

Reverse Mechanism Design

Nima Haghpanah
MIT
EECS and Sloan School of Management
nima@eecs.mit.edu

Jason Hartline
Northwestern University
EECS Department
hartline@northwestern.edu

March 23, 2015

Abstract

Myerson’s 1981 characterization of revenue-optimal auctions for single-dimensional agents follows from an *amortized analysis* of the incentives: Virtual values that account for expected revenue are derived using integration by parts and are optimized pointwise by an incentive compatible mechanism. A challenge of generalizing the approach to multi-dimensional agents is that a mechanism that pointwise optimizes “virtual values” resulting from a general application of integration by parts is not incentive compatible.

We give a framework for *reverse mechanism design*. Instead of solving for the optimal mechanism in general, we hypothesize a (natural) specific form of the optimal mechanism and identify conditions for existence of virtual values that prove the mechanism is optimal. As examples, we derive conditions for the optimality of mechanisms that sell each agent her favorite item or nothing for unit demand agents, and for the optimality of posting a single price for the grand bundle for additive agents.

1 Introduction

Optimal mechanisms for agents with multi-dimensional preferences are generally complex. This complexity makes them challenging to solve for and impractical to run. In a typical mechanism design approach, a model is posited and then the optimal mechanism is designed for the model. Successful mechanism design gives mechanisms that one could at least imagine running. By this measure, multi-dimensional mechanism design has had only limited success. In this paper we take the opposite approach, which we term *reverse mechanism design*. We start by hypothesizing the optimality of a particular form of mechanism that is simple and reasonable to run, then we solve for sufficient conditions for the mechanism to be optimal (among all mechanisms). Our approach is successful if the conditions under which the hypothesized mechanism is indeed optimal are broad and representative of relevant settings.

This paper has two main contributions. The first is in codifying the *method of virtual values* from single-dimensional auction theory and extending it to agents with multi-dimensional preferences. The second is in applying this method to two paradigmatic classes of multi-dimensional preferences. The first class is unit-demand preferences (e.g., a homebuyer who wishes to buy at most one house); for this class we give sufficient conditions under which posting a uniform price for each item is optimal. This result generalizes one of Alaei et al. (2013) for a consumer with values uniform on interval $[0, 1]$, and contrasts with an example of Thanassoulis (2004) for a consumer with values uniform on interval $[5, 6]$ where uniform pricing is not optimal. The second class is additive preferences, for this class we give sufficient conditions under which posting a price for the grand bundle is optimal. This result generalizes a recent result of Hart and Nisan (2012) and relates to work of Armstrong (1999). Similarly to an approach of Alaei et al. (2013), these results for single-agent pricing problems can be generalized naturally to multi-agent auction problems.

Myerson's (1981) characterization of revenue optimal auctions for single-dimensional agents is the cornerstone of modern auction theory and mechanism design. This characterization is successful in describing simple and practical mechanisms in simple environments where the agents preferences are independent and identically distributed according to a well-behaved distribution. In this case, the optimal auction is reserve price based. Myerson's characterization is also successful in describing the complex optimal mechanism for agents with preferences that are non-identically distributed or distributed according to an ill-behaved distribution. However, due to this complexity, the resulting mechanisms have limited application. The consequence of our work is similar in that we characterize simple optimal mechanisms for well-behaved preferences; but distinct in that it does not characterize optimal mechanisms beyond the class of well-behaved preferences.

Myerson's approach is based on mapping agent *values* to appropriately defined *virtual values* and then optimizing the *virtual surplus*, i.e., the sum of the virtual values of agents served. Importantly, this approach replaces the global objective of optimizing revenue in expectation over the distribution of agent values with the pointwise objective of optimizing virtual surplus on each profile of agent values. Furthermore, virtual surplus maximization leads to a simple and practical optimal mechanism in many environments. The simplicity of analysis by virtual values and of mechanisms resulting from optimizing virtual values has led to a rich single-dimensional auction theory. Our multi-dimensional virtual values similarly give a pointwise objective and their optimization results in simple optimal mechanisms.

In the remainder of this section we review virtual values in mechanism design for single-dimensional agents, describe the challenges in identifying virtual values for multi-dimensional preferences, and describe how restricting the form of the hypothesized optimal mechanism makes it

possible to solve for multi-dimensional virtual values. To keep the exposition simple, this introduction restricts its discussion to single-agent mechanisms.

Optimal Single-dimensional Mechanisms for Revenue. Revenue is a challenging objective for mechanism design because, due to the agents' incentives, it cannot be optimized pointwise. To illustrate this fact, consider a single agent wishing to buy a single item and whose single-dimensional value for the item is drawn uniformly from the $[0, 1]$ interval. If we post a take-it-or-leave-it price of 0.3 then the agent will buy when his value is greater than 0.3 and pay 0.3; if we post a price of 0.5 the agent will buy when his value is greater than 0.5 and pay 0.5. If our agent has value 0.3 then the first mechanism has the best possible revenue and the second mechanism has an inferior revenue of zero. If our agent has value 0.5 then the second mechanism has the best possible revenue of 0.5 and the first mechanism has an inferior revenue of 0.3. While we cannot rank these mechanisms by revenue pointwise, i.e., for all values the agent might possess, we can rank mechanisms and optimize revenue tradeoffs across values in expectation over the distribution from which the agent's value is drawn. With a uniform value on $[0, 1]$, the agent buys in the first mechanism at a price of 0.3 with probability 0.7 for an expected revenue of 0.21, and in the second mechanism at a price of 0.5 with probability 0.5 for an expected revenue of 0.25.

Myerson (1981) solved for revenue optimal auctions for agents with single-dimensional preferences. He gives a definition of virtual values, claims that optimization of virtual surplus (the expected virtual value of the allocation rule, where the allocation rule maps types to probability of service) gives the optimal auction, and proves this claim by utilizing two properties of the virtual surplus. The following two properties collectively imply that the mechanism that optimizes virtual surplus is indeed the optimal mechanism.¹ A *virtual value function* maps values pointwise to virtual values satisfying two properties:

- The pointwise optimization of virtual surplus gives an allocation rule that is *incentive compatible*. That is, there exist payments for this allocation rule that induce an agent to truthfully report his value.
- The virtual values are an *amortization* of the revenue. That is, the expected the virtual surplus of any incentive compatible mechanism is equal to the expected revenue of that mechanism.²

This definition of virtual value functions gives a roadmap for identifying the optimal mechanism: find a virtual value function (that satisfies the two conditions) and run the mechanism that maximizes virtual surplus pointwise. The identification of a virtual value function reduces the problem of optimization of the expected revenue (a global quantity) to the optimization of virtual surplus (a pointwise quantity).

In our single-dimensional example of the uniform distribution above, the appropriate virtual value function is $2v - 1$. Optimizing virtual surplus pointwise (for a single agent) means serving

¹Single-dimensional virtual values satisfy this conditions simultaneously only for regular distributions. Subsequently, in Section 3, we will define a more permissive version of the amortization property which corresponds to the *ironed virtual values* of Myerson (1981). The simpler definition here will, nonetheless, be sufficient for our introductory discussion.

²In the design and analysis of algorithms, an *amortized analysis* is one where the contributions of local decisions to a global objective are indirectly accounted for (see Borodin and El-Yaniv, 1998). The correctness of such an indirect accounting is often proven via a *charging argument*. For the analysis of the expected payment of a single-dimensional agent, a low type is charged for the loss of revenue from all higher types and this accounting gives rise to the familiar virtual value formulation.

the agent if his virtual value is positive and not serving him if his value is negative. For this virtual value function, the agent will be served if his value is least 0.5. This virtual value function satisfies the incentive compatibility condition: serving the agent if her value is above 0.5 is incentive compatible and the appropriate payment is 0.5 if the agent is served and zero otherwise. This virtual value function satisfies the amortization condition: for posting any price p , the expected virtual value of the agent served and the expected revenue are equal. The former can be calculated by $\int_p^1 (2v - 1) dv = p(1 - p)$ and the latter was calculated above as $p(1 - p)$.³

The Challenge of Multi-dimensional Preferences. As described above, if a virtual value function that satisfies the incentive compatibility and amortization conditions can be identified then the optimal mechanism design problem is solved. Integration by parts can be used on paths connecting each type to the origin to obtain an amortization. For agents with single-dimensional preferences, uniqueness of such paths results in a unique amortization. It then remains to check the incentive compatibility property, i.e., that pointwise optimization of virtual surplus is incentive compatible; by standard characterizations of incentive compatible mechanisms, this is a simple task as well. Integration by parts can be used in higher dimensions as well to derive amortizations, parameterized by a consistent choice of paths connecting each point in the type space to the origin. However, not all choices of consistent paths give an amortization that also satisfy the incentive compatibility requirements. The challenge is to find the right choice of paths, such that pointwise optimization of resulting virtual welfare gives an incentive compatible mechanism. This difficulty has prevented the design of mechanisms for multi-dimensional agents that follows the virtual-value-based approach (for example see Hart and Reny (2014) for a discussion on single-dimensional vs. multi-dimensional incentive constraints).

Multi-dimensional Virtual Values. To resolve the non-uniqueness of functions that satisfy the amortization property we consider additional constraints that the optimality of the hypothesized mechanism would place on virtual values. These additional constraints reduce a degree of freedom in the integration by parts and give result in a unique amortization. From this amortization, optimality of the hypothesized mechanism (and thus, incentive compatibility) can be checked.

We walk through this approach for the example of a unit-demand agent and the hypothesized optimality of mechanisms that post a uniform price for each item. On one hand, under uniform pricing, the agent will always choose to buy his favorite item, or no item if all values are below the price. On the other hand, a mechanism that optimizes a multi-dimensional virtual surplus (for a unit-demand agent) would serve the agent the item he has the highest positive virtual value for, or no item if all virtual values are negative. Synthesizing these constraints, the following conditions are sufficient for virtual surplus maximization to imply optimality of uniform pricing.

- The virtual value function is a *single-dimensional projection* if the virtual value for the favorite item corresponds to the single-dimensional virtual value for the distribution of the value for the favorite item (i.e., the distribution of the maximum value).⁴
- The virtual value function is *consistent with uniform pricing* if there is a price such that (a) when the value for the favorite item exceeds the price then the virtual value for the favorite

³Further discussion omitted, this amortized equivalence continues to hold for randomized mechanisms.

⁴Rationale: The mechanism has effectively projected the agent's multi-dimensional preference onto a single dimension. In this single dimension the unique function that satisfies the amortization property is the one given by the single-dimensional virtual values of Myerson (1981).

item is non-negative and at least the virtual value of any other items and (b) when the value for the favorite item is below the price then both virtual values are non-positive.

Any virtual value function that satisfies the consistency-with-uniform-pricing conditions satisfies the incentive compatibility requirement. Thus, the identification of a virtual value function becomes one of simultaneously resolving the three conditions of amortization, single-dimensional projection, and consistency with uniform pricing. Conditions on the distribution over values that guarantee the existence of such a virtual value function are sufficient for the optimality of uniform pricing.

For two-dimensional preferences the amortization and single-dimensional-projection restrictions pin down a two-dimensional virtual value function uniquely. Specifically, as the virtual value for the favorite item is fixed by the restriction, only the virtual value for the other item must be determined. Essentially, we are left with a single-dimensional problem and in a single dimension the function that satisfies the amortization condition is unique. Our task is then to give sufficient conditions under which this amortization is consistent with uniform pricing.

Proving optimality of selling a unit-demand agent his favorite item provides our main example of the framework of reverse mechanism design. As a second example, we apply the framework to agents with additive preferences and give sufficient conditions for pricing the grand bundle to be optimal. The techniques we develop can be similarly applied to other environments and appropriate restrictions, and these extensions are an important topic for future work.

The opening example of this introduction of selling one of many houses to a homebuyer with values distributed independently, identically, and uniformly from $[0, 1]$ satisfies the conditions for optimality of selling him only his favorite item. More generally, our framework identifies a positive correlation property as a sufficient condition for the distribution over agent values. Such correlation is natural when values correlate, e.g., with initial wealth. We demonstrate an example use of our main theorems for unit-demand agents below.

Example 1 (Pricing with Delay). Consider selling an item for immediate delivery, delayed delivery (e.g., by express mail or standard mail), or lotteries thereof. The buyer's has a private value v for receiving the item immediately and, for $\delta \leq 1$, value $v \times \delta$ for delayed receipt of the item. Since the value for the item with delay is less than the value for the item immediately, a natural guess for the optimal mechanism is that it only allocates the item immediately or not at all. When δ is known to the seller, this intuition is proved correct in several related models such as Stokey (1979) and Acquisti and Varian (2005). In our setting but with a public discount factor δ , the optimality of selling immediately (or not at all) is a consequence of the “no haggling” result of Riley and Zeckhauser (1983) (and also Myerson, 1981).

When δ is only privately known to the buyer, however, uniform pricing is no longer generally optimal. (The example from Thanassoulis (2004) can be adapted to give a correlated distribution on (v, δ) where the optimal mechanism randomizes delivery time.) A direct application of one of our main theorems (Theorem 11), however, states that uniform pricing is optimal when δ and v are positively correlated and the marginal distribution of v is *regular*. Our correlation condition requires that the conditional distribution of δ for any v is first order stochastically dominated by the conditional distribution of δ for $v' \geq v$. We remove the regularity requirement by demanding a form of positive correlation that is slightly more restrictive than stochastic dominance, but includes independence (Theorem 17). In the independent case, when v is a uniform draw on the interval $[0, 1]$, the two dimensional the virtual value function that satisfies the amortization and single-dimensional-projection constraints is pinned down as $2v - 1$ for the item today and $\delta(2v - 1)$ for

the item tomorrow. For $\delta \in [0, 1]$ as defined, these virtual values are consistent with selling the item immediately (or not at all). The existence of this virtual value function proves that uniform pricing is revenue optimal among all mechanisms.

1.1 Related Work

The starting point of work in multi-dimensional optimal mechanism design is the observation that an agent’s utility must be a convex function of his private type (e.g., Rochet, 1985, cf. the envelope theorem). The second step is in writing revenue as the difference between the surplus of the mechanism and the agent’s utility (e.g., McAfee and McMillan, 1988; Armstrong, 1996). The surplus can be expressed in terms of the gradient of the utility. The third step is in rewriting the objective in terms of either the utility (e.g., McAfee and McMillan, 1988; Manelli and Vincent, 2006; Hart and Nisan, 2012; Daskalakis et al., 2013; Wang and Tang, 2014; Giannakopoulos and Koutsoupias, 2014) or in terms of the gradient of the utility (e.g., Armstrong, 1996; Alaei et al., 2013; and this paper). This manipulation follows from an integration by parts. The first category of papers (rewriting objective in terms of utility) performs the integration by parts independently in each dimension, and the second category (rewriting objective in terms of gradient of utility, except for ours) does the integration along rays from the origin (see below). In our approach, in contrast, the integration by parts is performed in general and is dependent on the distribution and the form of the mechanism we wish to show is optimal.

Closest to our work are Armstrong (1996) and Alaei et al. (2013) which use integration by parts along paths that connect types with straight lines to the zero type (which has value zero for any outcome) to define virtual values. For the first work, Armstrong (1996) finds properties on valuation functions (beyond linear ones considered in our paper) and distributions, that when jointly satisfied, imply that the pointwise optimization of virtual surplus results in an incentive compatible mechanism. Armstrong gives some examples of mechanisms that result from this approach but does not generally interpret the form of the resulting mechanisms. Armstrong suggests generalizing his approach from rays from the origin to other kinds of paths; our approach, in contrast, proves the existence of appropriate paths over which to integrate without requiring the form of the path to be specified in advance. When Armstrong’s condition on the distribution is satisfied (which we refer to as independence in max-ratio coordinates), our solution is also equivalent to an integration along rays. For the second work, Alaei et al. (2013) show that optimal multi-agent mechanism design can be reduced to optimal single-agent mechanism design by the construction of a single-dimensional virtual value (that satisfies similar properties to ours) when the single-agent mechanism design problems satisfy a *revenue linearity* property. They prove that a unit-demand agent with values for items that are independently, identically, and uniformly distributed on the $[0, 1]$ interval is revenue linear; our results generalize this one. Moreover, our multi-dimensional virtual value construction constitutes a proof of revenue linearity; therefore, all of our optimal single-agent mechanisms automatically generalize to give optimal multi-agent mechanisms in the *service constrained environments* of Alaei et al. (2013).

There has been work looking at properties of single-agent mechanism design problems that are sufficient for optimal mechanisms to make only limited use of randomization. For context, the optimal single-item mechanism is always deterministic (e.g., Myerson, 1981; Riley and Zeckhauser, 1983), while the optimal multi-item mechanism is sometimes randomized (e.g., Thanassoulis, 2004; Pycia, 2006). For agents with additive preferences across multiple items, McAfee and McMillan (1988), Manelli and Vincent (2006), and Giannakopoulos and Koutsoupias (2014) find sufficient

conditions under which deterministic mechanisms, i.e., bundle pricings, are optimal. A recurring condition in these results is a form of hazard condition, $(m+1)f(\mathbf{t}) + \mathbf{t} \cdot \nabla f(\mathbf{t}) \geq 0$ (where m is the dimension of type space). This condition is orthogonal to our affiliation condition. Projected to a single dimension, this condition becomes regularity of the distribution. Our affiliation condition is automatically satisfied in a single dimension. Pavlov (2011) considers more general preferences and a more general condition; for unit-demand preferences, this condition implies that in the optimal mechanism an agent deterministically receives an item or not, though the item received may be randomized. Our approach is different from the works on multi-dimensional preferences in that it uses the properties of the given mechanism to pin down multi-dimensional virtual values using integration by parts.

A number of papers consider the question of finding closed forms for the optimal mechanism for an agent with additive preferences and independent values across the items. One such closed form is grand-bundle pricing. Our work for additive preferences contrasts in that we prove optimality of grand-bundle pricing and a particular family of correlated distributions. For the two item case, Hart and Nisan (2012) give sufficient conditions for the optimality of grand-bundle pricing; these conditions are further generalized by Wang and Tang (2014). Their results are not directly comparable to ours as our results apply to correlated distributions. Daskalakis et al. (2014) and Giannakopoulos and Koutsoupias (2014) give frameworks, similar to ours, for proving optimality of multi-dimensional mechanisms. Daskalakis et al. (2014) establish a strong duality theorem between the optimal mechanism design problem with additive preferences and an optimal transportation problem between measures. Using this duality they show that every optimal mechanism has a certificate of optimality in the form of transformation maps between measures. They use this result to show that when values for items are independently and uniformly distributed on $[c, c+1]$ for sufficiently large c , the grand bundling mechanism is optimal, extending a result of Pavlov (2011) for two items. Giannakopoulos and Koutsoupias (2014) give a closed form for the optimal mechanism when values are i.i.d. from the uniform distribution (with up to six items).

2 Preliminaries

2.1 The Setting

An agent is specified by a bounded set of possible types T normalized to be $T = [0, 1]^m$, where each $\mathbf{t} = (t_1, \dots, t_m) \in T$ is an m -dimensional vector of values for m items.⁵ The type of the agent is drawn from a known distribution with density f . For the special case that the type space is single-dimensional (i.e., $m = 1$), the cumulative distribution function of the type is denoted by F . We do not require that the values for items be drawn independently. The allocation $\mathbf{x} \in [0, 1]^m$ is a vector of probabilities in which x_i is the probability of receiving i . We consider unit-demand and additive agents. For unit-demand agents the allocation $\mathbf{x} \in [0, 1]^m$ must satisfy $\sum_i x_i \leq 1$; for additive agents $\mathbf{x}$ must satisfy $x_i \leq 1$ for all i . The utility of the agent with type $\mathbf{t}$ for allocation $\mathbf{x} \in [0, 1]^m$ and payment $p \in \mathbb{R}$ is $\mathbf{t} \cdot \mathbf{x} - p$.

A single-agent mechanism is a pair of functions, the allocation function $\mathbf{x} : T \rightarrow [0, 1]^m$ and the payment function $p : T \rightarrow \mathbb{R}$. A mechanism is *individually rational* if the utility of every type of the agent is at least zero,

⁵Throughout the paper we maintain the convention of denoting a vector $\mathbf{v}$ by a bold symbol and each of its components v_i by a non-bold symbol.

$$\mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t}) \geq 0, \quad \forall \mathbf{t} \in T.$$

A mechanism is *incentive compatible* if no type of the agent increases his utility by misreporting,

$$\mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t}) \geq \mathbf{t} \cdot \mathbf{x}(\hat{\mathbf{t}}) - p(\hat{\mathbf{t}}), \quad \forall \mathbf{t}, \hat{\mathbf{t}} \in T.$$

2.2 Multivariable Calculus Notation

For a function $h : \mathbb{R}^k \rightarrow \mathbb{R}$, we use $\partial_j h : \mathbb{R}^k \rightarrow \mathbb{R}$ to denote the *partial derivative* of function h with respect to its j 'th variable. The *gradient* of h is a vector field, denoted by $\nabla h : \mathbb{R}^k \rightarrow \mathbb{R}^k$, defined to be $\nabla h = (\partial_1 h, \dots, \partial_k h)$. The *divergence* of a vector field $\boldsymbol{\alpha} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is denoted by $\nabla \cdot \boldsymbol{\alpha} : \mathbb{R}^k \rightarrow \mathbb{R}$ and is defined to be

$$\nabla \cdot \boldsymbol{\alpha} = \partial_1 \alpha_1 + \dots + \partial_k \alpha_k.$$

We denote the integral of function $h : \mathbb{R}^k \rightarrow \mathbb{R}$ over a subset T of $\mathbb{R}^k$ as

$$\int_{\mathbf{t} \in T} h(\mathbf{t}) \, d\mathbf{t}.$$

Let ∂T be the boundary of set T and $\boldsymbol{\eta}(\mathbf{t})$ be the outward-pointing unit normal vector of T at point $\mathbf{t}$ on ∂T . The multi-variable *integration by parts* for functions $h : \mathbb{R}^k \rightarrow \mathbb{R}$ and $\boldsymbol{\alpha} : \mathbb{R}^k \rightarrow \mathbb{R}^k$ is as follows

$$\int_{\mathbf{t} \in T} (\nabla h \cdot \boldsymbol{\alpha})(\mathbf{t}) \, d\mathbf{t} = \int_{\mathbf{t} \in \partial T} h(\mathbf{t})(\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\mathbf{t}) \, d\mathbf{t} - \int_{\mathbf{t} \in T} h(\mathbf{t})(\nabla \cdot \boldsymbol{\alpha})(\mathbf{t}) \, d\mathbf{t}. \quad (1)$$

Setting h to be the constant function equal to 1 everywhere gives us the *divergence theorem*

$$\int_{\mathbf{t} \in T} (\nabla \cdot \boldsymbol{\alpha})(\mathbf{t}) \, d\mathbf{t} = \int_{\mathbf{t} \in \partial T} (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\mathbf{t}) \, d\mathbf{t}.$$

When the dimension $k = 1$, the integration by parts has the familiar form of

$$\int_{x=a}^b h'(x)g(x) \, dx = h(x)g(x) \Big|_{x=a}^b - \int_{x=a}^b h(x)g'(x) \, dx,$$

and the divergence theorem is the *fundamental theorem of calculus*

$$\int_{x=a}^b h'(x) \, dx = h(b) - h(a).$$

2.3 Problem Formulation

A single agent mechanism $(\mathbf{x}, p)$ defines a utility function $u(\mathbf{t}) = \mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - p(\mathbf{t})$. The following lemma connects the utility function of an IC mechanism with its allocation function.

Lemma 1 (Rochet, 1985). *Function u is the utility function of an agent in an incentive compatible mechanism if and only if u is convex, and in that case, the agent's allocation is $\mathbf{x}(\mathbf{t}) = \nabla u(\mathbf{t})$. Convexity of u with $\mathbf{x}(\mathbf{t}) = \nabla u(\mathbf{t})$ states that for all $\mathbf{t}, \mathbf{t}'$, $(\mathbf{x}(\mathbf{t}) - \mathbf{x}(\mathbf{t}')) \cdot (\mathbf{t} - \mathbf{t}') \geq 0$.*

Notice that the payment function can be defined using the utility function and the allocation function as $p(\mathbf{t}) = \mathbf{t} \cdot \mathbf{x}(\mathbf{t}) - u(\mathbf{t})$. Applying the above lemma, we can write payment to be $p(\mathbf{t}) = \mathbf{t} \cdot \nabla u(\mathbf{t}) - u(\mathbf{t})$. The revenue maximization problem can then be written as the following mathematical program, which is the starting point for the analysis of this paper.

$$\begin{aligned} \max_{\mathbf{x}, u} \quad & \int_{\mathbf{t}} [\mathbf{t} \cdot \nabla u(\mathbf{t}) - u(\mathbf{t})] f(\mathbf{t}) \, d\mathbf{t} \\ & u \text{ is convex,} \\ & \forall \mathbf{t}, \nabla u(\mathbf{t}) \text{ is feasible allocation.} \end{aligned} \tag{2}$$

Notice that when the dimension of the type space is $m = 1$, the above program is equivalent to the following familiar form from Myerson (1981) (we use v for the value, or type, of a single-dimensional agent),

$$\begin{aligned} \max_x \quad & \int_v \left[vx(v) - \int_{z \leq v} x(z) \, dz \right] f(v) \, dv \\ & x \text{ is monotone non-decreasing,} \\ & \forall v, x(v) \leq 1. \end{aligned}$$

In Section 5, we extend the above formulation and our results to multi-agent settings.

3 Amortization of Revenue

This section formalizes our codification of multi-dimensional virtual values for incentive compatible mechanism design and describes the working pieces of our framework. The main construct is the definition of multi-dimensional virtual value functions and the accompanying proposition, below.

Definition 1. A vector field $\bar{\phi} : [0, 1]^m \rightarrow \mathbb{R}^m$ that maps an an m -dimensional type to an m -dimensional vector is

- *incentive compatible* if the *virtual surplus maximizer* given by selecting the outcome $\mathbf{x}$ for type $\mathbf{t}$ that optimizes virtual surplus $\mathbf{x} \cdot \bar{\phi}(\mathbf{t})$ is incentive compatible;
- a *weak amortization of revenue* if, in expectation over types drawn from the distribution, the virtual surplus of any incentive compatible mechanism upper bounds is revenue, i.e., $\mathbf{E}[\bar{\phi}(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t})] \geq \mathbf{E}[p(\mathbf{t})]$, and with equality for the virtual surplus maximizer;
- a *strong amortization of revenue* if the inequality (of weak amortization) holds with equality for all incentive compatible mechanisms; and
- a *virtual value function* if it is incentive compatible and a weak or strong amortization.

Proposition 2. *In any environment for which a virtual value function exists, the virtual surplus maximizer is incentive compatible and revenue optimal.*

Proof. The expected revenue of the virtual surplus maximizer is equal to its expected virtual surplus (by weak amortization). This expected virtual surplus is at least the virtual surplus of any alternate mechanism (by definition of virtual surplus maximization). The expected virtual surplus is an upper bound on the expected revenue of the alternative mechanism (by weak amortization). Thus, the expected revenue of the virtual surplus maximizer is at least that of the alternative mechanism. Incentive compatibility follows directly from the definition of a virtual value function. $\square$

For a single-dimensional agent (i.e., $m = 1$), Myerson (1981) showed that the function $v - \frac{1-F(v)}{f(v)}$ is a strong amortization, when it is monotone it is incentive compatible, when it is non-monotone an ironing procedure can be applied to obtain from it a weak amortization function that is monotone and thus incentive compatible. Our approach will analogously enable the derivation of multi-dimensional virtual value functions (i.e., satisfying incentive compatibility and weak amortization) via the construction of a strong amortization function that is not necessarily incentive compatible.

We now give sufficient conditions for a vector field to be a strong amortization of revenue. At a high level, we derive first a strong amortization of utility and then, using the fact that revenue is value minus utility, derive a strong amortization of revenue. These strong amortizations will be building blocks for the derivation of virtual values for unit-demand and additive agents in the subsequent sections. The following lemma follows from integration by parts as per equation (1), the definition of strong amortization of utility (Definition 1, generalized to utility), and the fact that the gradient of utility is the allocation rule of the mechanism (Lemma 3).

Lemma 3. *For type space T and distribution f , vector field α/f is a strong amortization of utility, i.e., $\mathbf{E}[u(\mathbf{t})] = \mathbf{E}[\alpha(\mathbf{t})/f(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t})]$ for all incentive compatible allocation rules $\mathbf{x}$, if it satisfies*

- divergence density equality, i.e., that $\nabla \cdot \alpha(\mathbf{t}) = -f(\mathbf{t})$ at any point $\mathbf{t} \in T$, and
- boundary orthogonality, i.e., that $\alpha(\mathbf{t}) \cdot \eta(\mathbf{t}) = 0$ for all $\mathbf{t} \neq \mathbf{0}$ on the boundary of type space ∂T with normal vector $\eta(\mathbf{t})$.

Proof. Write the expectation of $\alpha(\mathbf{t})/f(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t})$ as the integral $\int_{\mathbf{t} \in T} \alpha(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t}) d\mathbf{t}$. From Lemma 3, substitute ∇u for allocation $\mathbf{x}$ and apply integration by parts.

$$\int_{\mathbf{t} \in T} \alpha(\mathbf{t}) \cdot \nabla u(\mathbf{t}) d\mathbf{t} = \int_{\mathbf{t} \in \partial T} u(\mathbf{t})(\alpha \cdot \eta)(\mathbf{t}) d\mathbf{t} - \int_{\mathbf{t} \in T} u(\mathbf{t})(\nabla \cdot \alpha(\mathbf{t})) d\mathbf{t} \quad (3)$$

$$= u(\mathbf{0}) + \int_{\mathbf{t} \in T} u(\mathbf{t}) f(\mathbf{t}) d\mathbf{t} \quad (4)$$

The second equality is derived from the first equality by employing the assumptions of the lemma on vector field α as follows.

- By divergence density equality, the second term simplifies by substituting $\nabla \cdot \alpha(\mathbf{t}) = -f(\mathbf{t})$.
- Recall that the divergence theorem is equivalent to setting $u(\mathbf{t}) = 1$ in the formula (3), this gives

$$\int_{\mathbf{t} \in \partial T} (\alpha \cdot \eta)(\mathbf{t}) d\mathbf{t} = \int_{\mathbf{t} \in T} f(\mathbf{t}) d\mathbf{t} = 1, \quad (5)$$

the total probability of any type. Boundary orthogonality implies that the integrand in the boundary integral of equation (5) is identically zero everywhere except $\mathbf{t} = \mathbf{0}$. To integrate to one on the boundary, the function must be the Dirac delta function at $\mathbf{t} = \mathbf{0}$; thus, the integral of the first term in equation (3) is $u(\mathbf{0})$.

Without loss of generality for revenue optimal mechanisms $u(\mathbf{0}) = 0$. We can interpret the left- and right-hand sides of equation (4) as expectations, which gives $\mathbf{E}[\alpha(\mathbf{t})/f(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t})] = \mathbf{E}[u(\mathbf{t})]$, the definition of strong amortization of utility for α/f . $\square$

For a single-dimensional agent with value v in type space $T = [0, 1]$, the only function that satisfies the conditions of Lemma 5 and gives a strong amortization of utility is $\frac{\alpha(v)}{f(v)} = \frac{1-F(v)}{f(v)}$. For this formula, notice that the divergence of $\alpha(v) = 1 - F(v)$ is simply its derivative $-f(v)$. The boundary $\partial T \setminus \{0\}$ is the point $v = 1$, the upper bound of the distribution, and thus trivially satisfies orthogonality as $\alpha(1) = 0$. In classical auction theory the amortization of utility $\frac{1-F(v)}{f(v)}$ is often referred to as the agent's *information rent*.

The following lemma is immediate from the fact that revenue is the agent's surplus minus the agent's utility. For a single-dimensional agent it implies that $\phi(v) = v - \frac{1-F(v)}{f(v)}$ is the strong amortization of revenue.

Lemma 4. *For type space T and distribution f , vector field ϕ is a strong amortization of revenue, i.e., $\mathbf{E}[p(\mathbf{t})] = \mathbf{E}[\phi(\mathbf{t}) \cdot \mathbf{x}(\mathbf{t})]$ for all incentive compatible allocation rules $\mathbf{x}$, if and only if $\phi(\mathbf{t}) = \mathbf{t} - \alpha(\mathbf{t})/f(\mathbf{t})$ for all $\mathbf{t}$ but a measure zero subset of T and α/f is a strong amortization of utility.*

Unlike the case of a single-dimensional agent, for multi-dimensional agents there are many strong amortizations of utility and, consequently, many strong amortizations of revenue. As an example, suppose we wish to show the optimality of a restricted form of mechanism via a strongly amortized virtual value function for an $m = 2$ dimensional agent. This virtual value function has two degrees of freedom. We can pin down one degree of freedom by equating virtual surplus to expected revenue for mechanisms with this restricted form. The divergence density equality for strong amortizations that Lemma 6 inherits from Lemma 5 gives a differential equation that then pins down the other degree of freedom. It remains to find sufficient conditions on the distribution under which virtual surplus maximization identically gives mechanisms of the restricted form.

The approach above can be generalized to show optimality of mechanisms via a weakly amortized virtual value function. For example, such a generalization can be used to give proofs of optimality under more permissive distributional assumptions. To substitute weak amortization for strong amortization we need a way to relate the differential equations (from divergence density equality) that govern strong amortizations to any given weak amortization. Such a relationship follows directly from the definitions of both weak and strong amortization in terms of the expected revenue of any incentive compatible mechanism (Definition 1) and is summarized below as Lemma 7.

Lemma 5. *For type space T and distribution f , vector field $\bar{\phi}$ is a weak amortization of revenue if and only if there exists a strong amortization of revenue ϕ such that $\mathbf{E}[\bar{\phi}(\mathbf{t}) \mathbf{x}(\mathbf{t})] \geq \mathbf{E}[\phi(\mathbf{t}) \mathbf{x}(\mathbf{t})]$, for all incentive compatible mechanisms $\mathbf{x}$, with equality for $\mathbf{x}$ that pointwise maximizes $\bar{\phi}$.*

For a single-dimensional agent when the strong amortization $\phi(v) = v - \frac{1-F(v)}{f(v)}$ is not monotone, it is not a virtual value function. In this case, virtual value function is a weak amortization which is derived using ironing.

In the above framework for understanding optimal mechanisms as virtual value optimizations, the class of mechanisms optimized over has not been constrained beyond incentive compatibility. In particular the quantifications over all incentive compatible mechanisms, both in the definition of weak and strong amortizations and in Lemma 7, can be restricted to all incentive compatible mechanisms within a given subclass of mechanisms. Such a restriction can be useful even when

looking for the optimal mechanism in the of the full class if it is known to always lie within the subclass. As one example, for symmetric environments (symmetric feasibility and identically distributed agents) there is always an optimal mechanism that is symmetric. Thus, to argue that a vector field is a weak amortization by Lemma 7, it suffices to show the inequality for symmetric incentive compatible mechanisms. The optimal symmetric mechanism is an optimal mechanism among all, possibly asymmetric, mechanisms.

Proposition 6. *If the optimal mechanism is in a certain class of mechanisms, then the weak amortization inequality is required to hold only for all mechanisms in that class.*

4 Optimality of Uniform Pricing for Unit Demand Preferences

In this section we study the existence of virtual values to prove optimality of uniform pricing for a single unit-demand agent. To simplify the exposition we focus on the case of two items and on the type space where item one is the favorite item, i.e., the lower right half of the unit square, $T = \{\mathbf{t} = (t_1, t_2) : 0 \leq t_2 \leq t_1 \leq 1\}$. Our conclusions extend easily to the $[0, 1]^2$ type space with symmetric distributions; other extensions are given at the end of the section. The general case of $m \geq 2$ items is considered in Appendix A.

The *single-dimensional projection for the favorite item* is given by distribution and density function for the agent's favorite item, $F_{\max}(v)$ and $f_{\max}(v)$. The distribution function $F_{\max}(v)$ is the integral of f over $\mathbf{t}$ with $t_1 \geq v$. The density function $f_{\max}(v)$ is the integral of f of $\mathbf{t}$ with $t_1 = v$, i.e., $f_{\max}(v) = \int_0^v f(v, z) dz$. As described in Section 3, the unique strong amortization of revenue for a single-dimensional agent (and thus for the single-dimensional projection) is $\phi_{\max}(v) = v - \frac{1 - F_{\max}(v)}{f_{\max}(v)}$. The strong amortization of utility $\alpha_{\max}/f_{\max}$ requires $\alpha_{\max}(v) = 1 - F_{\max}(v)$.

Definition 2. The two-dimensional extension ϕ of the favorite-item projection $\phi_{\max}$, satisfying $\phi_{\max}(\mathbf{t}) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)}$, is constructed as follows:

- (a) Set $\phi_1(\mathbf{t}) = \phi_{\max}(t_1)$ for all $\mathbf{t} \in T$.
- (b) Let $\alpha_1(\mathbf{t}) = (t_1 - \phi_1(\mathbf{t})) f(\mathbf{t}) = \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} f(\mathbf{t})$.
- (c) Let $\alpha_2(\mathbf{t}) = - \int_{y=0}^{t_2} (f(t_1, y) + \partial_1 \alpha_1(t_1, y)) dy$.
- (d) Set $\phi_2(\mathbf{t}) = t_1 - \alpha_2(\mathbf{t})/f(\mathbf{t})$.

In the remainder of this section we will show that to optimize revenue minus a fixed non-negative cost for selling either item, this two-item extension of the favorite-item projection is a strong amortization of revenue that proves the optimality of uniform pricing.⁶

An informal justification of the steps of the construction is as follows:

- (a) First, for fixed t_1 and as a function of t_2 , $\phi_1(\mathbf{t})$ must be constant (i.e., on a vertical line in T); otherwise, there is a cost c for which virtual surplus maximization with respect to ϕ serves a type $\mathbf{t}$ but not a type $\mathbf{t}'$ with $t'_1 = t_1$, which is not a uniform pricing. Second, the revenue of any mechanism that only ever sells the favorite item or nothing has revenue given by the favorite-item projection and must satisfy $\phi_1(\mathbf{t}) = \phi_{\max}(t_1)$ (given the first point).

⁶The extra constraint imposed by a non-negative cost of service will enable this method to be extended to multi-agent settings, see Section 5. This strong amortization is unique on the portion of type space for which $\phi_1(\mathbf{t}) > 0$.

- (b) We obtain α_1 from ϕ_1 by Lemma 6. Orthogonality of the right boundary ($t_1 = 1$) requires that $\alpha \cdot (1, 0) = 0$, and therefore $\alpha_1(1, t_2) = 0$. Boundary orthogonality of the favorite-item projection at $t_1 = 1$ implies $\alpha_{\max}(1) = 0$, and using the definition of α_1 , $\alpha_1(1, t_2) = \frac{\alpha_{\max}(1)}{f_{\max}(1)} f(1, t_2) = 0$.
- (c) The derivatives of α_1 (with respect to t_1) and α_2 (with respect to t_2) are related by the divergence density equality; integrating and employing boundary orthogonality on the bottom boundary ($t_2 = 0$) of the type space, which requires that $\alpha_2(t_2, 0) = 0$, gives the formula; these constraints are required by Lemma 5.
- (d) We obtain ϕ_2 from α_2 by Lemma 6.

For ϕ to prove optimality of uniform pricing, we need virtual surplus optimization with respect to ϕ to yield a uniform pricing. This requirement is simply $\phi_1(\mathbf{t}) \geq \phi_2(\mathbf{t})$ for any type $\mathbf{t} \in T$ for which either $\phi_1(\mathbf{t})$ or $\phi_2(\mathbf{t})$ is positive. A little algebra shows that this condition is implied by the angle of $\alpha(\mathbf{t})$ being at most the angle of $\mathbf{t}$ (with respect to the horizontal t_1 axis; see Lemma 10, below). In relation to the prior work of Armstrong (1996), the direction of α corresponds to the paths on which incentive compatibility constraints are considered. Importantly, the approach we are taking does not fix the direction, it allows any direction that satisfies the above constraint on angles. The condition on angles is equivalent to the dot product between α and the upward orthogonal vector to the ray crossing the zero type and $\mathbf{t}$ at $\mathbf{t}$ being non-positive. The following lemma is proved by the divergence theorem, and provides a closed form for the dot product between α and the upward orthogonal vector to $\mathbf{t}$. We will later identify conditions over the distribution of types such that the angle of α satisfies the required condition by analyzing this closed form.

Lemma 7. *Vector field α/f in the definition of the two-dimensional extension of the favorite-item projection is a strong amortization of utility and satisfies*

$$\theta \alpha_1(t_1, t_1\theta) - \alpha_2(t_1, t_1\theta) = (1 - F_{\max}(t_1)) \frac{d}{dt_1} \left[\frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)} \right]$$

for all $t_1, \theta \in [0, 1]$ (and thus $(t_1, t_1\theta) \in T$).

Proof. The informal justifications of Steps (b) and (c) show that α satisfies the divergence density equality and bottom and right boundary orthogonality. This proof starts with these assumptions and derives the identity of the lemma. Notice that in the identity, for $\theta = 1$ the numerator and the denominator in the derivative of the identity are equal for all t_1 , the right-hand side is zero, and therefore $\alpha_1(t_1, t_1) = \alpha_2(t_1, t_1)$, and boundary orthogonality holds for the diagonal boundary. Thus, α/f is a strong amortization of utility.

The strategy for the proof of the identity is as follows. We fix t_1 and θ and apply the divergence theorem to α on the trapezoidal subspace of type space defined by types $\mathbf{t}'$ with $t'_1 \geq t_1$, $t'_2/t'_1 \leq \theta$, $t'_2 \geq 0$, and $t'_1 \leq 1$ (Figure 2). The divergence theorem equates the integral of the orthogonal magnitude of vector field α on the boundary of the subspace to the integral of its divergence within the subspace. As the upper boundary of this trapezoidal subspace has slope t_2/t_1 , one term in this equality is the integral of $\alpha(\mathbf{t}')$ with the upward orthogonal vector to $\mathbf{t}'$. Differentiating this integral with respect to t_1 and evaluating at $\mathbf{t} = (t_1, t_1\theta)$ gives the desired quantity.

Applying the divergence theorem to α on the trapezoid and expressing the top boundary as the interior divergence minus the other three boundaries gives:

$$\int_{\mathbf{t}' \in \text{TOP}(t_1, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}' = \int_{\mathbf{t}' \in \text{INTERIOR}(t_1, \theta)} \nabla \cdot \alpha(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \{\text{RIGHT}, \text{BOTTOM}, \text{LEFT}\}(t_1, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}'.$$

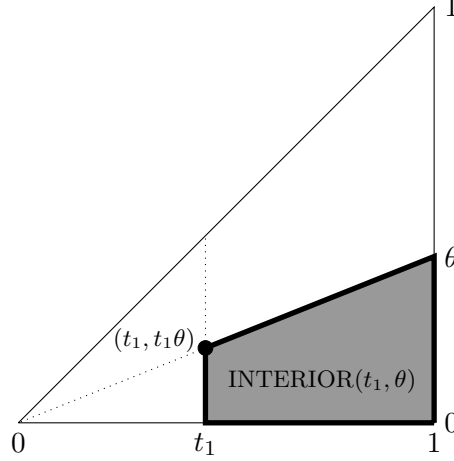


Figure 1: The trapezoidal set parameterized by t_1 and θ , and the four curves that define its boundary, $\{\text{TOP, RIGHT, BOTTOM, LEFT}\}_{(t_1, \theta)}$

Since α/f is a strong amortization of utility, the divergence density equality and boundary orthogonality of right and bottom boundaries imply that the integral over the interior simplifies and the integrals over the right and bottom boundaries are zero, respectively. We have,

$$\int_{\mathbf{t}' \in \text{TOP}(t_1, \theta)} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' = - \int_{\mathbf{t}' \in \text{INTERIOR}(t_1, \theta)} f(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \text{LEFT}(t_1, \theta)} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}'.$$

For the trapezoid at $\mathbf{t}$ these integrals are,

$$\begin{aligned} \int_{t'_1=t_1}^1 \left(-\theta \alpha_1(t'_1, t'_1\theta) + \alpha_2(t'_1, t'_1\theta) \right) dt'_1 \\ = - \int_{t'_1=t_1}^1 \int_{t'_2=0}^{t'_1\theta} f(\mathbf{t}') dt'_2 dt'_1 + \int_{t'_2=0}^{t_1\theta} \alpha_1(t_1, t'_2) dt'_2. \end{aligned}$$

Differentiating with respect to t_1 gives,

$$\theta \alpha_1(t_1, t_1\theta) - \alpha_2(t_1, t_1\theta) = \int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2 + \frac{d}{dt_1} \int_{t'_2=0}^{t_1\theta} \alpha_1(t_1, t'_2) dt'_2.$$

On the right-hand side, multiply first term by $\frac{f_{\max}(t_1)}{f_{\max}(t_1)} = 1$ and plug in the strong amortization of utility for the two-dimensional extension as $\alpha_1(\mathbf{t}) = \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)} f(\mathbf{t})$ to the second term. Notice that the integral of the second term is only on t'_2 therefore we can bring the terms related to t_1 outside the integral. These two terms then simplify by the product rule for differentiation to give the identity of the lemma.

$$\begin{aligned} \theta \alpha_1(t_1, t_1\theta) - \alpha_2(t_1, t_1\theta) &= f_{\max}(t_1) \frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)} + \frac{d}{dt_1} \left[(1 - F_{\max}(t_1)) \frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)} \right] \\ &= (1 - F_{\max}(t_1)) \frac{d}{dt_1} \left[\frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)} \right]. \end{aligned}$$

□

The following lemma gives sufficient conditions for uniform pricing to be the maximizer of virtual surplus given any cost c . These conditions imply that whenever $\phi_1(\mathbf{t}) \geq c$ then $\phi_1(\mathbf{t}) \geq \phi_2(\mathbf{t})$, and that $\phi_1(\mathbf{t}) \geq c$ if and only if t_1 is greater than a certain threshold (implied by monotonicity of $\phi_1(\mathbf{t}) \geq c$).

Lemma 8. *For ϕ and α defined by the two-dimensional extension of the favorite-item projection, if $\frac{t_2}{t_1} \alpha_1(t_1, t_2) - \alpha_2(t_1, t_2)$ is non-positive and $\phi_1(\mathbf{t})$ is monotone non-decreasing in $\mathbf{t}$, then virtual surplus maximization with respect to ϕ and any non-negative service cost c gives a uniform pricing.*

Proof. From the assumption $\frac{t_2}{t_1} \alpha_1(t_1, t_2) - \alpha_2(t_1, t_2) \leq 0$ and Definition 2 we have

$$\frac{t_2}{t_1} \phi_1(\mathbf{t}) = \frac{t_2}{t_1} \left(t_1 - \frac{\alpha_1(\mathbf{t})}{f(\mathbf{t})} \right) = \frac{t_2}{t_1} \left(t_1 - \frac{\alpha_1(\mathbf{t})}{f(\mathbf{t})} \right) \geq t_2 - \frac{\alpha_2(\mathbf{t})}{f(\mathbf{t})} = \phi_2(\mathbf{t}).$$

Thus, for $\mathbf{t}$ with $\phi_1(\mathbf{t}) \geq c$, $\phi_1(\mathbf{t}) \geq \phi_2(\mathbf{t})$ and virtual surplus maximization serves the agent item one. Since $\phi_1(\mathbf{t})$ is a function only of t_1 (Definition 2), its monotonicity implies that there is a smallest t_1 such that all greater types are served. Also, if $\phi_1(\mathbf{t}) \leq c$, again the above calculation implies that $\phi_2(\mathbf{t}) \leq c$ and therefore the type is not served. This outcome is a uniform pricing. □

We are now ready to state the main theorem of this section. In Section 4.1 we will give an interpretation of the main technical condition as a supermodularity condition on the density function.

Theorem 9. *For a unit demand agent with $m = 2$ items and any service cost $c \geq 0$, uniform pricing is revenue optimal for any distribution for which the favorite-item projection has monotone non-decreasing strong amortization $\phi_{\max}(t_1) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)}$ and $\frac{\int_0^{t_1 \theta} f(t_1, y) dy}{f_{\max}(t_1)}$ is a monotone non-increasing function of t_1 for all θ .*

We relax the monotonicity condition of the strong amortization of the favorite-item projection $\phi_{\max}$ in Section 4.2 by constructing a weak amortization $\bar{\phi}$ from the strong amortization ϕ , above, in a manner paralleling Myerson's ironing.

In Appendix A we give a natural generalization of the condition and extend the above results to $m > 2$ items.

4.1 Max-ratio Representation of Distributions

Theorem 11 in the preceding section require monotonicity of $\frac{\int_0^{t_1 \theta} f(t_1, y) dy}{f_{\max}(t_1)}$. In this section we further investigate this property and give simple sufficient conditions for it. Given t_1 and θ , the probability that a random draw $\mathbf{t}'$ conditioned on $t'_1 = t_1$ satisfies $\theta(\mathbf{t}') = t'_2/t'_1 \leq \theta$ is

$$\Pr_{\mathbf{t}'} \left[\theta(\mathbf{t}') \leq \theta | t'_1 = t_1 \right] = \frac{\int_{t'_2=0}^{t_1 \theta} f(t_1, t'_2) dt'_2}{\int_{t'_2=0}^{t_1} f(t_1, t'_2) dt'_2}.$$

The condition of Theorem 11 required that the distribution of θ conditioned on t_1 is first order stochastically dominated by the distribution of θ conditioned on $t'_1 \geq t_1$, or equivalently,

$$\frac{d}{dt_1} \mathbf{Pr}_{\mathbf{t}'} \left[\theta(\mathbf{t}') \leq \theta|_{t'_1 = t_1} \right] \leq 0.$$

We now define a simple sufficient condition implying the dominance requirement above. Define the max-ratio representation f^{MR} of a density function f is to be

$$f^{MR}(t_1, \theta) = f(t_1, \theta t_1), \quad \forall t_1, \theta \in [0, 1].$$

Equivalantly,

$$f(t_1, t_2) = f^{MR}(t_1, t_2/t_1) \quad \forall t_1, t_2, 0 \leq t_2 \leq t_1 \leq 1.$$

We call a distribution *MR-log-supermodular*, if its max-ratio representation is log-supermodular,

$$f^{MR}(t_1, \theta) \times f^{MR}(t'_1, \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t'_1, \theta), \quad \forall t_1 \leq t'_1, \theta \leq \theta'.$$

Notice that, for example, *MR-independent* distributions are MR-log-supermodular. A MR-independent distribution f^{MR} is such that $f^{MR}(t_1, \theta) = f_1(t_1) \times f_\theta(\theta)$ (for arbitrary f_1 and f_θ), and is MR-log-supermodular because

$$\begin{aligned} f^{MR}(t_1, \theta) \times f^{MR}(t'_1, \theta') &= f_1(t_1) f_\theta(\theta) f_1(t'_1) f_{\theta'}(\theta') \\ &= f_1(t_1) f_{\theta'}(\theta') f_1(t'_1) f_\theta(\theta) \\ &= f^{MR}(t_1, \theta') \times f^{MR}(t'_1, \theta). \end{aligned}$$

We now prove that $\mathbf{Pr}_{\mathbf{t}'}[\theta(\mathbf{t}') \leq \theta|_{t'_1 = t_1}]$ is monotone non-increasing in t_1 for MR-log-supermodular distributions. This result enables the interpretation of Theorem 11 in terms of MR-log-supermodularity. Recall that the density function of the favorite-item projection is defined from the two-item density function as $f_{\max}(t_1) = \int_0^{t_1} f(t_1, y) dy$.

Lemma 10. *If f is a MR-log-supermodular function, then for any t_1 and θ ,*

$$\frac{d}{dt_1} \mathbf{Pr}_{\mathbf{t}'} \left[\theta(\mathbf{t}') \leq \theta|_{t'_1 = t_1} \right] \leq 0,$$

with equality if distribution is MR-independent.

Proof. We prove that for any θ, t_1 , and t'_1 such that $t_1 < t'_1$,

$$\frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{\int_{t'_2=0}^{t_1} f(t_1, t'_2) dt'_2} \geq \frac{\int_{t'_2=0}^{t'_1\theta} f(t'_1, t'_2) dt'_2}{\int_{t'_2=0}^{t'_1} f(t'_1, t'_2) dt'_2}.$$

The proof first converts the above form into max-ratio coordinates, applies MR-log-supermodularity, and then transforms back to the standard form. Before applying MR-log-supermodularity, we break down the integral set into two sets, and apply MR-log-supermodularity to only one of the integrals. In particular notice that

$$\begin{aligned}
& \int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2 \times \int_{t'_2=0}^{t'_1} f(t'_1, t'_2) dt'_2 \\
&= \int_{\theta'=0}^{\theta} f^{MR}(t_1, \theta') t_1 d\theta' \times \int_{\theta''=0}^1 f^{MR}(t'_1, \theta'') t'_1 d\theta'' \quad (\text{change of variables}) \\
&= \int_{\theta'=0}^{\theta} \int_{\theta''=0}^{\theta} f^{MR}(t_1, \theta') t_1 f^{MR}(t'_1, \theta'') t'_1 d\theta'' d\theta' \\
&\quad + \int_{\theta'=0}^{\theta} \int_{\theta''=\theta}^1 f^{MR}(t_1, \theta') t_1 f^{MR}(t'_1, \theta'') t'_1 d\theta'' d\theta' \quad (\text{separate double integral into two sets}) \\
&\geq \int_{\theta'=0}^{\theta} \int_{\theta''=0}^{\theta} f^{MR}(t_1, \theta'') t_1 f^{MR}(t'_1, \theta') t'_1 d\theta' d\theta'' \quad (\text{rename variables } \theta' \text{ and } \theta'') \\
&\quad + \int_{\theta'=0}^{\theta} \int_{\theta''=\theta}^1 f^{MR}(t_1, \theta'') t_1 f^{MR}(t'_1, \theta') t'_1 d\theta'' d\theta' \quad (\text{apply MR-log-supermodularity}) \\
&= \int_{\theta''=0}^1 f^{MR}(t_1, \theta'') t_1 d\theta'' \int_{\theta'=0}^{\theta} f^{MR}(t'_1, \theta') t'_1 d\theta' \quad (\text{merge integrals}) \\
&= \int_{t'_2=0}^{t_1} f(t_1, t'_2) dt'_2 \times \int_{t'_2=0}^{t'_1\theta} f(t'_1, t'_2) dt'_2.
\end{aligned}$$

□

By combining the above lemma and Theorem 11 we get the following corollary.

Corollary 11. *Uniform pricing is revenue optimal for any service cost c and any max-ratio log-supermodular distribution for which the favorite-item projection has monotone non-decreasing strong amortization $\phi_{\max}(t_1) = t_1 - \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)}$.*

To understand MR-log-supermodular distributions better, we next show properties under which product distributions are MR-log-supermodular. A function g is geometric-geometric (GG) convex if

$$g(z_1^\lambda z_2^{1-\lambda}) \geq g(z_1)^\lambda g(z_2)^{1-\lambda}, \quad \forall \lambda \in [0, 1], z_1, z_2.$$

For example, the function $g(x) = x^k$ is GG-convex (it in fact satisfies the above condition with equality). More generally, for any convex function h and constant c , the function $g(x) = c \cdot e^{h(\log(x))}$ is GG-convex. The proof of the following lemma can be found in Appendix B.

Lemma 12. *The product distribution on t_1, t_2 from a distribution with geometric-geometric convex density is max-ratio log-supermodular.*

4.2 Quantile Space and Weak Amortization

As it happens in single-dimensional settings, the virtual value of the favorite item projection $\phi_{\max}$ may not be monotone. However, we can remove the monotonicity requirement by designing a weak amortization in a similar manner to ironing in a single dimension. In this section we will design a weak amortization $\bar{\phi}$ from the strong amortization ϕ , which satisfies the monotonicity of $\bar{\phi}_1$

without requiring regularity of the distribution of the favorite item projection. We will start by defining a mapping between the type space and a two-dimensional *quantile* space. We will then use Myerson's ironing to pin down the first coordinate $\bar{\phi}_1$ of the weak amortization. The second component $\bar{\phi}_2$ is then defined such that the expected virtual surplus with respect $\bar{\phi}$ upper bounds revenue for all incentive compatible mechanisms. To do this, we invoke integration by parts along curves defined by the quantile mapping, and then use incentive compatibility to identify a direction that the vector $\bar{\phi} - \phi$ may have for $\bar{\phi}$ to be an upper bound on revenue. We use this identity to solve for $\bar{\phi}_2$, and finally identify conditions such that optimization of $\bar{\phi}$ gives uniform pricing.

We first transform the value space to quantile space using following mappings. Recall from Section 4 that $F_{\max}$ and $f_{\max}$ are the distribution and the density functions of the favorite item projection. Define the first quantile mapping

$$q_1(t_1, t_2) = 1 - F_{\max}(t_1)$$

to be the probability that a random draw t'_1 from $F_{\max}$ satisfies $t'_1 \geq t_1$, and the second quantile mapping

$$q_2(t_1, t_2) = 1 - \frac{\int_{t'_2=0}^{t_2} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)}$$

where $f_{\max}(t_1) = \int_0^{t_1} f(t_1, t'_2) dt'_2$ to the probability that a random draw $\mathbf{t}'$ from a distribution with density f , conditioned on $t'_1 = t_1$, satisfies $t'_2 \geq t_2$. The determinant of the Jacobian matrix of the transformation is

$$\begin{vmatrix} \frac{\partial q_1}{\partial t_1} & \frac{\partial q_1}{\partial t_2} \\ \frac{\partial q_2}{\partial t_1} & \frac{\partial q_2}{\partial t_2} \end{vmatrix} = \begin{vmatrix} -f_{\max}(t_1) & 0 \\ \frac{\partial q_2}{\partial t_1} & -\frac{f(t_1, t_2)}{f_{\max}(t_1)} \end{vmatrix} = f(t_1, t_2).$$

As a result, we can express revenue in quantile space as follows

$$\int \int \mathbf{x}(\mathbf{t}) \cdot \phi(\mathbf{t}) f(\mathbf{t}) d\mathbf{t} = \int_{q_1=0}^1 \int_{q_2=0}^1 \mathbf{x}^Q(\mathbf{q}) \cdot \phi^Q(\mathbf{q}) d\mathbf{q},$$

where $\mathbf{x}^Q$ and ϕ^Q are representations of $\mathbf{x}$ and ϕ in quantile space. In particular, $\phi_1^Q(\mathbf{q}) = \phi_{\max}(t_1(q_1))$ might not be monotone in q_1 . In what follows we design a weak amortization $\bar{\phi}^Q$ using ϕ^Q .

Before we proceed, we note that quantile transformations are alternative representations of the probability distribution. In this section it is more convenient to work with this representation. Particularly useful are the *equi-quantile* curves, that is the set of types that mapped to the same second quantile q_2 , $\{\mathbf{t} | q_2(\mathbf{t}) = q_2\}$ (see Figure 3). We can express the technical conditions of Theorem 11, that

$$\frac{\int_{t'_2=0}^{t_2} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)}$$

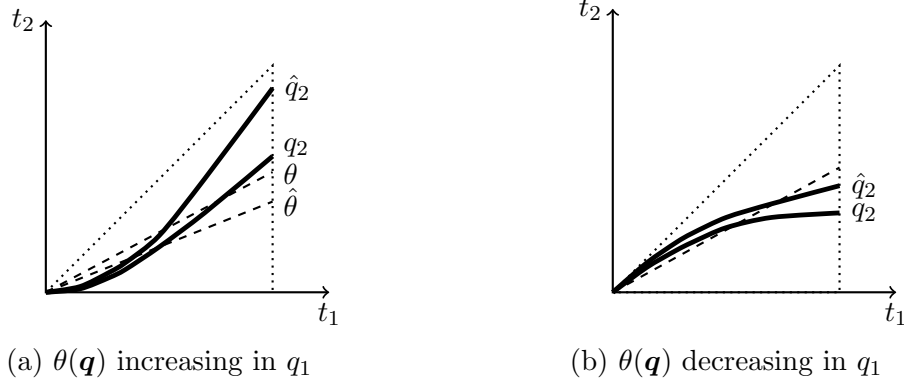


Figure 2: The quantile mapping defines the *equi-quantile* curves, that is, the set of point mapped to the same quantile q_2 . In each figure two curves are drawn mapping to q_2 and $\hat{q}_2$ such that $q_2 \geq \hat{q}_2$ (recall that lower t_2 are mapped to higher q_2). The fact that $q_2(t_1, t_1\theta)$ is monotone increasing in θ implies that as one moves along a ray with parameter θ , equi-quantile curves with higher q_2 will be met, see part (a). Alternatively, this means that as one moves along an equi-quantile curve parameterized by q_2 , rays with higher θ will be met. This is equivalent to saying that $\theta(\mathbf{q})$ is monotone increasing in q_1 . The same argument, with different directions of monotonicity, holds for part (b). As a result, the function $q_2(t_1, t_1\theta)$ is monotone increasing in t_1 if and only if $\theta(\mathbf{q})$ is monotone increasing in q_1 .

is monotone non-increasing function of t_1 for all θ , in terms of the quantile mappings. In particular note that

$$q_2(t_1, t_1\theta) = 1 - \frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)}.$$

As a result, monotonicity of $\frac{\int_{t'_2=0}^{t_1\theta} f(t_1, t'_2) dt'_2}{f_{\max}(t_1)}$ in t_1 is equivalent to monotonicity of $q_2(t_1, t_1\theta)$ in t_1 . Define $\theta(\mathbf{q}) = \frac{t_1(\mathbf{q})}{t_2(\mathbf{q})}$, where functions t_1 and t_2 are the inverses of the quantile transformation. The fact that $\frac{d}{dt_1} q_2(t_1, t_1\theta) \geq 0$ is equivalent to $\frac{d}{dq_1} \theta(\mathbf{q}) \geq 0$ (see Figure 3).

We now derive $\bar{\phi}^Q$ from the properties it must satisfy. In particular, we require $\bar{\phi}_1^Q(\mathbf{q}) = \bar{\phi}_1^Q(q_1)$ to be a monotone non-decreasing function of q_1 , and that $\bar{\phi}_1^Q(\mathbf{q}) \geq \bar{\phi}_2^Q(\mathbf{q})$ whenever either is positive. These properties will imply that a point-wise optimization of $\bar{\phi}^Q$ will result in an incentive compatible allocation of only the favorite item, such that $x_1^Q(\mathbf{q}) = x_1^Q(q_1)$, and $x_2^Q(\mathbf{q}) = 0$ (which is the case for the allocation of uniform pricing). Note that for any such allocation,

$$\int_{q_1=0}^1 \int_{q_2=0}^1 \mathbf{x}^Q(\mathbf{q}) \cdot \phi^Q(\mathbf{q}) d\mathbf{q} = \int_{q_1} x_1^Q(q_1) \phi_1^Q(q_1) dq_1.$$

Similarly, for any such allocation,

$$\int_{q_1=0}^1 \int_{q_2=0}^1 \mathbf{x}^Q(\mathbf{q}) \cdot \bar{\phi}^Q(\mathbf{q}) d\mathbf{q} = \int_{q_1} x_1^Q(q_1) \bar{\phi}_1^Q(q_1) dq_1.$$

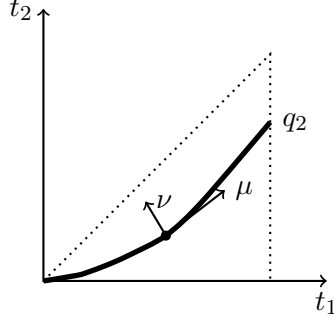


Figure 3: Incentive compatibility implies that $\frac{d}{dq_1} \mathbf{x}(\mathbf{q}) \cdot \mu$ is non-negative if μ is tangent to the equi-quantile curve crossing $\mathbf{q}$.

We can therefore use Myerson's ironing and define $\bar{\phi}_1^Q$ to be the derivative of the convex hull of the integral of ϕ_1^Q . This will imply that $\bar{\phi}^Q$ upper bounds revenue for any allocation that satisfies $x_1^Q(\mathbf{q}) = x_1^Q(q_1)$, and $x_2^Q(\mathbf{q}) = 0$, with equality for the allocation that optimizes $\bar{\phi}^Q$ pointwise.

We will next define $\bar{\phi}_2^Q$ such that $\bar{\phi}^Q$ upper bounds revenue for *all* incentive compatible allocations. That is, we require that for all incentive compatible $\mathbf{x}$,

$$\int \int \mathbf{x}^Q(\mathbf{q}) \cdot (\bar{\phi}^Q - \phi^Q)(\mathbf{q}) d\mathbf{q} \geq 0.$$

Using integration by parts we can write

$$\int \int \mathbf{x}^Q(\mathbf{q}) \cdot (\bar{\phi}^Q - \phi^Q)(\mathbf{q}) d\mathbf{q} = \int_{q_2} \int_{q_1} \frac{d}{dq_1} \mathbf{x}^Q(\mathbf{q}) \cdot \int_{q'_1 \geq q_1} (\bar{\phi}^Q - \phi^Q)(q'_1, q_2) dq'_1 dq_1 dq_2.$$

Incentive compatibility implies that the dot product of any vector and the change in allocation rule in the direction of that vector is non-negative (Lemma 3). In particular this must be true for the tangent vector to equi-quantile curve parameterized by q_2 . Thus incentive compatibility of $\mathbf{x}$ implies that the above expression is positive if the vector that is multiplied by $\frac{d}{dq_1} \mathbf{x}^Q(\mathbf{q})$ is tangent to the equi-quantile curve $(t_1(q'_1, q_2), t_2(q'_1, q_2))$, $0 \leq q'_1 \leq q_1$ at $q'_1 = q_1$ (see Figure 4),

$$\frac{\int_{q'_1 \geq q_1} (\bar{\phi}_2^Q - \phi_2^Q)(q'_1, q_2) dq'_1}{\int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1, q_2) dq'_1} = \frac{\frac{d}{dq_1} t_2(\mathbf{q})}{\frac{d}{dq_1} t_1(\mathbf{q})}.$$

We will set $\bar{\phi}_2^Q$ to satisfy the above equality. In particular, define for simplicity $\mu(\mathbf{q}) = \frac{\frac{d}{dq_1} t_2(\mathbf{q})}{\frac{d}{dq_1} t_1(\mathbf{q})}$ and take derivative of the above equality with respect to q_1

$$\bar{\phi}_2^Q(\mathbf{q}) = \phi_2^Q(\mathbf{q}) + (\bar{\phi}_1^Q - \phi_1^Q)(\mathbf{q}) \cdot \mu(\mathbf{q}) - \int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1, q_2) dq'_1 \cdot \frac{d}{dq_1} \mu(\mathbf{q}).$$

As a result, $\bar{\phi}^Q$ defined above is a weak amortization of revenue. The next lemma formally states the above discussion.

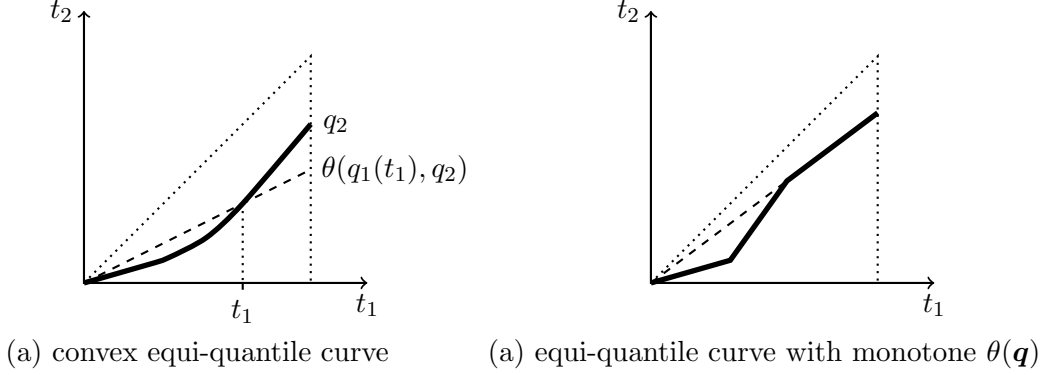


Figure 4: The connection between convexity of equi-quantile curves and monotonicity of $\theta(\mathbf{q})$ in q_1 . (a) Convexity implies monotonicity of $\theta(\mathbf{q})$. Convexity states that the line connecting any two points, namely $(0, 0)$ and $(t_1, t_2(q_1(t_1), q_2))$, lies above the curve for all $t'_1 \leq t_1$, and below the curve for all $t'_1 \geq t_1$. As a result, all the points on the curve with $t'_1 \leq t_1$ are mapped to a lower θ than $\theta(q_1(t_1), q_2)$, and all points with $t'_1 \geq t_1$ are mapped to higher θ than $\theta(q_1(t_1), q_2)$. (b) The function $\theta(\mathbf{q})$ can be monotone for a non-convex curve.

Lemma 13. *The virtual surplus, with respect to $\bar{\phi}^Q$ of any incentive compatible allocation $\mathbf{x}$ upper bounds its revenue. If x_1 is only a function of q_1 (equivalently t_1), $x'_1(q_1) = 0$ whenever $\int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1) dq'_1 > 0$, and $x_2(\mathbf{q}) = 0$ for all $\mathbf{q}$, the expected virtual surplus with respect to $\bar{\phi}^Q$ equals revenue.*

We will finally need to verify that $\bar{\phi}^Q$ also satisfies the properties required for ex-post optimization. The following lemma identifies convexity of equi-quantile curves as a sufficient condition (proof in Appendix B). Convexity of equi-quantile curve is more restrictive than monotonicity of $\theta(\mathbf{q})$ in q_1 (see Figure 5).

Lemma 14. *If the equi-quantile curves are convex for all q_2 , the weak amortization $\bar{\phi}^Q$ defined above satisfies $\theta(\mathbf{q})\bar{\phi}_1^Q(\mathbf{q}) \geq \bar{\phi}_2^Q(\mathbf{q})$. As a result, $\bar{\phi}_1^Q \geq \bar{\phi}_2^Q$ whenever either is positive.*

In particular for any MR-log-independent distribution the equi-quantile curves are straight rays, satisfying the above condition. By combining Lemma 15 and Lemma 16 we conclude the following theorem.

Theorem 15. *For a unit demand agent with $m = 2$ items, uniform pricing is optimal for any distribution for which the equi-quantile curves are convex for all q_2 .*

4.3 Necessary Conditions for Optimality of Uniform Pricing

We now provide a partial converse to Theorem 11 by specifying conditions under which uniform pricing is not revenue optimal. The proof follows the approach of Thanassoulis (2004) in showing that uniform pricing is not optimal for the uniform independent distribution on $[5, 6]$. Under the conditions of the theorem, the best uniform pricing mechanism can be improved by adding a lottery with slightly lower price, that offers each item with probability a half. The proof is deferred to Appendix B.

Theorem 16. *Uniform pricing is not revenue optimal if for any price t_1 that satisfies the first order condition of optimizing revenue $t_1(1 - F_{\max}(t_1))$,*

$$t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} = 0$$

the tradeoff between more sales and the loss of revenue from lowering the price as a result of adding a lottery to the menu is positive

$$t_1 - 2 \frac{\int_{t'_1=t_1}^1 f(t'_1, t'_1) dt'_1}{f(t_1, t_1)} > 0.$$

4.4 Bundle Pricing for General Distributions

We can use the generality of unit-demand settings and the above result to identify conditions under which bundle pricing is optimal for general distributions. More precisely, notice that any problem with finite outcome space can be converted to a unit-demand problem with dot-product utility function by letting $\mathbf{t}$ and $\mathbf{x}$ be vectors of dimension equal to the cardinality of the outcome space, with t_i being the value for outcome i . For multi-item settings with m items, the size of the outcome space is equal to 2^m . The input to the problem is a density function $f(t_1, \dots, t_k)$ over vectors of dimension $k = 2^m$. We assume *free disposal*, meaning that the density of a type is non-zero only if the valuation for each bundle is at least as much as the value for any other subset of that bundle. The max-ratio representation of the density function is a $f^{MR}(t_1, \theta_2, \dots, \theta_k)$, where t_1 , normalized to be at most 1, is the value for the grand bundle, and $\theta_i \leq 1$ is the ratio of the value for the i -th subset, over the value of the grand bundle. Now a bundle pricing mechanism corresponds to a mechanism in the unit demand setting which only sells the favorite outcome, which is the grand bundle. The theorem below follows immediately from the extension of Theorem 11 to the case of more than two items (see Appendix A).

Theorem 17. *Bundle pricing is optimal if the max-ratio representation of the density function is MR-log-supermodular for which the strong amortization of the favorite item projection is monotone.*

A density function is MR-log-supermodular if its max-ratio representation f^{MR} is log-supermodular, that is, for any two vectors $(t_1, \theta_2, \dots, \theta_k)$ and $(t'_1, \theta'_2, \dots, \theta'_k)$,

$$f^{MR}(t_1, \dots, \theta_k) \cdot f^{MR}(t'_1, \dots, \theta'_k) \leq f^{MR}(\min(t_1, t'_1), \dots, \min(\theta_k, \theta'_k)) \cdot f^{MR}(\max(t_1, t'_1), \dots, \max(\theta_k, \theta'_k)).$$

5 Multi-agent Extension

A multi-agent problem is defined by n agents, each agent κ associated with a distribution f^κ , and a feasibility setting. We focus on *service-constrained* feasibility settings. A service-constrained setting is parameterized by a set system $\mathcal{S} \subseteq 2^n$, where an integral allocation $\mathbf{x}$ is feasible if the set of served agents is an element of $\mathcal{S}$, that is, $\{\kappa \mid \sum_{j \leq m} x_{\kappa j} = 1\} \in \mathcal{S}$ (agents are unit demand, which implies that $\sum_{j \leq m} x_{\kappa j} \leq 1$ for any agent κ). Type of agent κ is drawn independently of other agents from f^κ . A multi-agent mechanism is a profile of feasible allocation functions $(\hat{\mathbf{x}}^1(\mathbf{t}^1, \dots, \mathbf{t}^n), \dots, \hat{\mathbf{x}}^n(\mathbf{t}^1, \dots, \mathbf{t}^n))$ in which $\hat{\mathbf{x}}^\kappa(\mathbf{t}^1, \dots, \mathbf{t}^n) \in \mathbb{R}^m$ is the allocation of agent

κ , together with a profile of payment function $(\hat{p}^1(\mathbf{t}^1, \dots, \mathbf{t}^n), \dots, \hat{p}^n(\mathbf{t}^1, \dots, \mathbf{t}^n)) \in \mathbb{R}^n$ in which $\hat{p}^\kappa(\mathbf{t}^1, \dots, \mathbf{t}^n) \in \mathbb{R}$ is the payment of agent κ . We can use linearity of expectation and Definition 1 to conclude that for any strong amortization ϕ the revenue of a multi-agent mechanism can be written as the expected virtual surplus

$$\int_{\mathbf{t}} \left[\sum_{\kappa} \hat{\mathbf{x}}^\kappa(\mathbf{t}^\kappa, \mathbf{t}^{-\kappa}) \cdot \phi^\kappa(\mathbf{t}^\kappa) \right] f(\mathbf{t}) \, d\mathbf{t}.$$

We say that allocations $\hat{\mathbf{x}}^1, \dots, \hat{\mathbf{x}}^n$ optimize an objective ϕ *ex-post*, if for any input $\mathbf{t}^1, \dots, \mathbf{t}^m$ they select an outcome that maximizes $\sum_{\kappa} \hat{\mathbf{x}}^\kappa(\mathbf{t}^1, \dots, \mathbf{t}^m) \cdot \phi^\kappa(\mathbf{t}^\kappa)$.

Recall that under assumptions of Theorem 11 and Theorem 17, the strong and weak amortizations ϕ and $\bar{\phi}$ satisfy the following conditions: First, ϕ_1 and $\bar{\phi}_1$ are functions of t_1 only, and are monotone in t_1 . Second, $\phi_2 \leq \phi_1$ whenever either is positive (and the same holds for $\bar{\phi}$). As a result, an ex-post optimization of either amortization in multi-agent unit demand settings will result in incentive compatible mechanisms. The allocation of each agent's favorite alternative is only a function of the value for the favorite alternative, $\hat{\mathbf{x}}_1^\kappa(\mathbf{t}^\kappa, \mathbf{t}^{-\kappa}) = \hat{\mathbf{x}}_1^\kappa(t_1^\kappa, \mathbf{t}^{-\kappa})$, and also $\hat{\mathbf{x}}_2^\kappa(\mathbf{t}^\kappa, \mathbf{t}^{-\kappa}) = 0$. Recall from Lemma 15 that for any such allocation, the expected virtual surplus with respect to $\bar{\phi}$ is equal to revenue, implying that $\bar{\phi}$ is indeed a weak amortization of revenue in multi-agent settings. We conclude with the following two theorems, generalizing Theorem 11 and Theorem 17 to multi-agent service-constrained settings.

Theorem 18. *In service-constrained environments with unit demand agents with $m = 2$ alternatives, allocating the favorite alternative to agents maximizing the sum of the virtual values for the favorite alternatives $\phi_{\max}$ is optimal if $\phi_{\max}(t_1) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)}$ is monotone non-decreasing and $\frac{\int_0^{t_1} f(t_1, y) \, dy}{f_{\max}(t_1)}$ is a monotone non-increasing function of t_1 for all θ .*

Theorem 19. *In service-constrained environments with unit demand agents with $m = 2$ alternatives, allocating the favorite alternative to agents maximizing the sum of the ironed virtual values for the favorite alternatives if each agent has a distribution for which the equi-quantile curves are convex for all q_2 .*

The above results extend to the case of $m > 2$ alternatives via the analysis of Appendix A.

6 Optimality of Bundle Pricing for Additive Preferences

In this section we provide sufficient conditions for optimality of grand bundle pricing for agents with additive utilities. Similar to Section 4, we only focus on constructing a proof assuming that item one is the favorite item, as the proof generalizes to symmetric distributions easily by mirroring the construction for the other half of set of types. It is also easiest to express the results of this section when the sum of the values for the items are normalized to be at most one. We thus define the set of types to be $T = \{(t_1, t_2) | t_1, t_2 \geq 0, t_1 \geq t_2, t_1 + t_2 \leq 1\}$.

The *single-dimensional projection for the sum of values* is given by distribution and density function for the agent's sum of values, $F_{\text{sum}}(v)$ and $f_{\text{sum}}(v)$. The distribution function is the integral of f over $\mathbf{t}$ with $t_1 + t_2 \leq v$. The density function $f_{\text{sum}}(v)$ is the derivative of $F_{\text{sum}}(v)$ with

respect to v . As described in Section 3, the unique strong amortization of revenue for a single-dimensional agent (and thus for the single-dimensional projection) is $\phi_{\text{sum}}(v) = v - \frac{1 - F_{\text{sum}}(v)}{f_{\text{sum}}(v)}$. The strong amortization of utility $\alpha_{\text{sum}}/f_{\text{sum}}$ requires $\alpha_{\text{sum}}(v) = 1 - F_{\text{sum}}(v)$.

As in Section 4 the sum-of-values projection, via the divergence density equality, pins down a strong amortization of revenue ϕ . Under the conditions we will identify, this strong amortization may fail to be a virtual value function because pointwise optimization does not yield an incentive compatible mechanism. For this reason, we will begin by directly defining a weak amortization of revenue $\bar{\phi}$ and then prove that it is a virtual value function by showing that it satisfies the conditions of Lemma 7 (for any incentive compatible mechanism, expected virtual surplus with respect to $\bar{\phi}$ is at least the virtual surplus with respect to ϕ).

Definition 3. The *two-dimensional extension of the sum-of-values projection* $\bar{\phi}$ is:

$$\begin{aligned}\bar{\phi}_1(\mathbf{t}) &= \frac{t_1}{t_1 + t_2} \phi_{\text{sum}}(t_1 + t_2) = t_1 - \frac{t_1}{t_1 + t_2} \frac{1 - F_{\text{sum}}(t_1 + t_2)}{f_{\text{sum}}(t_1 + t_2)}, \\ \bar{\phi}_2(\mathbf{t}) &= \frac{t_2}{t_1 + t_2} \phi_{\text{sum}}(t_1 + t_2) = t_2 - \frac{t_2}{t_1 + t_2} \frac{1 - F_{\text{sum}}(t_1 + t_2)}{f_{\text{sum}}(t_1 + t_2)}.\end{aligned}$$

The following lemma provides conditions on vector field $\bar{\phi}$ such that virtual surplus maximization according to vector field $\bar{\phi}$ gives a bundle pricing.

Lemma 20. *Virtual surplus maximization according to vector field $\bar{\phi}$ gives a bundle pricing p (and is incentive compatible) if and only if: $\bar{\phi}_1(\mathbf{t}), \bar{\phi}_2(\mathbf{t}) \geq 0$ when $t_1 + t_2 \geq p$ and $\bar{\phi}_1(\mathbf{t}), \bar{\phi}_2(\mathbf{t}) \leq 0$ otherwise.*

Recall that in order to prove that $\bar{\phi}$ is a weak amortization of revenue, we need to show that the revenue of mechanism that optimizes $\bar{\phi}$ is equal to its expected virtual surplus with respect to $\bar{\phi}$ (we will later also verify that $\bar{\phi}$ provides an upper bound on the revenue of any mechanism). The above lemma implies that the that mechanism is bundle pricing. The following lemma proves that the equivalence of virtual surplus and revenue equivalence is satisfied for $\bar{\phi}$ and bundle pricing.

Lemma 21. *The expected revenue of a bundle pricing is equal to its expected virtual surplus with respect to the two-dimensional extension of the sum-of-values projection $\bar{\phi}$ satisfying $\bar{\phi}_1(\mathbf{t}) + \bar{\phi}_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$, where $\phi_{\text{sum}}(v) = v - \frac{1 - F_{\text{sum}}(v)}{f_{\text{sum}}(v)}$ is the strong amortization for agent's sum-of-values projection.*

We will next argue that $\bar{\phi}$ provides an upper bound on revenue of any mechanism. To give sufficient conditions for $\bar{\phi}$ to be a weak amortization of revenue, we study the existence of a strong amortization ϕ that satisfies the following refinement of Lemma 7, for any incentive compatible allocation $\mathbf{x}$ and sum s ,

$$\mathbf{E} \left[\mathbf{x}(\mathbf{t}) \cdot (\bar{\phi}(\mathbf{t}) - \phi(\mathbf{t})) \mid t_1 + t_2 = s \right] \geq 0. \quad (6)$$

Consider the strong amortization that, like the weak amortization, sets $\phi_1(\mathbf{t}) + \phi_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$ but, unlike the weak amortization, splits this total amortized value across the two coordinates to satisfy the divergence density equality. It is clear then that equation (6) can be expressed in terms of this relative difference. We will first show that it is sufficient for ϕ , relative to $\bar{\phi}$, to place more value on the second coordinate (corresponding to the less preferred item). Later in the section we

will describe sufficient conditions on the distribution to guarantee this sufficient condition on the strong amortization ϕ is satisfied.

To calculate the expectation in equation (6), it will be convenient to change to sum-ratio coordinate space. For a function h on type space T , define h^{SR} to be its transformation to sum-ratio coordinates, that is

$$h(t_1, t_2) = h^{SR}(t_1 + t_2, \frac{t_2}{t_1}).$$

Our derivation of sufficient conditions for the two-dimensional extension of the sum-of-values projection to be a weak amortization exploits two properties. First, the change in allocation probabilities of an incentive compatible mechanism, for a fixed sum s as the ratio θ increases, must not shift allocation probability from coordinate one to coordinate two (recall that we are assuming that $t_1 \geq t_2$). Second, if a strong amortization shifts value from coordinate one to coordinate two relative to the vector field $\bar{\phi}$, then, it also shifts *expected* value from coordinate one to coordinate two, conditioned sum $t_1 + t_2 = s$ and ratio $t_2/t_1 \leq \theta$. Thus, the shift in expected value only hurts the virtual surplus of strong amortization ϕ relative to vector field $\bar{\phi}$ and equation (6) is satisfied. Below, the first observation is summarized in Lemma 24, the second observation is employed in the proof of Lemma 25 which formalizes the argument sketched above.

Lemma 22. *The allocation of any incentive compatible mechanism satisfies*

$$\frac{d}{d\theta} \mathbf{x}^{SR}(s, \theta) \cdot (-1, 1) \geq 0$$

Proof. Incentive compatibility implies that the dot product of any vector and the change in allocation rule in the direction of that vector is non-negative (Lemma 3). In particular this must be true by fixing s and increasing θ . That is, for any s, θ , and ϵ ,

$$(\mathbf{x}^{SR}(s, \theta + \epsilon) - \mathbf{x}^{SR}(s, \theta)) \cdot (-1, 1) \frac{\theta \epsilon}{s} \geq 0.$$

Letting ϵ approach zero implies the claim. $\square$

Lemma 23. *The two-dimensional extension of the sum-of-values projection $\bar{\phi}$ is a weak amortization of revenue if there exists strong amortization ϕ with $\phi_1(\mathbf{t}) + \phi_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$ satisfies $\phi_1(\mathbf{t}) \frac{t_2}{t_1} \leq \phi_2(\mathbf{t})$.*

Proof. By symmetry of the distribution, there exists an optimal mechanism that is also symmetric. Therefore, using Proposition 8 we need to prove the lemma only for symmetric incentive compatible allocations (in particular, we assume that $x_1(t_1, t_1) = x_2(t_1, t_1)$ for all t_1).

Fix the sum $s = t_1 + t_1$. Denote the expected difference between $\bar{\phi}$ and ϕ conditioned on $t_2/t_1 \leq \theta$ by:

$$\mathbf{\Gamma}(s, \theta) = \int_{\theta'=0}^{\theta} [\bar{\phi} - \phi]^{SR}(s, \theta') f^{SR}(s, \theta') \frac{s}{1 + \theta} d\theta'.$$

We will only be interested in three properties of $\mathbf{\Gamma}$:

- (a) $\Gamma_2(s, \theta) = -\Gamma_1(s, \theta)$, i.e., this is the expected amount of value shifted from coordinate one to coordinate two of $\bar{\phi}$ relative to ϕ . This follows from the fact that $\phi_1(\mathbf{t}) + \phi_2(\mathbf{t}) = \bar{\phi}_1(\mathbf{t}) + \bar{\phi}_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$.

(b) $\Gamma_2(s, \theta) \geq 0$, i.e., this shift is non-negative according to the assumption of the lemma.

(c) $\Gamma(s, 0) = \mathbf{0}$, as the range of the integral is empty at $\theta = 0$.

Write the left-hand side of equation (6) as:

$$\begin{aligned} & \mathbf{E} \left[\mathbf{x}(\mathbf{t}) \cdot (\bar{\phi}(\mathbf{t}) - \phi(\mathbf{t})) \mid t_1 + t_2 = s \right] \\ &= \int_{\theta=0}^1 \mathbf{x}^{SR}(s, \theta) \cdot [\bar{\phi} - \phi]^{SR}(s, \theta) f^{SR}(s, \theta) \frac{s}{1 + \theta} d\theta \\ &= \int_{\theta=0}^1 \mathbf{x}^{SR}(s, \theta) \cdot \frac{d}{d\theta} \int_{\theta'=0}^{\theta} [\bar{\phi} - \phi]^{SR}(s, \theta') f^{SR}(s, \theta') \frac{s}{1 + \theta'} d\theta' d\theta. \end{aligned}$$

Substituting Γ into the integral above, we have

$$\begin{aligned} &= \int_{\theta=0}^1 \mathbf{x}^{SR}(s, \theta) \cdot \frac{d}{d\theta} \Gamma(s, \theta) d\theta \\ &= \mathbf{x}^{SR}(s, \theta) \cdot \Gamma(s, \theta) \Big|_{\theta=0}^1 - \int_{\theta=0}^1 \frac{d}{d\theta} \mathbf{x}^{SR}(s, \theta) \cdot \Gamma(s, \theta) d\theta. \\ &= - \int_{\theta=0}^1 \frac{d}{d\theta} \mathbf{x}^{SR}(s, \theta) \cdot \Gamma(s, \theta) d\theta ds \\ &\geq 0. \end{aligned}$$

The second equality is integration by parts. The third equality follows because the first term on the left-hand side is zero: For $\theta = 0$, $\Gamma(s, \theta) = \mathbf{0}$ by property (c); for $\theta = 1$, $x_1^{SR}(s, \theta) = x_2^{SR}(s, \theta)$ by symmetry, and $\Gamma_1(s, \theta) = -\Gamma_2(s, \theta)$ by property (a). The final inequality follows from Lemma 24 which shows that $-\frac{d}{d\theta} \mathbf{x}^{SR}(s, \theta) \cdot (1, -1) \geq 0$ and properties (a) and (b). $\square$

To identify sufficient conditions for $\bar{\phi}$ to be a weak amortization it now suffices to derive conditions under which there exists strong amortization ϕ satisfying $\phi_1(\mathbf{t}) + \phi_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$ and the condition of Lemma 25, i.e., $\phi_1(\mathbf{t}) \frac{t_2}{t_1} \leq \phi_2(\mathbf{t})$. Notice that $\alpha_1 \frac{t_2}{t_1} \geq \alpha_2$ implies that $\phi_1 \frac{t_2}{t_1} \leq \phi_2$ because

$$\frac{t_2}{t_1} \phi_1(\mathbf{t}) = \frac{t_2}{t_1} \left(t_1 - \frac{\alpha_1(\mathbf{t})}{f(\mathbf{t})} \right) = \frac{t_2}{t_1} \left(t_1 - \frac{\alpha_1(\mathbf{t})}{f(\mathbf{t})} \right) \leq t_2 - \frac{\alpha_2(\mathbf{t})}{f(\mathbf{t})} = \phi_2(\mathbf{t}).$$

Thus, it suffices to identify conditions under which the strong amortization of utility α/f satisfies $\alpha_1 \frac{t_2}{t_1} \geq \alpha_2$. Define distribution function $F_{\text{sum}}(\cdot, \theta)$ and density function $f_{\text{sum}}(\cdot, \theta)$ for the distribution conditioned on $t_2/t_1 \leq \theta$. In other words, $F_{\text{sum}}(s, \theta)$ is the probability of the set of types $\mathbf{t}$ such that $t_1 + t_2 \geq s$ and $t_2/t_1 \leq \theta$, and $f_{\text{sum}}(s, \theta) = \frac{d}{ds} F_{\text{sum}}(s, \theta)$. These definitions imply that $F_{\text{sum}}(s) = F_{\text{sum}}(s, 1)$ and $f_{\text{sum}}(s) = f_{\text{sum}}(s, 1)$.

Lemma 24. *The strong amortization $\phi = \mathbf{t} - \alpha/f$ satisfying $\phi_1(\mathbf{t}) + \phi_2(\mathbf{t}) = \phi_{\text{sum}}(t_1 + t_2)$ is unique and also satisfies*

$$\theta \alpha_1^{SR}(s, \theta) - \alpha_2^{SR}(s, \theta) = -(1 + \theta)(1 - F_{\text{sum}}(s)) \frac{d}{ds} \left[\frac{f_{\text{sum}}(s, \theta)}{f_{\text{sum}}(s)} \right].$$

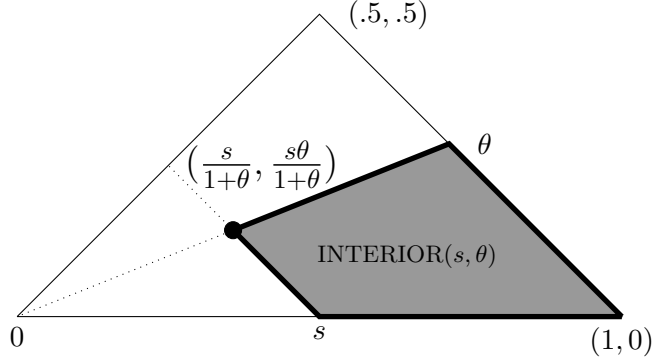


Figure 5: The trapezoidal set parameterized by s and θ , and the four curves that define its boundary, $\{\text{TOP, RIGHT, BOTTOM, LEFT}\}(s, \theta)$

Proof. We will here sketch the proof of the above lemma, and the complete proof appears in Appendix B.

The proof is similar to the proof of Lemma 9, and takes the derivative of the divergence theorem. We assume ϕ exists, use the divergence theorem and properties of ϕ to derive a closed form for it, and then verify that ϕ indeed satisfies all the required properties. We apply the divergence theorem to α on the trapezoidal subspace of type space defined by types $\mathbf{t}'$ with $t'_1 + t'_2 \geq s$, $t'_2/t'_1 \leq \theta$, $t'_2 \geq 0$, and $t'_1 + t'_2 \leq 1$ (Figure 6). The divergence theorem equates the integral of the vector field α on the boundary of the subspace to the integral of its divergence within the subspace. The integral of the vector field over right and bottom boundaries equates to zero by boundary orthogonality, and its value over the left boundary will be specified given the equation $\phi_1 + \phi_2 = \phi_{\text{sum}}(t_1 + t_2)$ (the outward pointing vector is $(-1, -1)$). As the upper boundary of this trapezoidal subspace has slope t_2/t_1 , one term in this equality is the integral of the dot product of $\alpha(\mathbf{t}')$ with the upward orthogonal vector to $\mathbf{t}$. Differentiating this integral and evaluating at $\mathbf{t}' = (\frac{s}{1+\theta}, \frac{s\theta}{1+\theta})$ gives a closed form expression for $\frac{t_2}{t_1}\alpha_1 - \alpha_2$. $\square$

We will next identify a class of distributions for which $\frac{f_{\text{sum}}(s, \theta)}{f_{\text{sum}}(s)}$ is a monotone non-decreasing function of s , which will together Lemma 26 imply the right direction on the angle of α . A distribution f is SR-log-submodular if it is log-submodular in sum-ratio coordinates, that is,

$$f^{SR}(s, \theta) \times f^{SR}(s', \theta') \leq f^{SR}(s', \theta) \times f^{SR}(s, \theta'), \quad \forall s \leq s', \theta \leq \theta',$$

and is SR-independent if the above holds with equality everywhere. Similar to Section 4 we prove the following lemma. (Notice that the sign is the opposite of the sign in Section 4, which is the reason that supermodularity is replaced by submodularity)

Lemma 25. *If f is a SR-log-submodular function, then for any s and θ ,*

$$\frac{d}{ds} \left[\frac{f_{\text{sum}}(s, \theta)}{f_{\text{sum}}(s)} \right] \geq 0,$$

with equality if distribution is SR-independent.

Theorem 26. *For additive agents with $m = 2$, bundle pricing is revenue optimal for any SR-log-submodular distribution for which the sum-of-values projection has monotone strong amortization ϕ_{sum} .*

References

- Acquisti, A. and Varian, H. R. (2005). Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381.
- Alaei, S., Fu, H., Haghpanah, N., and Hartline, J. (2013). The simple economics of approximately optimal auctions. In *FOCS*.
- Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 64(1):51–75.
- Armstrong, M. (1999). Price discrimination by a many-product firm. *The Review of Economic Studies*, 66(1):151–168.
- Borodin, A. and El-Yaniv, R. (1998). *Online Computation and Competitive Analysis*. Cambridge University Press, New York, NY, USA.
- Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2013). Mechanism design via optimal transport. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 269–286. ACM.
- Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2014). Strong duality for a multiple-good monopolist. *CoRR*, abs/1409.4150.
- Giannakopoulos, Y. and Koutsoupias, E. (2014). Duality and optimality of auctions for uniform distributions. In *ACM Conference on Electronic Commerce*.
- Hart, S. and Nisan, N. (2012). Approximate revenue maximization with multiple items. In *ACM Conference on Electronic Commerce, EC ’12, Valencia, Spain, June 4-8, 2012*, page 656.
- Hart, S. and Reny, P. J. (2014). Maximal revenue with multiple goods: Nonmonotonicity and other observations.
- Manelli, A. M. and Vincent, D. R. (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35.
- McAfee, R. P. and McMillan, J. (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic Theory*, 46(2):335–354.
- Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):pp. 58–73.
- Pavlov, G. (2011). A property of solutions to linear monopoly problems. *The BE Journal of Theoretical Economics*, 11(1).
- Pycia, M. (2006). Stochastic vs deterministic mechanisms in multidimensional screening.
- Riley, J. and Zeckhauser, R. (1983). Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289.
- Rochet, J. (1985). The taxation principle and multi-time hamilton-jacobi equations. *Journal of Mathematical Economics*, 14(2):113 – 128.
- Stokey, N. L. (1979). Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371.

Thanassoulis, J. (2004). Hagglng over substitutes. *J. Economic Theory*, 117(2):217–245.

Wang, Z. and Tang, P. (2014). Optimal mechanisms with simple menus. In *ACM Conference on Electronic Commerce*.

A Extension for $m > 2$ Items

In this section, we extend the definition of α given in Section 4 from two items to $m > 2$. The approach to derive such general α is similar to that of Section 4, and therefore in this section we only define α and verify its properties. Assume that we are in a partition of the space in which $t_1 \geq t_i$ for all i . Define the distribution of agent's favorite item $F_{\max}$ to be $F_{\max}(t_1) = \int_{\{t' | t'_1 \geq t_1\}} f(t') dt'$, and its density function $f_{\max}(t_1) = \frac{d}{dt_1} F_{\max}(t_1)$. Now define

$$\alpha_1(\mathbf{t}) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)}$$

Given α_1 we will next define α_i for $i \geq 2$ to automatically satisfy the density divergence and boundary orthogonality conditions. We define α_i for $i \geq 2$ as follows

$$\alpha_i(\mathbf{t})f(\mathbf{t}) = t_i f(\mathbf{t}) + \frac{1}{m-1} \int_{y=0}^{t_i} f(y, \mathbf{t}_{-i}) + \frac{d}{dt_1} [(t_1 - \alpha_1(y, \mathbf{t}_{-i}))f(y, \mathbf{t}_{-i})] dy$$

The above definition implies that

$$\frac{d}{dt_i} (t_i f(\mathbf{t}) - \alpha_i(\mathbf{t})f(\mathbf{t})) = -\frac{1}{m-1} \left(f(\mathbf{t}) + \frac{d}{dt_1} [(t_1 - \alpha_1(\mathbf{t}))f(\mathbf{t})] \right).$$

As a result,

$$\begin{aligned} \nabla \cdot \phi &= \sum_i \frac{d}{dt_i} (t_i f(\mathbf{t}) - \alpha_i(\mathbf{t})f(\mathbf{t})) \\ &= \frac{d}{dt_1} [(t_1 - \alpha_1(\mathbf{t}))f(\mathbf{t})] - (m-1) \times \frac{1}{m-1} \left(f(\mathbf{t}) + \frac{d}{dt_1} [(t_1 - \alpha_1(\mathbf{t}))f(\mathbf{t})] \right) \\ &= -f \end{aligned}$$

We now verify that ϕ satisfies the boundary conditions. This holds because when $t_i = 0$, $\alpha_i(\mathbf{t}) = 0$, and also when $t_1 = 1$, $\alpha_1(\mathbf{t}) = f(\mathbf{t})$.

Finally, we verify that $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$. This is again done in a manner similar to Section 4. Fix values of $\theta_3, \dots, \theta_m$, let $T_{\theta_3, \dots, \theta_n}(t_1, \theta_2)$ be the projection of type space into set of types such that each type $\mathbf{t}'$ satisfies $t'_1 \geq t_1$, $t'_1 \leq t'_2 \theta$, and $t'_i = t_1 \theta_i$. Now we can invoke the divergence theorem to conclude that $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$ if

$$\frac{d}{dt_1} \left(\frac{\int_{y=0}^{t_1 \theta_2} f(t_1, y, \theta_3, \dots, \theta_n) dy}{\int_{y=0}^{t_1} f(t_1, y, \theta_3, \dots, \theta_n) dy} \right) \leq 0.$$

Notice that fixing $\theta_3, \dots, \theta_n$, the above property is exactly what was required in two dimensions, which we showed follows from MR-log-supermodularity. As a result, if the function is MR-log-supermodular in every pair of variables t_1, θ_i for $i \geq 2$, then we $\alpha_i(\mathbf{t}) \leq \frac{t_i}{t_1} \alpha_1(\mathbf{t})$. Notice that this property is implied by MR-log-supermodularity of the distribution in all its variables, and therefore is a less demanding condition. We have therefore proved the following lemma.

Lemma 27. *If the distribution is MR-log-supermodular in every pair of variables t_1, θ_i for $i \geq 2$, then the revenue of an allocation can be upper bounded by*

$$\int_{\mathbf{t}} \mathbf{x} \cdot (1, \frac{t_2}{t_1}, \dots, \frac{t_n}{t_1}) \alpha(t_1) f(\mathbf{t}) \, d\mathbf{t},$$

where

$$\alpha_1(\mathbf{t}) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)}.$$

We conclude with the main theorem of this section.

Theorem 28. *For a single unit-demand agent with $m \geq 2$ items, uniform pricing is optimal for any MR-log-supermodular distribution for which the strong amortization of the favorite item projection is monotone.*

B Missing Proofs

B.1 Proof from Section 4

Proof of Lemma 16. We will start with the following two technical lemmas.

Lemma 29. *The first component of weak amortization $\bar{\phi}_1$ satisfies $\bar{\phi}_1(\mathbf{t}) \leq t_1$.*

Proof. In *un-ironed* regions, that is whenever $\bar{\phi}_1 = \phi_1$, by definition we have $\bar{\phi}_1(\mathbf{t}) = t_1 - \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \leq t_1$. If the curve is ironed between q_1 and $q'_1 \geq q_1$, then $\bar{\phi}_1^Q$ is the derivative of convex hull of ϕ_1^Q , which is $\int_0^q t_1(q') - \frac{q}{f_{\max}(t_1(q))} \, dq' = qt_1(q)$. Thus, for all q''_1 with $q_1 \leq q''_1 \leq q'_1$ we have

$$\begin{aligned} \bar{\phi}_1^Q(q''_1) &= \frac{q'_1 t'_1(q'_1) - q_1 t_1(q_1)}{q'_1 - q_1} \\ &\leq \frac{q'_1 t_1(q'_1) - q_1 t_1(q'_1)}{q'_1 - q_1} \\ &= t_1(q'_1) \leq t_1(q''_1). \end{aligned}$$

□

Lemma 30. *The strong amortization ϕ and the tangent to the equi-quantile curve at $\mathbf{q}$, $\mu(\mathbf{q}) =$ satisfy $\phi_1^Q(\mathbf{q})\mu(\mathbf{q}) - \phi_2^Q(\mathbf{q}) = t_1(\mathbf{q})\mu(\mathbf{q}) - t_2(\mathbf{q})$.*

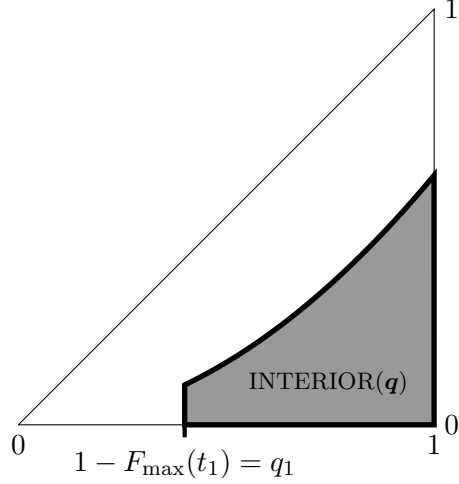


Figure 6: Given $\mathbf{q}$, $T(\mathbf{q})$ is the set of points with first quantile at most q_1 and second quantile at least q_2 . The four curves that define the boundary of $T(\mathbf{q})$ are $\{\text{TOP, RIGHT, BOTTOM, LEFT}\}(\mathbf{q})$.

Proof. We will first show that

$$\alpha_1^Q(\mathbf{q})\mu(\mathbf{q}) - \alpha_2^Q(\mathbf{q}) = 0.$$

We prove the above equation, similar to Lemma 9, by invoking the divergence theorem on the set of types $T(\mathbf{q}) = \{\mathbf{t} | q_1(\mathbf{t}) \leq q_1; q_2(\mathbf{t}) \geq q_2\}$ (see Figure 7),

$$\int_{\mathbf{t}' \in \text{TOP}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' = \int_{\mathbf{t}' \in \text{INTERIOR}(\mathbf{q})} \nabla \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \{\text{RIGHT, BOTTOM, LEFT}\}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}'.$$

Using divergence density equality and boundary orthogonality the right hand side becomes

$$\begin{aligned} &= - \int_{\mathbf{t}' \in \text{INTERIOR}(\mathbf{q})} f(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \{\text{LEFT}\}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' \\ &= -q_2(1 - F_{\max}(t_1(\mathbf{q}))) - \int_{\mathbf{t}' \in \{\text{LEFT}\}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' \\ &= -q_2(1 - q_1) - \int_{\mathbf{t}' \in \{\text{LEFT}\}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}', \end{aligned}$$

where the last two equalities followed directly from definitions of the quantile transformation. By definition of $\boldsymbol{\alpha}$,

$$\begin{aligned} \int_{\mathbf{t}' \in \{\text{LEFT}\}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' &= - \frac{1 - F_{\max}(t_1(\mathbf{q}))}{f_{\max}(t_1(\mathbf{q}))} \int_{t'_2, q_2(t_1(\mathbf{q}), t'_2) \geq q_2} f(t_1, t'_2) dt'_2 \\ &= -(1 - q_1)q_2. \end{aligned}$$

As a result,

$$\int_{\mathbf{t}' \in \text{TOP}(\mathbf{q})} \boldsymbol{\eta}(\mathbf{t}') \cdot \boldsymbol{\alpha}(\mathbf{t}') d\mathbf{t}' = 0.$$

Since the above equation must hold for all $\mathbf{q}$, we conclude that α is tangent to the equi-quantile curve at any $\mathbf{q}$. We can therefore write

$$\phi_1^Q(\mathbf{q})\mu(\mathbf{q}) - \phi_2^Q(\mathbf{q}) = t_1(\mathbf{q})\mu(\mathbf{q}) - t_2(\mathbf{q}) + \alpha(\mathbf{q}) \cdot (\mu(\mathbf{q}), -1) = t_1(\mathbf{q})\mu(\mathbf{q}) - t_2(\mathbf{q}),$$

completing the proof. $\square$

We can now complete the proof of Lemma 16. By rearranging the definition of ϕ_2 we get

$$\begin{aligned} \bar{\phi}_1^Q(\mathbf{q})\mu(\mathbf{q}) - \bar{\phi}_2^Q(\mathbf{q}) &= \phi_1^Q(\mathbf{q})\mu(\mathbf{q}) - \phi_2(\mathbf{q}) + \int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1, q_2) \, dq'_1 \cdot \frac{d}{dq_1}\mu(\mathbf{q}) \\ &= t_1(\mathbf{q})\mu(\mathbf{q}) - t_2(\mathbf{q}) + \int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1, q_2) \, dq'_1 \cdot \frac{d}{dq_1}\mu(\mathbf{q}) \\ &\geq t_1(\mathbf{q})\mu(\mathbf{q}) - t_2(\mathbf{q}), \end{aligned}$$

where the inequality followed since by definition of $\bar{\phi}_1^Q$, we have $\int_{q'_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1, q_2) \, dq'_1 \geq 0$, and $\frac{d}{dq_1}\mu(\mathbf{q}) \geq 0$ by the assumption of the lemma. We can now rearrange the above inequality and write

$$\begin{aligned} t_2(\mathbf{q}) - \bar{\phi}_2^Q(\mathbf{q}) &\geq \mu(\mathbf{q})(t_1(\mathbf{q}) - \bar{\phi}_1^Q(\mathbf{q})) \\ &\geq \theta(\mathbf{q})(t_1(\mathbf{q}) - \bar{\phi}_1^Q(\mathbf{q})), \end{aligned}$$

where the inequality followed since convexity of equi-quantile curves imply that $\mu(\mathbf{q}) \geq \theta(\mathbf{q})$, and by Lemma 31, $t_1(\mathbf{q}) - \bar{\phi}_1^Q(\mathbf{q}) \geq 0$.

We can now use the above inequality to write

$$\begin{aligned} \theta(\mathbf{q})\bar{\phi}_1^Q(\mathbf{q}) &= \theta(\mathbf{q})(t_1(\mathbf{q}) + (\bar{\phi}_1^Q(\mathbf{q}) - t_1(\mathbf{q}))) \\ &= t_2(\mathbf{q}) + \theta(\mathbf{q})(\bar{\phi}_1^Q(\mathbf{q}) - t_1(\mathbf{q})) \\ &\geq t_2(\mathbf{q}) + \bar{\phi}_2^Q(\mathbf{q}) - t_2(\mathbf{q}) \\ &= \bar{\phi}_2^Q(\mathbf{q}). \end{aligned}$$

$\square$

Proof of Theorem 18. Assume for contradiction that a uniform price p is optimal. This price should be the optimal price and therefore satisfies the first order condition $\frac{d}{dp}p(1 - F_{\max}(p)) = 0$ which implies that $pf_{\max}(p) = 1 - F_{\max}(p)$. Consider adding a lottery with price $p - \epsilon$, with ϵ approaching zero, that assigns each item with probability a half. We will study the change in revenue after this change, arguing that it must be positive under the assumptions of the theorem, thus contradiction optimality of uniform pricing p . The following set of types, $T(+)$, will switch from selecting nothing to selecting this new lottery (see Figure 8)

$$t_1 - p \leq 0, (t_1, t_2) \cdot (0.5, 0.5) - (p - \epsilon) \geq 0.$$

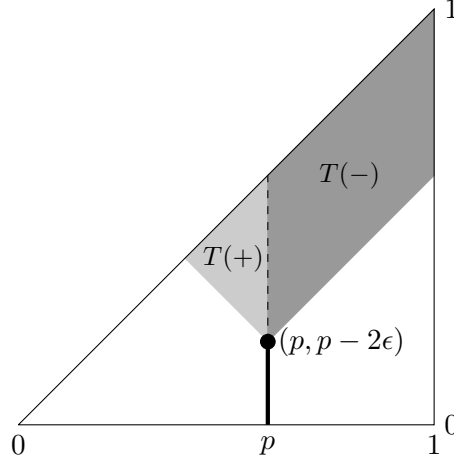


Figure 7: The effect of adding a lottery $(0.5, .0.5)$ with price $p - \epsilon$ to the menu consisting of lottery $(1, 0)$ with price p . Types in $T(+)$ will contribute an additional $p - \epsilon$ to revenue, and types in $T(-)$ will contribute to a revenue loss of ϵ .

Each type in this set contributes $p - \epsilon$ to the difference in revenue. The following set of types, $T(-)$, will switch from purchasing the item deterministically with price p to the half-half lottery with price $p - \epsilon$.

$$t_1 - p \geq 0, (t_1, t_2) \cdot (0.5, 0.5) - (p - \epsilon) \geq t_1 - p.$$

Each type in this set will contribute $-\epsilon$ to the difference in revenue. For the price p to be the optimal mechanism, the difference in revenue, weighted by the measure of the sets, must be negative. That is,

$$(p - \epsilon)f(T(+)) \leq \epsilon f(T(-)).$$

Expanding the above inequality and ignoring the terms with high orders of ϵ ,

$$pf(p, p)\epsilon^2/4 \leq 2\epsilon^2 \int_{x=p}^1 f(x, x),$$

which simplifies to

$$pf(p, p) \leq 2 \int_{x=p}^1 f(x, x).$$

This is in contradiction with the assumptions of the theorem. □

Proof of Lemma 14. We need to prove

$$f^{MR}(t_1, \theta) \times f^{MR}(t'_1, \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t'_1, \theta), \quad \forall t_1 \leq t'_1, \theta \leq \theta'.$$

Recall that $f^{MR}(t_1, \theta) = f(t_1, t_1\theta)$. Since the distribution is a product one, this implies that $f^{MR}(t_1, \theta) = f_1(t_1)f_2(t_1\theta)$. Notice that pair of values $t\theta'$ and $t'\theta$ have the same geometric mean as the pair $t\theta$, $t'\theta'$. Also given the assumptions, $t\theta \leq t'\theta$, $t\theta' \leq t'\theta'$. GG-convexity implies that

$$f_2(t_1\theta) \times f_2(t'_1\theta') \geq f_2(t\theta') \times f_2(t'\theta).$$

Multiplying both sides by $f_1(t_1) \times f_1(t'_1)$ we get

$$f_1(t_1)f_2(t_1\theta) \times f_1(t'_1)f_2(t'_1\theta') \geq f_1(t_1)f_2(t_1\theta') \times f_1(t'_1)f_2(t'_1\theta),$$

which since the distribution is a product distribution implies that

$$f^{MR}(t_1, \theta) \times f^{MR}(t'_1, \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t'_1, \theta).$$

□

B.2 Proofs from Section 6

Proof of Lemma 23. Let $\mathbf{x}^p$ be the allocation corresponding to posting price p for the bundle, that is $x_1^p(\mathbf{t}) = x_2^p(\mathbf{t}) = 1$ if $t_1 + t_2 \geq p$, and $x_1^p(\mathbf{t}) = x_2^p(\mathbf{t}) = 0$ otherwise. We will show that the virtual surplus of $\mathbf{x}^p$ is equal to the revenue of posting price p , $R(p) = p(1 - F_{sum}(p))$. The virtual surplus is

$$\begin{aligned} \int_{\mathbf{t} \in T} (\mathbf{x}^p \cdot \phi f)(\mathbf{t}) \, d\mathbf{t} &= \int_{\mathbf{t} \in T} \mathbf{x}^p(t_1, t_2) \cdot \bar{\phi}(t_1, t_2) f(t_1, t_2) \, d\mathbf{t} \\ &= \int_{\mathbf{t} \in T, t_1 + t_2 \geq p} \phi_{sum}(t_1 + t_2) f(t_1, t_2) \, d\mathbf{t}. \end{aligned}$$

For a function h on T , define h^{SR} to be its transformation to sum-ratio coordinates, that is

$$h(t_1, t_2) = h^{SR}(t_1 + t_2, \frac{t_2}{t_1})$$

By performing a change of variables $s = t_1 + t_2$, and $\theta = t_2/t_1$, the virtual surplus of $\mathbf{x}^p$ can be written as (the Jacobian of the transformation is $\frac{s}{(1+\theta)^2}$)

$$\begin{aligned} \int_{s=p}^1 \int_{\theta=0}^1 \phi_{sum}(s) f^{SR}(s, \theta) \frac{s}{(1+\theta)^2} \, d\theta ds &= \int_{s=p}^1 \phi_{sum}(s) \int_{\theta=0}^1 f^{SR}(s, \theta) \frac{s}{(1+\theta)^2} \, d\theta ds \\ &= \int_{s=p}^1 \phi_{sum}(s) f_{sum}(s) \, ds \end{aligned}$$

Replacing ϕ_{sum} by its definition,

$$\begin{aligned} &= \int_{s \geq p} t_1 f_{sum}(s) - (1 - F_{sum}(s)) \, ds \\ &= - \int_{s \geq p} \frac{d}{ds} (s(1 - F_{sum}(s))) \, ds \\ &= R(p) - R(1) = R(p). \end{aligned}$$

□

Proof of Lemma 22. We need to show that for the uniform price p , the allocation function $\mathbf{x}$ of posting a price p for the bundle optimizes ϕ pointwise. Pointwise optimization of $\mathbf{x} \cdot \bar{\phi}$ will result in $\mathbf{x} = (1, 1)$ whenever $\bar{\phi}_1, \bar{\phi}_2 \geq 0$, and will result in $\mathbf{x} = (0, 0)$ whenever $\phi_1, \phi_2 \leq 0$. □

Proof of Lemma 26. We assume that ϕ satisfying the requirements of the lemma exists, derive the closed form suggested in the lemma, and then verify that the derived ϕ indeed satisfies all the required properties. We apply the divergence theorem to α on the trapezoidal subspace of type space defined by types $\mathbf{t}'$ with $s \leq t'_1 + t'_2 \leq 1$, $t'_2/t'_1 \leq \theta$, and $0 \leq t'_1, t'_2 \leq 1$ (Figure 6). The divergence theorem equates the the integral of the vector field α on the boundary of the subspace to the integral of its divergence within the subspace. As the upper boundary of this trapezoidal subspace has slope t_2/t_1 , one term in this equality is the integral of $\alpha(\mathbf{t}')$ with the upward orthogonal vector to $\mathbf{t}$. Differentiating this integral gives the desired quantity.

Applying the divergence theorem to α on the trapezoid and expressing the top boundary as the interior divergence minus the other three boundaries gives:

$$\int_{\mathbf{t}' \in \text{TOP}(s, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}' = \int_{\mathbf{t}' \in \text{INTERIOR}(s, \theta)} \nabla \cdot \alpha(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \{\text{RIGHT}, \text{BOTTOM}, \text{LEFT}\}(s, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}'.$$

Since α/f is a strong amortization of utility, the divergence density equality and boundary orthogonality imply that the integral over the interior simplifies and the integrals over the right and bottom boundary are zero, respectively. We have,

$$\int_{\mathbf{t}' \in \text{TOP}(s, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}' = - \int_{\mathbf{t}' \in \text{INTERIOR}(s, \theta)} f(\mathbf{t}') d\mathbf{t}' - \int_{\mathbf{t}' \in \text{LEFT}(s, \theta)} \eta(\mathbf{t}') \cdot \alpha(\mathbf{t}') d\mathbf{t}'.$$

For the trapezoid parameterized by (s, θ) these integrals are (recall that the Jacobian of the transformation from $\mathbf{t}$ to (s, θ) is $\frac{s}{(1+\theta)^2}$),

$$\int_{s'=s}^1 \frac{\alpha^{SR}(s', \theta) \cdot (-\theta, 1)}{1 + \theta} ds' = - \int_{s'=s}^1 \int_{\theta'=0}^{\theta} \frac{f^{SR}(s', \theta') \cdot s}{(1 + \theta')^2} d\theta' ds' + \int_{\theta'=0}^{\theta} \frac{\alpha^{SR}(s, \theta') \cdot (-1, -1)s}{(1 + \theta')^2} d\theta'.$$

Differentiating with respect to s gives,

$$\frac{\alpha^{SR}(s, \theta) \cdot (-\theta, 1)}{1 + \theta} = \int_{\theta'=0}^{\theta} \frac{f^{SR}(s, \theta') \cdot s}{(1 + \theta')^2} d\theta' - \frac{d}{ds} \int_{\theta'=0}^{\theta} \frac{\alpha^{SR}(s, \theta') \cdot (-1, -1)s}{(1 + \theta')^2} d\theta'.$$

On the right-hand side, multiply first term by $\frac{f_{sum}(s)}{f_{sum}(s)} = 1$. The assumption that $\phi_1 + \phi_2 = \phi_{sum}(t_1 + t_2)$ implies that $\alpha_1^{SR}(s, \theta) + \alpha_2^{SR}(s, \theta) = \frac{1 - F_{sum}(s)}{f_{sum}(s)} f^{SR}(s, \theta)$ for all s and θ . These two terms then simplify by the product rule for differentiation to give the identity of the lemma.

$$\frac{\alpha^{SR}(s, \theta) \cdot (-\theta, 1)}{1 + \theta} = f_{sum}(s) \frac{\int_{\theta'=0}^{\theta} \frac{f^{SR}(s, \theta') \cdot s}{(1 + \theta')^2} d\theta'}{f_{sum}(s)} + \frac{d}{ds} \left[(1 - F_{max}(s)) \frac{\int_{\theta'=0}^{\theta} \frac{f^{SR}(s, \theta') \cdot s}{(1 + \theta')^2} d\theta'}{f_{max}(t_1)} \right].$$

By definition and change of variables, $F(s, \theta) = \int_{\mathbf{t}: t_1 + t_2 \leq s, t_2/t_1 \leq \theta} f(\mathbf{t}) d\mathbf{t} = \int_{s' \leq s} \int_{\theta' \leq \theta} f^{SR}(s', \theta') \frac{s'}{(1 + \theta')^2} d\theta' ds'$. Therefore, $f(s, \theta) = \int_{\theta' \leq \theta} f^{SR}(s, \theta') \frac{s}{(1 + \theta')^2} d\theta'$. Plugging this definition into the above equation, we get

$$\begin{aligned} \frac{\alpha^{SR}(s, \theta) \cdot (-\theta, 1)}{1 + \theta} &= f_{sum}(s) \frac{f_{sum}(s, \theta)}{f_{sum}(s)} + \frac{d}{ds} \left[(1 - F_{sum}(s)) \frac{f_{sum}(s, \theta)}{f_{sum}(s)} \right] \\ &= (1 - F_{sum}(s)) \frac{d}{ds} \left[\frac{f_{sum}(s, \theta)}{f_{sum}(s)} \right]. \end{aligned}$$

As a result,

$$\alpha^{SR}(s, \theta) \cdot (-\theta, 1) = (1 + \theta)(1 - F_{sum}(s)) \frac{d}{ds} \left[\frac{f_{sum}(s, \theta)}{f_{sum}(s)} \right].$$

We can now use the above equation, together with $\alpha_1^{SR}(s, \theta) + \alpha_2^{SR}(s, \theta) = \frac{1 - F_{sum}(s)}{f_{sum}(s)} f^{SR}(s, \theta)$ to solve for α_1 . $\square$

Proof of Lemma 27. We prove that for any θ, s , and s' such that $s < s'$,

$$\frac{f_{sum}(s, \theta)}{f_{sum}(s, 1)} \leq \frac{f_{sum}(s', \theta)}{f_{sum}(s', 1)}$$

The proof first converts the above form into max-ratio coordinates, applies SR-log-submodularity, and then transforms back to the standard form. Before applying SR-log-submodularity, we break down the integral set into two set, and apply SR-log-submodularity to only one of the integrals. More particularly, notice that

$$\begin{aligned} & f_{sum}(s, \theta) \times f_{sum}(s', 1) \\ &= \int_{\theta'=0}^{\theta} f^{SR}(s, \theta') \frac{s}{(1 + \theta')^2} d\theta' \times \int_{\theta''=0}^1 f^{SR}(s', \theta'') \frac{s'}{(1 + \theta'')^2} d\theta'' \quad (\text{change of variables}) \\ &= \int_{\theta'=0}^{\theta} \int_{\theta''=0}^{\theta} f^{SR}(s, \theta') \frac{s}{(1 + \theta')^2} f^{SR}(s', \theta'') \frac{s'}{(1 + \theta'')^2} d\theta'' d\theta' \\ &\quad + \int_{\theta'=0}^{\theta} \int_{\theta''=\theta}^1 f^{SR}(s, \theta') \frac{s}{(1 + \theta')^2} f^{SR}(s', \theta'') \frac{s'}{(1 + \theta'')^2} d\theta'' d\theta' \quad (\text{separate double integral into two sets}) \\ &\leq \int_{\theta'=0}^{\theta} \int_{\theta''=0}^{\theta} f^{SR}(s, \theta'') \frac{s}{(1 + \theta'')^2} f^{SR}(s', \theta') \frac{s'}{(1 + \theta')^2} d\theta' d\theta'' \quad (\text{rename variables } \theta' \text{ and } \theta'') \\ &\quad + \int_{\theta'=0}^{\theta} \int_{\theta''=\theta}^1 f^{SR}(s, \theta'') \frac{s}{(1 + \theta'')^2} f^{SR}(s', \theta') \frac{s'}{(1 + \theta')^2} d\theta'' d\theta' \quad (\text{apply MR-log-supermodularity}) \\ &= \int_{\theta''=0}^1 f^{SR}(s, \theta'') \frac{s}{(1 + \theta'')^2} d\theta'' \int_{\theta'=0}^{\theta} f^{SR}(s', \theta') \frac{s'}{(1 + \theta')^2} d\theta' \quad (\text{merge integrals}) \\ &= f_{sum}(s, 1) \times f_{sum}(s', \theta'). \end{aligned}$$

$\square$