part of a candidate ironing interval of a consecutive item. The third property is a weaker monotonicity property. “Not too shuffled” requires that values for types in a candidate ironing interval are monotone for at least one adjacent item.

Finally, we call a distribution *compatibly* weakly monotone and mostly regular if it is weakly monotone and mostly regular with respect to the same cutoffs. Figure 3 shows an instance of such a distribution over a monotone MRS type space.<sup>2</sup>

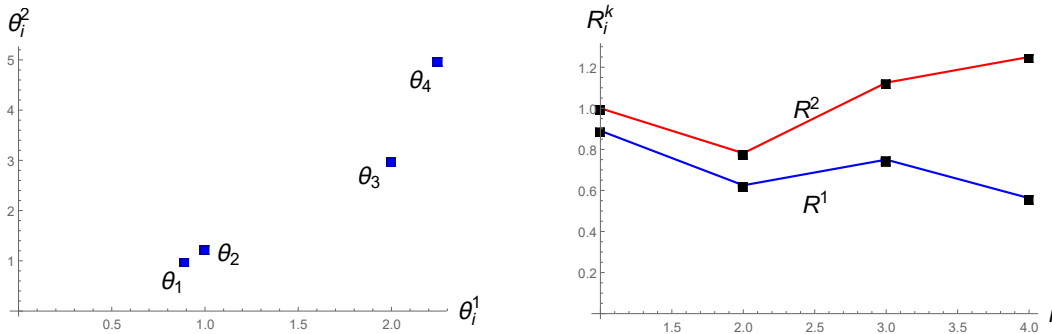


Figure 3: Left: Dimension-wise types. Right: Pseudo-revenues.

Conversely, Figure 4 shows an instance of a distribution over a monotone MRS type space that is *not* Mostly Regular. In particular, this example fails the first condition, because it involves overlapping ironing intervals.

<sup>2</sup>In Figure 3, the type space is given by  $\Theta = \{(57/64, 1), (1, 5/4), (2, 3), (9/4, 5)\}$  and the type distribution is given by  $f = (3/8, 1/4, 1/8, 1/4)$ . The optimal mechanism sells good 1 at a price of  $57/64$ , and good 2 as an upgrade at a price of 5. All types buy good 2, and only type  $\theta_4$  buys good 1.

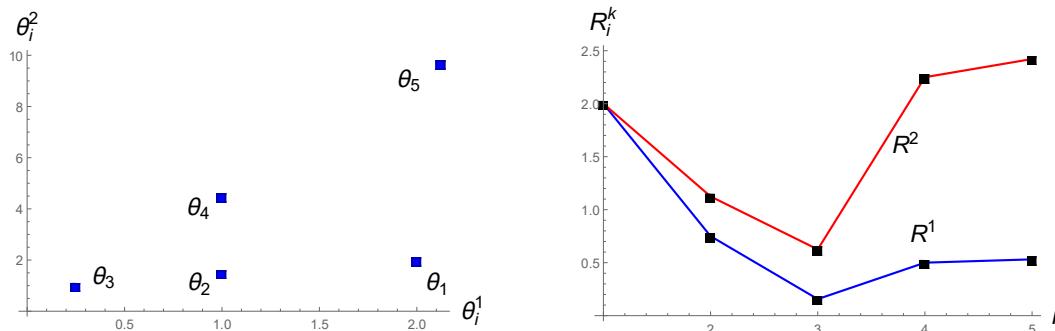


Figure 4: Left: Dimension-wise types. Right: Pseudo-revenues.

**Theorem 2.** *Let  $\Theta$  have monotone marginal rates of substitution. If the type distribution  $F$  is compatibly weakly monotone and mostly regular, then upgrade pricing is optimal. In particular, the following mechanism is optimal:*

$$q_i^k := \begin{cases} 1 & \overline{R}_i^k \geq \overline{R}_{i+1}^k, \\ 0 & \text{else.} \end{cases}, \quad i \in [n], k \in [d]. \quad (6)$$

Note that the allocation (6) is implementable through upgrade pricing because  $(\overline{R}_i^k)_{i \in [n]}$  is quasi-concave for any  $k$  with respect to the same order. Also note that, because for each  $k \in [d]$ ,  $\max R_i^k = \max \overline{R}_i^k$ , this is the allocation that arises from separate monopoly pricing.<sup>3</sup>

In Theorem 2, the condition of monotone marginal rates of substitution serves to link virtual valuations of different items, as our proof will show. Conversely, the weak monotonicity condition ensures the implementability of the allocation in (6), i.e., the existence of a price vector  $(t_i)_{i \in [n]}$  such that the mechanism  $(q^*, t)$  is incentive compatible and individually rational.

To prove Theorem 2, we construct a flow  $\lambda$  that induces (for some items) a revenue sequence equal to  $\overline{R}^k$ . In particular, we construct an Ironing Algorithm that, if well-defined (Lemma 6), produces a flow  $\lambda$  that is both non-negative and feasible (Lemma 4). The rest of the proof mainly consists of showing that the output of the Ironing Algorithm satisfies virtual welfare maximization (Lemma 3 and Lemma 7), as well as complementary slackness (Lemma 7).

The next lemma is a main structural tool to link different items. For  $k \in [d]$ ,  $i \in [n]$  and

<sup>3</sup>In section 5, we further explore the relationship between upgrade pricing and separate pricing, by showing conditions under which the allocation (6) can be implemented by a vector of single-item prices.

a flow  $\lambda$ , denote the normalized virtual value by

$$\nu_i^{\lambda,k} := \frac{\phi_i^{k,\lambda}}{\theta_i^k}.$$

The property that we will use repeatedly is that  $\nu_i^{k,\lambda}$  has the same sign as  $\phi_i^{\lambda,k}$ . We call a flow *downward* if  $\lambda_{ji} \geq 0$  for  $i, j \in [n]$  implies that  $j > i$ .

**Lemma 2.** *Let  $\Theta$  have monotone MRS. For any non-negative downward flow  $\lambda$ ,  $\nu_i^{\lambda,k} \leq \nu_i^{\lambda,l}$  for any  $1 \leq k \leq l \leq d$  and  $i \in [n]$ .*

*Proof.* It follows from definitions and monotone marginal rates of substitution that

$$\begin{aligned} \frac{\phi_i^{\lambda,k}}{\theta_i^k} &= \frac{\theta_i^k - \sum_{j=1}^n \lambda_{ji}(\theta_j^k - \theta_i^k)}{\theta_i^k} = 1 + \sum_{j=i}^n \lambda_{ji} - \sum_{j=i}^n \lambda_{ji} \frac{\theta_j^k}{\theta_i^k} \\ &\leq 1 + \sum_{j=i}^n \lambda_{ji} - \sum_{j=i}^n \lambda_{ji} \frac{\theta_j^l}{\theta_i^l} = \frac{\theta_i^l - \sum_{j=1}^n \lambda_{ji}(\theta_j^l - \theta_i^l)}{\theta_i^l} = \frac{\phi_i^{\lambda,l}}{\theta_i^l}. \end{aligned}$$

□

The next Lemma shows that virtual welfare maximization reduces to virtual welfare maximization for the marginally bought and marginally not bought items.

**Lemma 3.** *Assume that  $\Theta$  has monotone MRS and there exists a non-negative downward flow  $\lambda$  such that*

$$\phi_i^{\lambda,k} \geq 0 \qquad \phi_i^{\lambda,k+1} \leq 0, \quad \text{for any } i \in [i^k, i^{k+1}].$$

*Then, the allocation in (6) maximizes virtual welfare.*

*Proof.* Fix  $i \in [n]$  such that  $i^k \leq i \leq i^{k+1}$ . Note that, to show virtual welfare maximization of the allocation (6), it is sufficient to show

$$\phi_i^{\lambda,l} \geq 0, \quad l \geq k \qquad \phi_i^{\lambda,l} \leq 0, \quad l \leq k.$$

This is implied by

$$\phi_i^{\lambda,k+1} \leq 0 \implies \phi_i^{\lambda,l} \leq 0, \quad l \geq k+1 \tag{7}$$

and

$$\phi_i^{\lambda,k} \geq 0 \implies \phi_i^{\lambda,l} \geq 0, \quad l < k. \tag{8}$$

As  $\phi_i^{\lambda,k}$  and  $\nu_i^{\lambda,k}$  are positive multiples of each other, implications (7) and (8) follow directly from Lemma 2. □

For  $k = 0$  and  $k = d$  this Lemma reduces ironing in our problem to (well-studied) one-dimensional ironing. Finding a flow that maximizes virtual welfare reduces by Lemma 3 to ironing the (one-dimensional) virtual values for items 1 and  $d$ ,  $\phi_i^1$  and  $\phi_i^d$ , respectively. In these cases we can iron the pseudo-revenue to its concave closure in a discrete variant of Myerson (1981)'s procedure.

From now, our discussion focusses on  $k \in [d-1]$  and  $i \in [i^k+1, i^{k+1}]$ , i.e. types where an ironing that ensures virtual welfare maximization for both item  $k$  and item  $k+1$  is needed.

The following algorithm will make use of  $\hat{\lambda}$  as defined in (5), the *initial flow* and of a generalization of the pseudo-revenue. Before we go into the algorithm, we define a (pseudo-)revenue sequence associated to a flow  $\lambda$ ,  $R_i^{\lambda,k}$  as

$$R_i^{k,\lambda} = \sum_{j=i}^n f_j \phi_j^{k,\lambda}.$$

Intuitively,  $R_i^{\lambda,k}$  can be seen as (the negative of) an anti-derivative in the sense of virtual values being slopes of the revenue curve,

$$\frac{R_i^{\lambda,k} - R_{i+1}^{\lambda,k}}{f_i} = \frac{\sum_{j=i}^n f_j \phi_j^{k,\lambda} - \sum_{j=i+1}^n f_j \phi_j^{k,\lambda}}{f_i} = \phi_i^{k,\lambda}.$$

In this intuition, our algorithm will adjust a flow by raising one point in a revenue sequence at a time, from right to left. We will prove that this will yield slopes of revenue sequences, which are virtual values, which have the correct sign for virtual welfare maximization.

$\lambda \leftarrow \hat{\lambda};$

**for**  $i = n$  **to** 1 **do**

    Let  $\gamma_i \in [0, 1]$  be maximal such that for

$$\begin{aligned} \lambda'_{ji} &:= \gamma_i \lambda_{ji}, \quad \forall n > j > i \\ \lambda'_{j(i-1)} &:= \lambda_{j(i-1)} + (1 - \gamma_i) \lambda_{ji}, \quad \forall n > j > i \\ \lambda'_{i(i-1)} &:= \lambda_{i(i-1)} - (1 - \gamma_i) \sum_{i'=i}^n \lambda_{i'i} \end{aligned} \tag{9}$$

    it holds that  $R_i^{\lambda',\kappa(i)} = \overline{R^{\kappa(i)}_i};$

$\lambda \leftarrow \lambda';$

**Return**  $\lambda';$

### **Algorithm:** Ironing

The flow (9) was used earlier in Haghpanah and Hartline (2020). An important difference is that Haghpanah and Hartline (2020) choose  $\gamma_i$  to iron the revenue sequence of the grand



bundle to the concave closure of pseudo-revenue, we iron to the quasi-concave closure of the pseudo-revenue of an item  $\kappa(i)$ . The parameter  $\gamma$  can be found as solution to a system of linear equations. We show that a solution  $\gamma_i \in [0, 1]$  exists in Lemma 6.

We first observe that the Ironing Algorithm outputs a flow which is non-negative and feasible.

**Lemma 4.** *The output flow of the Ironing Algorithm is non-negative and feasible.*

*Proof.* We first prove feasibility by induction from  $i = n$  to  $i = 1$ . Note that  $\hat{\lambda}$  is feasible as argued in the proof of Theorem 1, which starts the induction. We are done if we show that  $\lambda$ 's feasibility implies feasibility of  $\lambda'$ . This involves checking that the in- and out-flows into  $i, i - 1$ , and  $j > i$ , do not change. For  $j > i$ , the change is

$$\lambda'_{ji} - \lambda_{ji} + \lambda'_{j(i-1)} - \lambda_{j(i-1)} = \gamma\lambda_{ji} - \lambda_{ji} + \lambda_{j(i-1)} + (1 - \gamma_i)\lambda_{ji} - \lambda_{j(i-1)} = 0.$$

For  $i$  and  $i - 1$ , checking feasibility involves similar calculations which we omit.

Now consider non-negativity. Each  $\lambda_{ij}$  reduces at most once. More specifically, only if  $j = i - 1$  and during iteration  $i$ . In this case,

$$\lambda'_{i(i-1)} = \lambda_{i(i-1)} - (1 - \gamma_i) \sum_{i'=i}^n \lambda_{i'i} \geq \lambda_{i(i-1)} - \sum_{i'=i}^n \lambda_{i'i} = f_i \geq 0,$$

where we used that  $\lambda_{ir} = \hat{\lambda}_{ir}$ ,  $r < i - 1$ , which in particular implies that  $\lambda_{ir} = 0$ , and feasibility of the flow.  $\square$

Next observe that in the Ironing Algorithm, iteration  $i$  changes the revenue (for any item  $k$ ) only for type  $i$ .

**Lemma 5.** *For any iteration  $i$ ,  $R_j^{k,\lambda'} = R_j^{k,\lambda}$  for any  $j \neq i$ . In particular,  $\phi_j^{k,\lambda'} = \phi_j^{k,\lambda}$  for  $j \notin \{i - 1, i\}$ .*

*Proof.* First note that as the in-flow for higher types remains unchanged in iteration  $i$ ,  $\lambda'_{rj} = \lambda_{rj}$ ,  $r \in [n]$  and  $j > i$ , the revenue does not change,  $R_j^{k,\lambda'} = R_j^{k,\lambda}$ . For types  $j \leq i - 1$ , we check that the changes to virtual welfare on the types whose inflows do change,  $i$  and

$i - 1$ , cancel out. By definition of virtual values,

$$\begin{aligned}\phi_i^{\lambda',k} &= \phi_i^{\lambda,k} + \frac{1 - \gamma_i}{f_i} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) \\ \phi_{i-1}^{\lambda',k} &= \phi_{i-1}^{\lambda,k} + \frac{1 - \gamma_i}{f_{i-1}} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) - \frac{1 - \gamma_i}{f_{i-1}} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k) \\ &= \phi_i^{\lambda,k} - \frac{1 - \gamma_i}{f_i} \sum_{j=i}^n \lambda_{ji}(\theta_j^k - \theta_{i-1}^k).\end{aligned}$$

Hence,

$$f_{i-1}\phi_{i-1}^{\lambda',k} + f_i\phi_i^{\lambda',k} = f_{i-1}\phi_{i-1}^{\lambda,k} + f_i\phi_i^{\lambda,k}.$$

□

Before showing that  $\gamma_i$  in the algorithm always exists, which requires that there are no “changes in the item being ironed while ironing,” we define the function  $\kappa(i)$  that will be used in the algorithm.

By the second condition of mostly regular, we can define the function  $\kappa(i)$  separately for  $k \in [d]$  and  $i^k < i \leq i^{k+1}$ . Consider for  $l = k, k + 1$  the interval subsets within

$$M^k = \{i \in [i^k, i^{k+1}] \cap \mathbb{N} \mid R_i^k \neq \overline{R^k}_i \text{ or } R_i^{k+1} \neq \overline{R^{k+1}}_i\}.$$

By the first condition in the definition of mostly regular  $F$ , these have disjoint inclusion-maximal intervals. For each maximal interval  $A$ , it must either be that  $R_i^k \neq \overline{R^k}_i$  or  $R_i^{k+1} \neq \overline{R^{k+1}}_i$ . For  $i \in A$ , define  $\kappa(i)$  as any of  $k, k + 1$  that satisfy either inequality.

**Lemma 6.** *For each  $i \in [n]$ ,  $\gamma_i$  such that  $R_i^{\lambda_i(\gamma_i), \kappa(i)} = \overline{R^{\kappa(i)}}_i$  exists. In particular, the Ironing Algorithm is well-defined.*

*Proof.* Denote by

$$\overline{\phi^k}_i = \frac{\overline{R^k}_i - \overline{R^k}_{i+1}}{f_i}$$

the slope of the quasi-concave closure of pseudo-revenue of item  $k$  at type  $i$ . Note that by definition of the quasi-concave closure, and by definition of  $\kappa(i)$ , we have

$$\overline{\phi^k}_i \leq 0.$$

As all types are non-negative, we get that  $\overline{\phi^k}_i \leq 0 \leq \theta_i^k$ . Note that for  $\gamma_i = 0$ ,  $\phi_i^{k, \lambda'} = \theta_i^k$ .

This implies that for  $\gamma_i = 0$ , we get that

$$\begin{aligned}\overline{R}_i^k &= f_i \overline{\phi}_i^k + \overline{R}_{i+1}^k = f_i \overline{\phi}_i^k + R_{i+1}^{k,\lambda} \\ &\leq f_i \phi_i^{k,\lambda} + R_{i+1}^{k,\lambda} \leq R_i^{k,\lambda},\end{aligned}$$

where we used that  $\overline{R}_{i+1}^k = R_{i+1}^{k,\lambda}$ , which follows from the definition of  $\kappa(i)$ .

For  $\gamma_i = 1$ , we get by definition that  $R_i^{k,\lambda'} = R_i^{k,\lambda'} = R_i^{k,\lambda} \leq \overline{R}_i^k$ . Note also that  $\gamma_i \mapsto R_i^{k,\lambda}$  is a continuous function. The existence of  $\gamma_i$  as desired in the algorithm follows from the Intermediate Value Theorem.  $\square$

We conclude by showing that the output of the algorithm satisfies complementary slackness and the condition of Lemma 3 sufficient for virtual welfare maximization.

**Lemma 7.** *Assume that  $\Theta$  is has monotone MRS, and that  $F$  is mostly regular. Then,  $q^*$  maximizes virtual welfare and satisfies complementary slackness with respect to  $\lambda'$ , the output of the Ironing Algorithm.*

*Proof.* Denote the final output of the Ironing Algorithm by  $\lambda$ . We would like to show that for any  $i^k \leq i \leq i^{k+1}$  we have that  $\phi_i^{k,\lambda} \geq 0$  and  $\phi_i^{k+1,\lambda} \leq 0$ . Recall that the Ironing Algorithm chooses which item to iron according to ironing mapping  $\kappa: I \rightarrow [d]$ . By Lemma 3, virtual welfare maximization follows. We call the piece-wise constant intervals of the ironing interval  $\kappa$  *ironing intervals*. Fix an ironing interval  $M$ . Define  $M' := M \setminus \{\max M\}$ . Then for  $i \in M'$ , as

$$\phi_i^{\kappa(i),\lambda} = \nu_i^{\kappa(i),\lambda} = 0$$

we have for  $\kappa(i) = k$  that  $\nu_i^{\kappa(i),\lambda} \geq 0$  by Lemma 2 and hence  $\phi_i^{\kappa(i),\lambda} \geq 0$ . Similarly, we have for  $\kappa(i) = k + 1$  that  $\nu_i^{\kappa(i),\lambda} \leq 0$  by Lemma 2 and hence  $\phi_i^{\kappa(i),\lambda} \leq 0$ .

For any ironing interval, it remains to consider  $i = \min M - 1$  or  $i = \max M$ .

**Case 1:**  $i = \max M$  **and**  $\kappa(i) = k$  By definition of the quasi-concave closure,  $\overline{R}_i^k = R_i^k$ , the algorithm chooses  $\gamma_i = 1$  and hence  $\phi_i^{\lambda,k+1} = \phi_i^{k+1} \leq 0$  as by the first part of mostly regularity,  $i$  cannot be a part of an ironing interval.

**Case 2:**  $i = \min M - 1$  **and**  $\kappa(i) = k + 1$  By definition of the quasi-concave closure,  $\overline{R}_i^{k+1} = R_i^{k+1}$ , the algorithm chooses  $\gamma_i = 1$  and hence  $\phi_i^{\lambda,k} = \phi_i^k \geq 0$  as by the first part of mostly regularity,  $i$  cannot be a part of an ironing interval.

**Case 3:**  $i = \max M$  **and**  $\kappa(i) = k + 1$  By definition of the quasi-concave closure, we know

that  $\phi_i^{\kappa(i)} \leq \phi_i^{\kappa(i),\lambda} = \overline{\phi^{\kappa(i)}}_i \leq 0$ . Note that in this case, the derivative is

$$\frac{\partial \phi_i^{\kappa(i),\lambda'}}{\partial \gamma_i} = -\frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} (\theta_j^{\kappa(i)} - \theta_i^{\kappa(i)}),$$

which must be non-positive because  $\phi_i^{\kappa(i)} \leq \phi_i^{\kappa(i),\lambda}$ ,  $\phi_i^{\kappa(i)}$  is linear in  $\gamma_i$ , and the algorithm chooses a value  $\gamma_i < 1$ . (At  $\gamma_i = 1$ ,  $\phi_i^{\kappa(i)} = \phi_i^{\kappa(i),\lambda}$ .) Furthermore, the monotone MRS property implies

$$\frac{\theta_j^{\kappa(i)}}{\theta_i^{\kappa(i)}} = \frac{\theta_j^{k+1}}{\theta_i^{k+1}} \leq \frac{\theta_j^k}{\theta_i^k}, \quad \text{for all } j \geq i$$

and hence

$$0 \geq \frac{1}{\theta_i^{\kappa(i)}} \frac{\partial \phi_i^{\kappa(i),\lambda'}}{\partial \gamma} = -\frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} \left( \frac{\theta_j^{\kappa(i)}}{\theta_i^{\kappa(i)}} - 1 \right) \geq -\frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} \left( \frac{\theta_j^k}{\theta_i^k} - 1 \right) = \frac{1}{\theta_i^k} \frac{\partial \phi_i^{k,\lambda'}}{\partial \gamma}.$$

Because the algorithm chooses  $\gamma_i < 1$ , this implies

$$\phi_i^{k,\lambda} \geq \phi_i^k \geq 0.$$

**Case 4:**  $i = \min M - 1$  and  $\kappa(i) = k$  The derivative of the virtual value of the next item at the left end of the ironing interval, type  $i$  is given by

$$\begin{aligned} \frac{\partial \phi_i^{k+1,\lambda'}}{\partial \gamma} &= \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} (\theta_j^{k+1} - \theta_i^{k+1}) - \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} (\theta_{i+1}^{k+1} - \theta_i^{k+1}) \\ &= \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} (\theta_j^{k+1} - \theta_{i+1}^{k+1}) \end{aligned}$$

Because  $\Theta$  and  $F$  are not too shuffled, we have

$$\frac{\partial \phi_i^{k+1,\lambda'}}{\partial \gamma} = \frac{1}{f_i} \sum_{j=i+1}^n \lambda_{ji} (\theta_j^{k+1} - \theta_{i+1}^{k+1}) \geq 0$$

Note that  $0 \geq \phi_i^{k+1}$  by no overlap. Moreover,  $\phi_i^{k+1} \geq \phi_i^{k+1,\lambda}$  because  $\partial \phi_i^{k+1,\lambda'} / \partial \gamma \geq 0$  and the algorithm chooses  $\gamma_i < 1$ . Combining these observations, we obtain  $\phi_i^{k+1,\lambda} \leq 0$ .

To show complementary slackness, observe that whenever  $\overline{R^{\kappa(i)}}_i = R_i^{\kappa(i)}$ , the algorithm chooses  $\gamma_i = 1$  and the final output of the algorithm  $\lambda$  is such that for  $j > i > r$  and for  $j > i + 1$  and  $i = r$ , we have  $\lambda_{jr} = 0$ .

By the second condition in the assumption of a mostly regular distribution, this implies

that for  $j > i^k > r$ ,  $\lambda_{jr} = 0$ .

Moreover, the algorithm's output satisfies for  $j < i$ ,  $\lambda_{ji} = 0$  by construction.

Finally, complementary slackness requires that downward incentive constraints for  $i, j \in [n]$  such that  $i^k \leq i, j \leq i^{k+1}$  are tight, which follows from these types getting the same allocation.  $\square$

We are now ready to finish the proof of Theorem 2.

*Proof of Theorem 2.* Implementability follows from weak monotonicity and the definition of the optimal mechanism, (6). Non-negativity and feasibility of flow are properties of the Ironing Algorithm shown in Lemma 4. Virtual welfare maximization and complementary slackness have been shown in Lemma 7.  $\square$

## 5 Upgrade Pricing and Separate Pricing

In both Theorem 1 and Theorem 2, we established the optimality of an upgrade pricing mechanism that yields the same allocation as separate (item by item) monopoly pricing, though not necessarily the same transfers.

We say that the type space  $\Theta$  is *monotone* if  $\theta_i^k \leq \theta_j^k$  for any  $i < j \in [n]$  and  $k \in [d]$ .

In some instances, including the examples in Figures 1 and 3 above, the optimal upgrade pricing mechanism can be implemented through separate pricing, i.e., bundling is redundant. We now explore the relationship between upgrade pricing and separate pricing, and we show that type monotonicity is both necessary and sufficient to establish an equivalence result between these two classes of mechanisms.

We call a mechanism separate pricing if a type separately chooses to buy (or not) each item  $k$  at a price  $p_k$ . Formally, a mechanism satisfies separate pricing if it can be written as:

$$q_i^k = \begin{cases} 1 & \theta_i^k \geq p_k \\ 0 & \text{else,} \end{cases} \quad t_i = \sum_{k=1}^d p_k \mathbb{1}_{q_i^k=1}.$$

**Theorem 3.** *If the type space  $\Theta$  is monotone, then the allocation of any upgrade pricing mechanism can be implemented via separate pricing, and conversely. When the type space is not monotone, neither implication needs to hold.*

*Proof.* We first assume types are monotone and show that the allocation and revenue of any upgrade pricing mechanism can be obtained through a separate pricing mechanism, and vice versa.

Let  $\theta_1 \leq \dots \leq \theta_i \leq \dots \leq \theta_n$ , and fix an upgrade pricing mechanism  $\mathcal{M}$ . This mechanism admits an indirect representation as (a) a collection of bundles ranked by set inclusion  $\{b_k\}_{k=0}^K$ , with  $b_0 = \emptyset$  and  $K \leq d$ , and (b) a vector of prices  $t_k$  that are increasing in  $k$ , with  $t_0 = 0$ . Let  $\underline{\theta}_k$  and  $\bar{\theta}_k$  denote the lowest and highest types who choose bundle  $b_k$  under mechanism  $\mathcal{M}$ . Because types are monotone, buyer self-selection implies  $\underline{\theta}_k \geq \bar{\theta}_{k-1}$ .

We now construct a separate pricing mechanism, i.e., a vector of prices  $\{p_j\}_{j=1}^d$  that yields the same allocation and payments as our upgrade pricing mechanism. To do so, define the collection of *upgrades*  $u_k := b_k \setminus b_{k-1}$  and the upgrade prices  $\tau_k := t_k - t_{k-1}$ . For each upgrade bundle  $k$  and every good  $j \in u_k$ , let the single-item prices  $p_j$  satisfy

$$p_j \in [\bar{\theta}_{k-1}^j, \underline{\theta}_k^j] \quad \text{and} \quad \sum_{j \in u_k} p_j = \tau_k.$$

Under monotonicity, such a vector of prices always exists. By consumer self-selection in the original mechanism  $\mathcal{M}$ , we have

$$\begin{aligned} \bar{\theta}_{k-1} b_k - t_k &\leq \bar{\theta}_{k-1} b_{k-1} - t_{k-1}, \\ \underline{\theta}_k b_{k-1} - t_{k-1} &\leq \underline{\theta}_k b_k - t_k. \end{aligned}$$

In turn, this implies

$$\bar{\theta}_{k-1} u_k \leq \tau_k \leq \underline{\theta}_k u_k.$$

With the prices so constructed, each type purchases the same goods as under  $\mathcal{M}$  and pays the same total price. Notice first that each type's choice from the original mechanism  $\mathcal{M}$  is still available at the same price, i.e., each bundle  $b_k$  can still be purchased for a total price  $t_k$ . Moreover, by monotonicity, no type  $\theta$  who buys bundle  $b_k$  under the upgrade pricing mechanism  $\mathcal{M}$  derives positive net surplus from any object  $j \in u_{k'}$  with  $k' > k$  under the separate prices constructed above. And finally, no such type  $\theta$  derives positive net surplus by removing any object  $j \in u_{k'}$  with  $k' \leq k$  from her consumption bundle.

The other direction of this result is immediate: if types are monotone, the goods purchased by two different types under any separate pricing mechanism are ranked by set inclusion. Thus, replacing the separate pricing mechanism with the resulting upgrade pricing mechanism yields the same outcome.

Finally, we now show by means of two counterexamples that, without type monotonicity, separate pricing is not equivalent to upgrade pricing.

In particular, there exist type spaces and vectors of separate prices that do not induce

an upgrade pricing allocation. For example, let

$$\Theta = \{(1, 1), (1, 3), (3, 3), (4, 1)\}$$

and consider the separate prices  $p = (2, 2)$ : type  $\theta_2$  buys good 2 only, type  $\theta_3$  buys both goods, and type  $\theta_4$  buys good 1 only.

Likewise, for the same type space, consider the upgrade pricing mechanism where  $q = (0, 1)$  is sold for  $t = 2$  and  $q = (1, 1)$  is sold for  $t = 4$ , i.e., good  $j = 1$  is only sold as an upgrade, for an additional price  $\tau = 2$ . Under this mechanism, type  $\theta_2$  buys good 2 only, while types  $\theta_3$  and type  $\theta_4$  buy both goods. However, as we saw above, the vector of separate prices  $p = (2, 2)$  yields a different allocation (and a lower revenue for the seller).  $\square$

Whenever an upgrade pricing mechanism implements the allocation of optimal separate pricing, by construction, each marginal type  $\theta_k$  is indifferent between the two consecutive bundles  $b_{k-1}$  and  $b_k$ . Theorem 3 then implies that the outcome of this mechanism can be implemented by the separate monopoly prices.

**Corollary 1.** *If  $\Theta$  is monotone,  $q$  is an allocation of an optimal upgrade pricing mechanism, and  $q$  is the allocation of separate monopoly pricing, then separate monopoly pricing is optimal.*

Adding a monotonicity condition to both of our main theorems, Theorem 1 and Theorem 2, we hence obtain two sets of sufficient conditions under which separate monopoly pricing is optimal.

**Corollary 2.** *If  $\Theta$  is monotone and  $F$  is regular, separate monopoly pricing is optimal.*

**Corollary 3.** *If  $\Theta$  is monotone and has a monotone marginal rates of substitution, and  $F$  is mostly regular, then separate monopoly pricing is optimal.*

Finally, we show by means of examples that a monotone marginal rate of substitution and monotonicity are important yet distinct conditions in establishing the optimality of upgrade pricing and separate pricing, respectively. Consider first the following example:

**Example 1** (MRS without Monotonicity). *Let  $\Theta = \{(7/8, 1/8), (5/4, 3/4), (3/4, 5/4)\}$  and  $f = (1/2, 1/4, 1/4)$ . Note that  $f$  satisfies regularity and weak monotonicity, and that  $\Theta$  has the monotone MRS property, as shown by Figure 5 below.*

*In the optimal mechanism, the seller charges  $7/8$  for good 1 (which every type buys) and  $3/4$  for adding good 2, which types  $\theta_2$  and  $\theta_3$  do. This is also the optimal allocation under separate pricing, as indicated by the virtual values. However, offering each item separately*

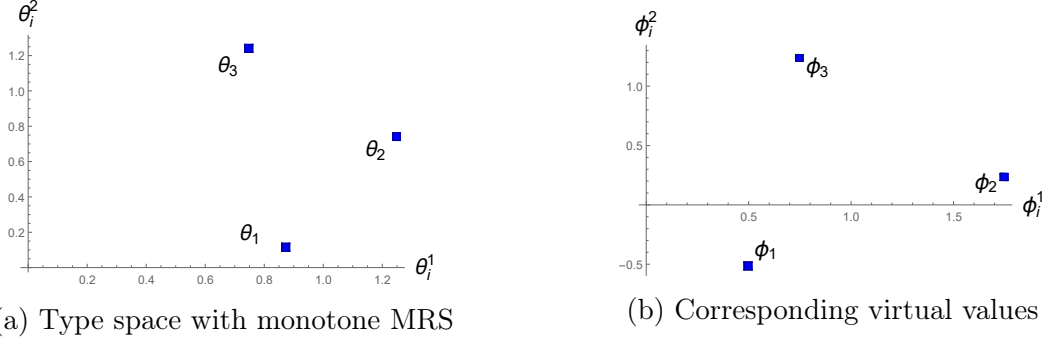


Figure 5: A type space with monotone MRS and its corresponding virtual values

at prices  $(7/8, 3/4)$  means type  $\theta_3$  will deviate and buy good 2 only. In other words, upgrade pricing allows the seller to raise the price of good 1 from the standalone monopoly price of  $3/4$  to  $7/8$ .

Conversely, there are cases, such as Example 2 below, where types are monotone but their marginal rates of substitution are not ordered, and the optimal mechanism is not upgrade pricing.<sup>4</sup>

**Example 2** (Monotonicity without MRS). Let  $f = \{1/4, 1/4, 1/2\}$  and  $\Theta = \{(5/8, 7/16), (3/2, 1/2), (2, 1)\}$ . Note that these types are monotone, but do not satisfy monotone MRS, since  $\theta_2^1/\theta_2^2 > \theta_1^1/\theta_1^2$  and  $\theta_2^1/\theta_2^2 > \theta_3^1/\theta_3^2$ . The optimal allocation is described by the virtual values in Figure 6 below. In the optimal mechanism, the seller charges  $7/16$  for

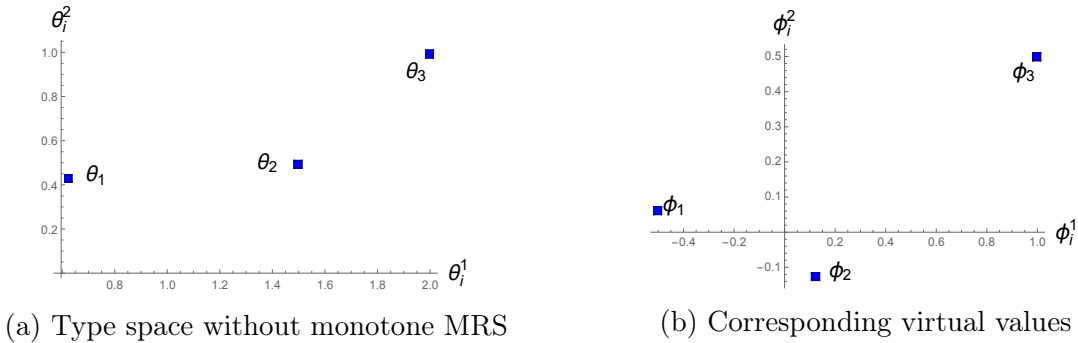


Figure 6: A type space without monotone MRS and its corresponding virtual values

good 2,  $23/16$  for good 1, and  $39/16$  for the grand bundle. Type  $\theta_1$  buys good 2, type  $\theta_2$  buys good 1, and  $\theta_3$  buys the grand bundle. The seller's expected revenue is equal to  $27/16$ . No upgrade pricing scheme can implement a mechanism with this revenue.

<sup>4</sup>Likewise, it is clear how to construct examples with non-monotone types, whose MRS are ordered and where upgrade pricing is not optimal.



## 6 Beyond the Sufficient Conditions

In this section, we explore the scope of upgrade pricing optimality beyond our sufficient conditions by means of numerical analysis. We consider an environment with three buyer types,  $\Theta = (\theta_1, \theta_2, \theta_3)$ . The first two types are fixed to be either  $\{\theta_1, \theta_2\} = \{(2, 2), (4, 4)\}$ , or  $\{\theta_1, \theta_2\} = \{(1, 2), (4, 4)\}$ , whereas the third type  $\theta_3$  takes values in  $[0, 6]^2$ . Regarding the prior type distribution, we consider two scenarios. In the first scenario, the prior distribution

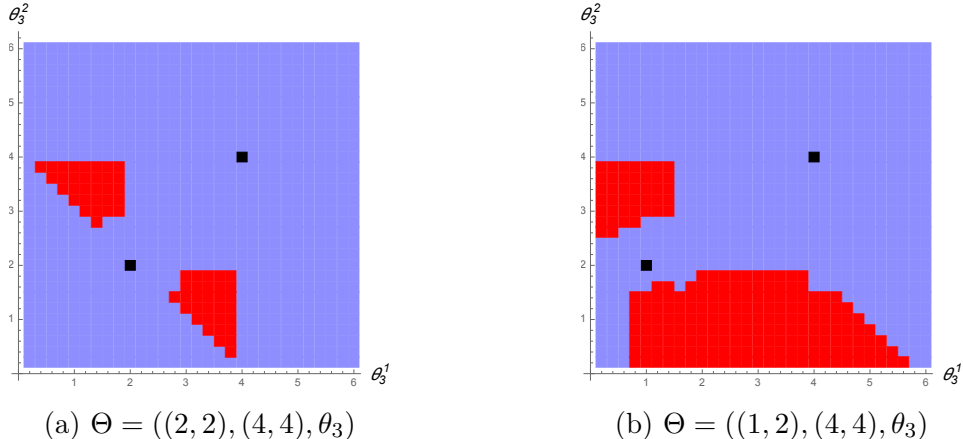


Figure 7: Optimality of upgrade pricing mechanisms for a uniform prior distribution.

is fixed to be uniform  $f = (1/3, 1/3, 1/3)$ . The type  $\theta_3$  takes strictly positive values over the uniform grid of size  $1/5$ ,  $\theta_3 \in ((1/5, 1/5), \dots, (6, 6))$ . For each type  $\theta_3$ , we calculate an optimal mechanism as an exact numerical solution to the seller’s linear program. We color the corresponding point blue if the mechanism features upgrade pricing, i.e., the optimal allocation admits order by inclusion, and color the point red otherwise. Figure 7 presents the results for the two cases of the first two types. They suggest that for a given prior distribution, an upgrade pricing may be optimal for many type configurations, well beyond our sufficient conditions. In the second scenario, we explore the “robust” optimality of upgrade pricing with respect to the prior distribution, in the spirit of Haghpahan and Hartline (2020). Type  $\theta_3$  takes the same values as in the first scenario. For each  $\theta_3$ , we calculate an optimal mechanism for each prior distribution located on the uniform grid over the interior of  $\Delta(\Theta)$  of size  $1/25$ , i.e.,  $f \in ((23/25, 1/25, 1/25), \dots, (1/25, 1/25, 23/25))$ . We color the corresponding point green if *all* of these mechanisms feature upgrade pricing and red otherwise.

Figure 8 presents the results for both  $\{\theta_1, \theta_2\} = \{(2, 2), (4, 4)\}$  and  $\{\theta_1, \theta_2\} = \{(1, 2), (4, 4)\}$ . These calculations suggest two important conclusions. First, type monotonicity and monotone MRS are sufficient for the optimality of upgrade pricing regardless of the distribution of types. In that sense, there seems to be scope for relaxing our Most Regularity condition. Second, for monotone type spaces, the monotone MRS condition appears to be, in fact,

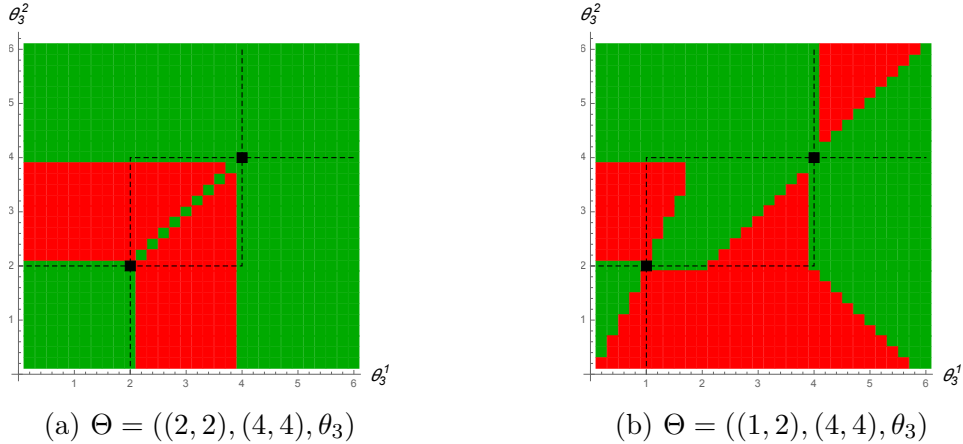


Figure 8: Robust optimality of upgrade pricing across various prior distributions.

*necessary* for upgrade pricing to be robustly optimal. Finally, as we have already seen, the monotone MRS property is neither necessary nor sufficient when types are not monotone.

## 7 Conclusion

It is a common practice for a seller to offer bundles of products or services that are ordered in a way that more expensive bundles contain all items from less expensive bundles as well as some extra items. In this paper, we provide sufficient conditions under which such “upgrade pricing” schemes are exactly optimal for a monopolist seller.

There are several ways in which the current analysis could be extended. First, our conditions could be relaxed to account for richer type distributions. One natural extension can be obtained immediately: assume that a type distribution can be split into several type cohorts such that each type cohort satisfies the conditions of our theorems. Our results imply that the optimal mechanisms in each respective cohort are upgrade pricing. Furthermore, if optimal prices are the same in all those schemes, then the upgrade pricing with those prices is an optimal mechanism for the compound type distribution. We leave it to future work to characterize when such decomposition exists.

Second, our sufficient conditions for the optimality of upgrade pricing may be complemented by necessary conditions. In doing so, one may want to distinguish between conditions on type distributions and type spaces. For the latter, one may ask which type spaces guarantee that upgrade pricing is optimal irrespective of the type distribution. Figure 7 provides a first step in this direction.

Finally, throughout the paper we highlight the interplay between optimality of different pricing schemes: bundling, upgrade pricing, and separate sales. It would be instructive to

provide a more complete characterization of the cases in which one of these schemes strictly outperforms another.

## References

- W. J. Adams and J. L. Yellen. Commodity Bundling and the Burden of Monopoly. *Quarterly Journal of Economics*, 90(3):475–498, 1976.
- M. Babaioff, N. Immorlica, B. Lucier, and S. M. Weinberg. A Simple and Approximately Optimal Mechanism for an Additive Buyer. *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 21–30, 2014. doi: 10.1109/focs.2014.11.
- S. Bikhchandani and D. Mishra. Selling two identical objects. *arXiv:2009.11545*, 2020.
- S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. ISBN 9780521833783. doi: 10.1198/jasa.2005.s41.
- Y. Cai, N. R. Devanur, and S. M. Weinberg. A Duality Based Unified Approach to Bayesian Mechanism Design. *Proceedings of the Annual ACM Symposium on Theory of Computing*, 19-21 June, 2016. ISSN 07378017. doi: 10.1145/2897518.2897645.
- C. Daskalakis, A. Deckelbaum, and C. Tzamos. Strong Duality for a Multiple-Good Monopolist. *Econometrica*, 85(3):735–767, 2017. ISSN 0012-9682. doi: 10.3982/ecta12618.
- N. R. Devanur, K. Goldner, R. R. Saxena, A. Schwartzman, and S. M. Weinberg. Optimal mechanism design for single-minded agents. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 193–256, 2020.
- G. Ellison. A Model of Add-On Pricing. *Quarterly Journal of Economics*, 120(2):585–637, 2005.
- A. Fiat, K. Goldner, A. R. Karlin, and E. Koutsoupias. The fedex problem. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 21–22, 2016.
- S. Ghili. A Characterization for Optimal Bundling of Products with Interdependent Values. Technical report, Yale University, 2021.
- N. Haghpahan and J. Hartline. When Is Pure Bundling Optimal? *The Review of Economic Studies*, 88(3):1127–1156, 08 2020. ISSN 0034-6527. doi: 10.1093/restud/rdaa053. URL <https://doi.org/10.1093/restud/rdaa053>.
- S. Hart and N. Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347, 2017.
- B. H. Korte and J. Vygen. *Combinatorial optimization*, volume 1. Springer, 2011.

- A. M. Manelli and D. R. Vincent. Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35, 2006.
- R. P. McAfee, J. McMillan, and M. D. Whinston. Multiproduct Monopoly, Commodity Bundling, and Correlation of Values. *Quarterly Journal of Economics*, 104(2):371–383, 1989.
- R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.
- G. Pavlov. Optimal mechanism for selling two goods. *The BE Journal of Theoretical Economics*, 11(1), 2011.
- J. Philips. Don't look now, but the great unbundling has spun into reverse. *New York Times*, 2017.