Multi-dimensional Virtual Values and Second-degree Price Discrimination*

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

Jason Hartline
Northwestern University
EECS Department
hartline@northwestern.edu

July 22, 2016

Abstract

We consider a multi-dimensional screening problem of selling a product with multiple quality levels and specify conditions that imply optimality of only selling highest quality. A challenge of identifying optimal mechanisms for multi-dimensional agents is that the paths on which the incentive compatibility constraints bind are endogenous to the mechanism, and therefore the solution to an arbitrary relaxation of the problem may not be incentive compatible. We develop a methodology to identify an appropriate relaxation of the problem. Our methodology is general, and as a second application we use it to derive conditions for optimality of selling only the grand bundle of items to an agent with additive preferences.

*We would like to thank Nageeb Ali, Gabriel Carroll, Robert Kleinberg, Vijay Krishna, Alexey Kushnir, Preston McAfee, Ron Siegel, Rakesh Vohra, and seminar participants at Northwestern University, University of Chicago, Penn State University, University of Rochester, University of Pennsylvania, Simons Institute for Theoretical Computer Science, and Harvard University, as well as the participants at NBER Market Design 2014, Stony Brook Game Theory Conference and ACM Conference on Economics and Computation 2015, and the North American Econometric Society Meeting 2016.

1 Introduction

A monopolist seller can extract more of the surplus from consumers with heterogeneous tastes through second-degree price discrimination. While the optimal mechanism for selling a non-differentiated product to an agent with linear utility is a posted pricing, optimal mechanisms for a differentiated product can be complex and even generally require the pricing of lotteries over the variants of the product. This paper gives sufficient conditions under which the simple pricing of a non-differentiated product is optimal even when product differentiation is possible. These conditions allow multi-dimensional tastes to be projected to a single dimension where the pricing problem is easily solved by the classic theory. The identified conditions are natural and far more comprehensive than the previous known conditions.

Consider a monopolist who can sell a high-quality or low-quality product. The values of a consumer for these differentiated products can be seen as a point in the plane. It will be convenient to write the consumer's value for these two versions of the product as a base value for the high-quality product and the same base value times a discount factor for the low-quality product. It is a standard result of Stokey (1979) and Riley and Zeckhauser (1983) (and of Myerson, 1981, more generally) that when the base value is private but the discount factor is fixed, i.e., the values of the agent for the two qualities of products are distributed on a line through the origin, then selling only the high-quality good is optimal (and it is done by a posted price). The analysis of Armstrong (1996), applied to this setting, generalizes this result to the case where the base value and discount factor are independently distributed but both private to the agent. His result follows relaxing the problem to include incentive compatibility constraints only between types with the same discount factor. Since the discount factor is independent of the base value, the solution to the relaxed problem is globally incentive compatible, and therefore the same mechanism is optimal even when the discount factor is private. In contrast, our approach does not fix a relaxation in advance and therefore obtains more general conditions of optimality.

Our sufficient conditions generalize these results further to distributions where the base value and discount factor are positively correlated.[1] Notice that allowing arbitrary correlations between base value and discount factor is completely general as a multi-dimensional screening problem for a high- and low-quality product. Further, an example of Thanassoulis (2004) shows that the single-dimensional projection, i.e., selling only the high-quality prod-

[1]In this paragraph we assume that the marginal distribution of the base value is regular, i.e., Myerson's virtual value is monotone; and positive correlation is defined by first-order stochastic dominance. Generalizations are given later in the paper.

uct is not generally optimal with correlated base value and discount factor. Consider the special case where base value and discount factor are *perfectly correlated*, i.e., the discount factor is a function of the base value. Such an instance can be represented by a curve on which values for the differentiated products lie. We prove that if the curve only crosses lines from the origin from below, i.e., the discount factor is monotonically non-decreasing in the base value, then selling only the high-quality product is optimal. On the other hand, if the discount factor is not monotone in the base value then we show that there exists a distribution for the base value for which it is not optimal to sell only the high-quality product. Perfect correlation with a monotone discount factor is a special case of positive correlation which we show remains a sufficient condition for optimality of selling only the high-quality product.

From the analysis of the perfectly correlated case, we see that the analyses of Armstrong (1996) where the discount factor is independent of the base value, and Stokey (1979) and Riley and Zeckhauser (1983) where the discount factor is known, are at the boundary between optimality and non-optimality of selling only the high-quality product. Thus, these results are non-robust with respect to perturbations in the model, that is, optimality of selling only the high-quality product no longer generally holds if a distribution with a fixed discount factor is locally perturbed. Our result shows that pricing only the high-quality product remains optimal for any positive correlation; the more positively correlated the model is the more robust the result is to perturbations of the model.

Our characterization of positive correlation of the base value and discount factor as sufficient for the optimality of selling only the high-quality product is intuitive. Price discrimination can be effective when high-valued consumers are more sensitive to quality than low-valued consumers (Figure 1). These high-valued consumers would then prefer to pay a higher price for the high-quality product than to obtain the low-quality product at a lower price. Positive correlation between the base value and discount factor eliminates this possibility. It implies that high-valued agents are less sensitive to quality than low-valued agents.

As a qualitative conclusion from this work, optimal second-degree price discrimination, which is complex in general, cannot improve a monopolist's revenue over a non-differentiated product unless higher-valued types are more sensitive (with respect to the ratio of their values for high- and low-quality products) to product differentiation than lower-valued types. For example, a manufacturer has no advantage of intentionally damaging a good in order to price discriminate if the correlation property holds (Deneckere and McAfee, 1996, provide

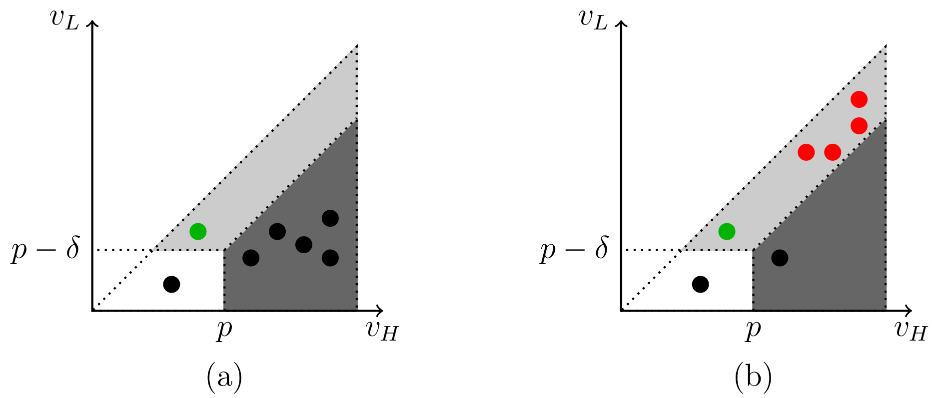


Figure 1: a) Types with high value for high-quality product v_H are relatively more sensitive to quality. Offering a discounted price $p - \delta$ for the low-quality product in addition to a price p for the high-quality product has a positive net effect of attracting the green type. b) Types with high value for high-quality product v_H are relatively less sensitive to quality. Offering a discounted price $p - \delta$ for the low-quality product in addition to a price p for the high-quality product has a positive effect of attracting the green type, and a negative effect of extracting lower payments from the red types.

examples of such practice). This simplification, generalizes to auction design with multiple agents. For example, a (monopolist) auctioneer on eBay has no advantage of discriminating based on expedited or standard delivery method if high-valued bidders discount delayed delivery less than low valued bidders (and if the costs are equal).

Our result above applies generally to a risk-neutral agent with quasi-linear utility over multiple alternatives, and identifies conditions for optimality of a mechanism that simply posts a uniform price for all alternative (i.e., the only non-trivial alternative assigned to each type is its favorite alternative). Applied a setting with a multi-product seller where the consumer can buy multiple items and alternatives correspond to bundles of items, this result indirectly gives conditions for optimality of posting a price for the grand bundle of items. If a uniform price is posted for all bundles, the consumer will only buy the grand bundle, or nothing (assuming free disposal).

The special case of this bundle pricing problem where the consumer's values are additive across the items has received considerable attention in the literature (Adams and Yellen, 1976; Hart and Nisan, 2012; Daskalakis et al., 2014) and our framework for proving optimality of single-dimensional projections can be applied to it directly. For this application, we employ a more powerful method of virtual values which is analogous to the ironing approach of Myerson (1981) to obtain more a more general condition for optimality of grand bundle pricing. We show that, for selling two items to a consumer with additive value, grand-bundle pricing is optimal when higher value for the grand bundle is negatively correlated with the ratio of values for the two items, i.e., when higher valued consumers have more heterogeneity in their tastes. This result formalizes a connection that goes back to Adams and Yellen (1976). This second application of our framework for proving the optimality of simple mechanisms further demonstrates its general applicability.

Our Approach. The main technical contribution of the paper, from which these sufficient conditions are identified, is a method for proving the optimality of a family of mechanisms for agents with multi-dimensional preferences, that extends the single-dimensional theory of virtual values of Myerson (1981). In settings with single-dimensional preferences, it is possible to order types and relax all but local incentive compatibility constraints (i.e., the constraints between a type and its higher and lower neighbors in the ordering). The relaxed problem can be solved by using the local incentive compatibility constrains to express expected profit in terms of the expected virtual surplus of allocation (i.e., virtual value minus the cost of allocation), and observing that the allocation function of an incentive compatible mechanism maximizes virtual surplus pointwise. The main challenge of multi-dimensional mechanism design is that the paths (in the agent's type space) on which the incentive constraints bind is a variable; thus a straightforward attempt to generalize single-dimensional virtual values to multi-dimensional agents by selecting an arbitrary relaxation of the problem fails since the optimal mechanism of the relaxed problem may not be incentive compatible. To prove optimality of a given mechanism, we develop a methodology to verify existence of a relaxation of the problem to local incentive compatibility such that the mechanism is an optimal solution to the relaxed problem. Importantly, our framework leaves the paths parameterizing a relaxation as a variable and solves for them.

The above characterizations show that the multi-dimensional pricing problem reduces to a single-dimensional projection where the agent's type is, with respect to the examples above, her base value. Our proof method instantiated for this problem is the following. We need to show the existence of a virtual value function for which (a) point-wise optimization of virtual surplus gives a mechanism that posts a price for the high-quality product and (b) expected virtual surplus equals expected revenue when the agent's type is drawn from the distribution. If the single-dimensional projection is optimal and (b) holds then it must be that the virtual value of the high-quality product is equal to the single-dimensional virtual value according to the marginal distribution of agent's value for the high-quality product. This pins down a degree of freedom in problem of identifying a virtual value function (which is generally given by integration by parts on the paths in type space, e.g., Rochet and Chone, 1998); the virtual value for the low-quality product can then be solved for from the high-quality virtual value and a differential equation that relates them. It then suffices to check that (a) holds, which in this case requires that, at any pair of values for the high and low qualities, (a.1) if the virtual value for the high-quality product is positive then it is at least the virtual value of the low-quality product and (a.2) if it is negative then they are both negative. Analysis of the constraints imposed by (a.1) and (a.2) then gives sufficient conditions on the distribution on types for optimality of the single-dimensional projection.

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, and that its gradient is equal to the allocation (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. In our approach, in contrast, 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 Wilson (1993), 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. Wilson (1993) and Armstrong (1996) gave closed form solutions for multi-dimensional screening problems. Their results are for nonlinear problems that are different from our model. Alaei et al. (2013) used integration by parts to get closed form solutions with independent and uniformly distributed values; our results generalize this one. Importantly, the paths for integration by parts in all these works is fixed a priori. In contrast, the choice of paths in our setting varies based on the distribution. Rochet and Chone (1998) showed that the general application of integration by parts (with parameterized choice of paths) characterizes the solutions of the relaxed problem where all but local incentive constraints are removed. However, the characterization is implicit and includes the choice of paths as parameters. They use the characterization to show that since bunching can happen, the solution to the relaxed problem is generically not incentive compatible.[2] Importantly, the observation is based on the placement of the outside option, in the form of a price for a certain allocation, that is the zero allocation in our setting. Compared to the above papers, our work is the first to use the variability of paths to derive explicit conditions of optimality (see Rochet and Stole, 2003, for an accessible survey).

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. 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 these works on multi-dimensional mechanism design in that it uses properties of a pre-specified family of mechanisms to pin down multi-dimensional virtual values that prove that mechanisms from the family are optimal.

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. 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 for proving optimality of multi-dimensional mechanisms, and find the optimal mechanism when values are i.i.d. from the uniform distribution (with up to six items). 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 (similar to the characterization of Rochet and Chone, 1998. 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

[2]Bunching refers to the case where different types are assigned the same allocation.

when values for $m \geq 2$ 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 $m = 2$ items. In comparison, a simple corollary of our theorem states that grand bundling is optimal for uniform draws from $[a, b]$ truncated such that the sum of the values is at most $a + b$, for *any* $a \leq b$.

2 Preliminaries

We consider a single-agent mechanism design problem with allocation space $X \subseteq [0,1]^m$ for a finite m, where the cost to the seller for producing allocation $\boldsymbol{x} \in X$ is $c(\boldsymbol{x})$ (our main theorems assume additional structure on the cost function).[3] The agent has a convex, continuous, and bounded type space $T \subset \mathbb{R}^m$ with Lipschitz continuous boundary. The type $\boldsymbol{t} \in T$ of the agent is drawn at random from a distribution with density $f > 0$. The utility of the agent with type $\boldsymbol{t} \in T$ for allocation $\boldsymbol{x} \in X$ and payment $p \in \mathbb{R}$ is $\boldsymbol{t} \cdot \boldsymbol{x} - p$.

We use the revelation principle and focus on direct mechanisms. A single-agent mechanism is a pair of functions, the allocation function $\boldsymbol{x} : T \to X$ and the payment function $p : T \to \mathbb{R}$. A mechanism is *incentive compatible* (IC) if no type of the agent increases its utility by misreporting,

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

A mechanism is *individually rational* (IR) if the utility of every type of the agent is at least zero,

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

The problem is to design an IC and IR mechanism that maximizes the *expected profit* of the seller, defined to the the seller's expected revenue minus cost

$$\mathbf{E}_{\boldsymbol{t} \sim f}\left[p(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))\right].$$

(We will subsequently drop $\boldsymbol{t} \sim f$ and simply write $\mathbf{E}[p(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))]$.)

A single agent mechanism $(\boldsymbol{x}, p)$ defines a utility function $u(\boldsymbol{t}) = \boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}) - p(\boldsymbol{t})$. The

[3]Throughout the paper we denote vectors by bold symbols, e.g., $\mathbf{v}$, and a component of a vector by a non-bold symbol, e.g., v_i.

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 individually-rational incentive-compatible mechanism if and only if u is convex, non-negative, and non-decreasing. The allocation is* $\boldsymbol{x}(\boldsymbol{t}) = \nabla u(\boldsymbol{t})$, wherever the gradient $\nabla u(\boldsymbol{t})$ is defined.[4]

Even though the above lemma completely characterizes incentive compatibility in our setting, we will mainly use the *envelope equality* $\boldsymbol{x}(\boldsymbol{t}) = \nabla u(\boldsymbol{t})$ as a necessary condition for incentive compatibility of a mechanism. This equality is obtained from the first order conditions of the utility maximization problem solved by the agent.

In what follows, we first develop our framework for a general allocation space X. Our main results consider the following allocation spaces.

The multi-alternative setting. (Section 4): We assume $X = \{\boldsymbol{x} \in [0,1]^m \mid \sum_i x_i \leq 1\}$. Here m is the number of alternatives, and x_i specifies the probability or the quantity of alternative i assigned ($1-\sum_i x_i$ is the probability of selecting an outside alternative for which the agent has zero value). For example, m may be the number of possible configurations, e.g., quality or delivery method, of a product to be sold to a unit-demand agent.[5] Another example is a multi-product seller that can offer different bundles of products. In this case, each $\boldsymbol{x}$ is a distribution over bundles, with x_i the probability of assigning bundle i, and m is be the number of possible bundles of the products.

In the multi-alternative setting we study optimality of the *uniform pricing* mechanism in which the *same price* is posted on all non-trivial alternatives (Section 4). In such a mechanism, the only relevant information a type contains is the value for the favorite alternative.

The multi-product setting with additive preferences. (Section 5): We assume $X = [0,1]^m$. Here m is the number of items, and an allocation $\boldsymbol{x}$ specifies the probability x_i of receiving each item i.[6]

In the multi-product setting with additive preferences we study optimality of the *grand bundle pricing* mechanism in which a price is posted for the grand bundle of items only

[4]The gradient $\nabla u(\boldsymbol{t})$ is a vector $(\partial_1 u(\boldsymbol{t}), \ldots, \partial_m u(\boldsymbol{t}))$, where $\partial_i u(\boldsymbol{t})$ is the partial derivative of the function u with respect to the i'th variable t_i. If u is convex, $\nabla u(\boldsymbol{t})$ is defined almost everywhere, and the mechanism corresponding to u is essentially unique. As a result, when writing the expected profit we will simply assume that $\nabla u(\boldsymbol{t})$ is defined everywhere.

[5]Even if the seller can produce and assign multiple configurations, since the agent is unit demand, then any deterministic allocation without loss of generality assigns only one alternative, and any randomized allocation is a distribution over alternatives.

[6]This setting is a special case of the multi-alternative setting with 2^m alternatives, but with the extra structure that the agent's valuation for a subset of items is equal to the sum of the values for the items in the bundle. The additivity structure allows us to focus on the lower dimensional space of items instead of alternatives.

(Section 5). In such a mechanism, the only relevant information a type contains is the value for the grand bundle $\sum_i t_i$.

3 Virtual Values and IC Relaxations

This section codifies the approach of incentive compatible mechanism design via virtual values and extends it to agents with multi-dimensional type spaces. We start with an overview. A standard approach to understanding optimal mechanisms is as follows: in the single-dimensional special case of the problem, i.e., if $m = 1$, types of the agent can be ordered such that *local* incentive compatibility (i.e., a type reporting to be either of its two adjacent neighbors in the ordering) and individual rationality for the lowest type is sufficient to guarantee incentive compatibility and individual rationality of a mechanism. Now consider a relaxation of the problem where only *downward local* incentive compatibility (a type reporting to be its lower adjacent neighbor) and individual rationality of the lowest type are kept and all other IC and IR constraints are relaxed; the envelope equality ($\nabla u(\boldsymbol{t}) = \boldsymbol{x}(\boldsymbol{t})$ in our setting) can be used to express profit in terms of expected *virtual surplus* of the allocation. Assuming that a regularity condition on the distribution of types holds, there exists an incentive compatible mechanism whose allocation function optimizes virtual surplus, implying that the mechanism is an optimal solution to the relaxed problem (with only downward local IC conditions and IR at the lowest type), and thus it is a solution to the original problem.[7]

The challenge with multidimensional preferences is that the paths on which incentive compatibility binds varies depending on the mechanism, and therefore it is not possible to order types such that only local incentive compatibility, implied by the ordering, of any mechanism would guarantee global incentive compatibility of the mechanism. Nevertheless, to prove optimality of a mechanism, it is sufficient to argue that the mechanism is the optimal solution to some relaxation of the constraints, even though those constraints alone do not imply global incentive compatibility. An example is a single-dimensional setting discussed above: when a regularity condition on the distribution holds, the solution to the relaxed problem where only downward local IC conditions are maintained is an IC and IR mechanism, even though those downward local IC conditions alone are not sufficient to guarantee incentive compatibility of all mechanisms. We extend this idea to our setting with multidimensional preferences as follows: consider a partial ordering on the set of types

[7]This approach applies more generally to settings where a *single-crossing* condition on preferences is satisfied. Importantly, preferences satisfying the single-crossing condition can be ordered and the approach discussed above extends.

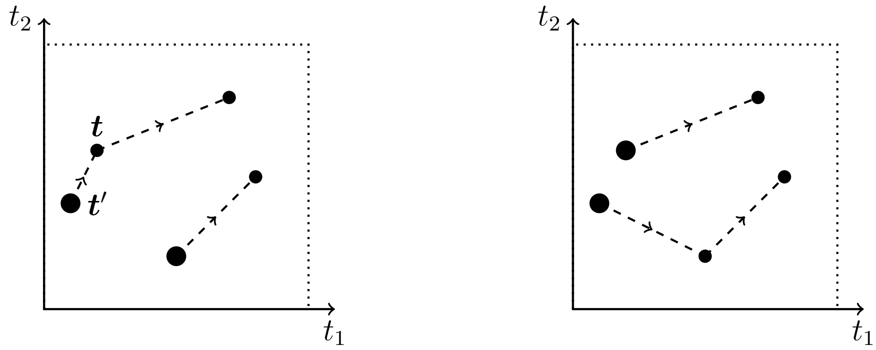


Figure 2: A relaxation of the problem is obtained by partitioning the type space into a family of paths, and imposing only local downward IC conditions along each path, e.g., type $\boldsymbol{t}$ reporting $\boldsymbol{t}'$, and an IR condition on the lowest type on each path, e.g., type $\boldsymbol{t}'$.

induced by a partitioning of the type space into a family of directed paths (see for example Figure 2), and a relaxation of the problem where all but one adjacent "downward" IC constraint for each type is relaxed (or the IR constraint if the type is the "lowest" type on a path). The downward IC constraints can be used (by applying the envelope equality) to express profit in the relaxed problem in terms of expected virtual surplus of allocation. If there exists an IC and IR mechanism whose allocation function pointwise optimizes virtual surplus, then the mechanism is the optimal solution to the relaxed problem, and therefore to the original problem.

To prove optimality of a given IC and IR mechanism, we answer the following question: does there exist a relaxation of the problem defined by a partitioning of type space into a family of paths as discussed above, whose induced virtual surplus is pointwise optimized by the allocation of the given mechanism? If the answer is positive, the mechanism is the solution to the relaxed problem and is therefore optimal. The rest of this section develops the methodology to formally verify the existence of such relaxation. Our main technical contribution is to identify a set of conditions that will imply that a given function is indeed a virtual surplus function induced by some relaxation (Lemma 3), and therefore can be used as an upper bound on revenue. Among the set of such functions, we need to verify existence of one that is optimized by the allocation function of the given mechanism.

3.1 Illustration: An Instance with Two Types

To make the above discussion more concrete, we analyze a simple scenario with two types. A reader eager to see our general treatment can skip directly to Section 3.2.

Consider an instance of the problem with two types $\boldsymbol{t}$ and $\boldsymbol{t}'$ with probabilities $f(\boldsymbol{t})$ and $f(\boldsymbol{t}')$ such that $f(\boldsymbol{t}) + f(\boldsymbol{t}') = 1$.[8] Assume for simplicity that costs are zero. The objective is to find a mechanism $(\boldsymbol{x}, p)$ to maximize the expected revenue

$$\mathbf{E}\left[p(\boldsymbol{t})\right] = p(\boldsymbol{t})f(\boldsymbol{t}) + p(\boldsymbol{t}')f(\boldsymbol{t}'), \tag{1}$$

subject to two incentive compatibility constraints (one for each ordered pair of types) and two individual rationality constraints (one for each type). Consider the individual rationality constraint for type $\boldsymbol{t}'$,

$$\boldsymbol{t}' \cdot \boldsymbol{x}(\boldsymbol{t}') - p(\boldsymbol{t}') \geq 0, \tag{2}$$

and the incentive compatibility constraint for type $\boldsymbol{t}$,

$$\boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}) - p(\boldsymbol{t}) \geq \boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}') - p(\boldsymbol{t}'). \tag{3}$$

Together, inequalities (2) and (3) imply an upper bound on the expected revenue (1) as explained below. In particular, (2) implies that

$$p(\boldsymbol{t}') \leq \boldsymbol{t}' \cdot \boldsymbol{x}(\boldsymbol{t}'),$$

and (3) implies that

$$p(\boldsymbol{t}) - p(\boldsymbol{t}') \leq \boldsymbol{t} \cdot (\boldsymbol{x}(\boldsymbol{t}) - \boldsymbol{x}(\boldsymbol{t}')).$$

We can now bound the expected revenue in terms of the allocation function given $f(\boldsymbol{t}) + f(\boldsymbol{t}') = 1$ and the above two inequalities,

$$\begin{aligned}\mathbf{E}\left[p(\boldsymbol{t})\right] = p(\boldsymbol{t})f(\boldsymbol{t}) + p(\boldsymbol{t}')f(\boldsymbol{t}') &= (p(\boldsymbol{t}) - p(\boldsymbol{t}'))f(\boldsymbol{t}) + p(\boldsymbol{t}') \\ &\leq [\boldsymbol{t} \cdot (\boldsymbol{x}(\boldsymbol{t}) - \boldsymbol{x}(\boldsymbol{t}'))]f(\boldsymbol{t}) + \boldsymbol{t}' \cdot \boldsymbol{x}(\boldsymbol{t}').\end{aligned}$$

Note for future reference that the above inequality is tight if both inequalities (2) and (3) are tight. To get a more tractable expression, we can rearrange the terms in the right hand

[8] In this subsection only we consider a distribution with discrete support and let f denote the probability of types.

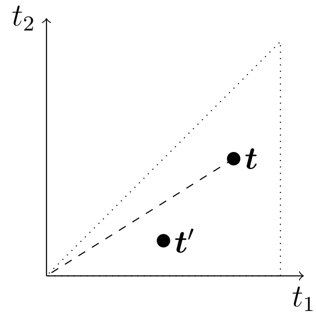


Figure 3: Type $\boldsymbol{t}'$ is below the ray that connects type $\boldsymbol{t}$ to the origin. Our analysis shows that either posting a price t_1 or t'_1 for alternative 1 is the optimal mechanism.

side of the above inequality to get

$$\begin{aligned}\mathbf{E}\left[p(\boldsymbol{t})\right] &\leq \boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t})f(\boldsymbol{t}) + [\boldsymbol{t}' - \boldsymbol{t}f(\boldsymbol{t})]\cdot\boldsymbol{x}(\boldsymbol{t}')\\ &= \boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t})f(\boldsymbol{t}) + [\boldsymbol{t}'f(\boldsymbol{t}') - (\boldsymbol{t}-\boldsymbol{t}')f(\boldsymbol{t})]\cdot\boldsymbol{x}(\boldsymbol{t}')\\ &= \boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t})f(\boldsymbol{t}) + [\boldsymbol{t}' - \frac{(\boldsymbol{t}-\boldsymbol{t}')f(\boldsymbol{t})}{f(\boldsymbol{t}')}]\cdot\boldsymbol{x}(\boldsymbol{t}')f(\boldsymbol{t}'),\end{aligned}$$

where the first equality followed from $f(\boldsymbol{t}) + f(\boldsymbol{t}') = 1$.

Define function $\bar{\boldsymbol{\phi}}$ as follows: $\bar{\boldsymbol{\phi}}(\boldsymbol{t}) = \boldsymbol{t}$ and $\bar{\boldsymbol{\phi}}(\boldsymbol{t}') = \boldsymbol{t}' - (\boldsymbol{t}-\boldsymbol{t}')f(\boldsymbol{t})/f(\boldsymbol{t}')$. The above analysis argues that for any mechanism that satisfies (2) and (3), we must have

$$\mathbf{E}\left[p(\boldsymbol{t})\right] \leq \mathbf{E}\left[\bar{\boldsymbol{\phi}}(\boldsymbol{t})\cdot\boldsymbol{x}(\boldsymbol{t})\right], \tag{4}$$

and that if the mechanism satisfies constraints (2) and (3) with equality, the above inequality holds with equality.

We now use the above analysis to prove the optimality of certain mechanisms for a class of instances. The example below is a special case of our main result Theorem 6.

Example 1. Consider a multi-alternative two-dimensional instance of the problem, that is, the two alternatives may be allocated with probabilities x_1 and x_2 ($x_1, x_2 \geq 0$ and $x_1 + x_2 \leq 1$), and assume alternative 1 is the favorite alternative for both types $t_1 \geq t_2$, $t'_1 \geq t'_2$, and that $t_1 \geq t'_1$. In addition, assume that $\boldsymbol{t}'$ is *below the ray* that connects $\boldsymbol{t}$ to the origin, that is, $t'_2/t'_1 \leq t_2/t_1$ (see Figure 3). We argue that a uniform pricing mechanism (with appropriate price) is the optimal solution to the relaxed problem with constraints (2) and (3), and therefore is optimal.

Recall that by definition, $\bar{\boldsymbol{\phi}}(\boldsymbol{t}) = \boldsymbol{t}$, and since $t_1 \geq t_2 \geq 0$, for any feasible $\boldsymbol{x}$ we must have

$$\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) = \boldsymbol{t} \cdot \boldsymbol{x}(\boldsymbol{t}) \leq t_1 = \bar{\phi}_1(\boldsymbol{t}). \tag{5}$$

The assumption that $t'_2/t'_1 \leq t_2/t_1$ implies that

$$\begin{aligned} \frac{t'_2}{t'_1}\bar{\phi}_1(\boldsymbol{t}') = \frac{t'_2}{t'_1}[t'_1 - (t_1 - t'_1)\frac{f(\boldsymbol{t})}{f(\boldsymbol{t}')}] = t'_2 - (t_1\frac{t'_2}{t'_1} - t'_2)\frac{f(\boldsymbol{t})}{f(\boldsymbol{t}')} \\ \geq t'_2 - (t_2 - t'_2)\frac{f(\boldsymbol{t})}{f(\boldsymbol{t}')} = \bar{\phi}_2(\boldsymbol{t}'). \end{aligned} \tag{6}$$

Now given the above inequality, we show that a mechanism that only offers the favorite alternative, alternative 1, is an optimal mechanism by arguing that it is the optimal solution with relaxed constraints (i.e., maintain (4) and (5) and relax the other two IC and IR constraints). Consider two cases:

- $\bar{\phi}_1(\boldsymbol{t}') \geq 0$. Given (6) and since $t'_2/t'_1 \leq 1$ by assumption, we have $\bar{\phi}_1(\boldsymbol{t}') \geq \bar{\phi}_2(\boldsymbol{t}')$, and thus for any feasible function $\boldsymbol{x}$, we have $\bar{\boldsymbol{\phi}}(\boldsymbol{t}') \cdot \boldsymbol{x}(\boldsymbol{t}') \leq \bar{\phi}_1(\boldsymbol{t}')$. Therefore, by (4) and (5), the optimal revenue is at most $f(\boldsymbol{t})\bar{\phi}_1(\boldsymbol{t}) + f(\boldsymbol{t}')\bar{\phi}_1(\boldsymbol{t}')$. Now consider a mechanism that posts a price t'_1 for alternative 1 (i.e., $\boldsymbol{x} = (1, 0)$ and $p = t'_1$ for both types). Since this mechanism satisfies constraints (2) and (3) with equality, its revenue is *equal to* $f(\boldsymbol{t})\bar{\phi}_1(\boldsymbol{t}) + f(\boldsymbol{t}')\bar{\phi}_1(\boldsymbol{t}')$, and thus is the optimal mechanism.

- $\bar{\phi}_1(\boldsymbol{t}') \leq 0$. Given (6), we have $\bar{\phi}_1(\boldsymbol{t}'), \bar{\phi}_2(\boldsymbol{t}') \leq 0$, and thus for any feasible function $\boldsymbol{x}$, we have $\bar{\boldsymbol{\phi}}(\boldsymbol{t}') \cdot \boldsymbol{x}(\boldsymbol{t}') \leq 0$. Therefore, by (4) and (5), the optimal revenue is at most $f(\boldsymbol{t})\bar{\boldsymbol{\phi}}_1(\boldsymbol{t})$. Now consider a mechanism that posts a price t_1 for alternative 1 (i.e., $\boldsymbol{x}(\boldsymbol{t}) = (1, 0)$, $p(\boldsymbol{t}) = t_1, \boldsymbol{x}(\boldsymbol{t}') = (0, 0), p(\boldsymbol{t}') = 0$). Since this mechanism satisfies constraints (2) and (3) with equality, its revenue is *equal to* $f(\boldsymbol{t})\bar{\phi}_1(\boldsymbol{t})$, and thus is the optimal mechanism.

Let us now summarize the above analysis to obtain a general approach for proving optimality of a given IC and IR mechanism $(\boldsymbol{x}, p)$. First, identify a set of IC and IR constraints (in our illustration, constraints (2) and (3)) and relax all other IC and IR constraints. Second, use the unrelaxed constraints to bound optimum value of the relaxed problem in terms of *expected virtual surplus of allocation* $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))]$ as in (4) (recall that the above analysis assumed that costs are zero and thus expected virtual surplus simplified to

$\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})]$). Third, argue that $(\boldsymbol{x}, p)$ is the optimum solution to the relaxed problem and therefore the optimum solution to the original problem by arguing that $\boldsymbol{x}$ optimizes virtual surplus subject to the feasibility condition $\boldsymbol{x}(\boldsymbol{t}) \in X$ for all $\boldsymbol{t}$, and that the revenue of $(\boldsymbol{x}, p)$ is indeed equal to the maximum of $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))]$ by a tightness argument.

To make the above approach operational, a methodology is required to identify the "right" relaxation of the IC and IR constraints, if such a relaxation exists. Indeed, not all relaxations would work even in our simple example with two types. A bit of analysis shows that if in Example 1 we had instead kept the other two IC and IR constraints and relaxed constraints (2) and (3), the solution to the relaxed problem would not have been an IC and IR mechanism. In the single-dimensional version of our problem, the right relaxation is the most natural one: keep the downward local IC constraints, and the IR constraint for the lowest type. Assuming a regularity condition holds, a mechanism that posts a price for deterministic allocation of the alternative optimizes the relaxed problem. On the other hand, in the fully general version of our setting, no natural ordering on types exists, and therefore it is not clear what the right relaxation would be. We therefore need a framework that allows us to search over the relaxations of the problem. We develop such a framework below.

3.2 Definitions: Amortizations and Virtual Values

Inspired by the above analysis, this section formally defines the main concepts used in our framework. The illustration above constructed a function $\bar{\boldsymbol{\phi}}$ such that the expected profit of any IC and IR mechanism $(\boldsymbol{x}, p)$ is upper bounded by its *expected virtual surplus* $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))]$, we call this condition on $\bar{\boldsymbol{\phi}}$ *amortization*;[9] if in addition, virtual surplus is optimized pointwise by the allocation function of an incentive compatible mechanism, we call $\bar{\boldsymbol{\phi}}$ a *virtual value function.* Existence of a virtual value function is a certificate to optimality of the mechanism.

Definition 1. A vector field[10] $\bar{\boldsymbol{\phi}} : T \to \mathbb{R}^m$ is an *amortization of revenue* if expected virtual surplus of any individually-rational incentive-compatible mechanism $(\hat{\boldsymbol{x}}, \hat{p})$ upper bounds the mechanism's expected profit, that is, $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \hat{\boldsymbol{x}}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}(\boldsymbol{t}))] \geq \mathbf{E}[\hat{p}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}(\boldsymbol{t}))]$, or

[9] This terminology comes from the design and analysis of algorithms in which an *amortized analysis* is one where the contributions of local decisions to a global objective are indirectly accounted for (see the textbook of Borodin and El-Yaniv, 1998). The correctness of such an indirect accounting is often proven via a *charging argument.* Myerson's construction of virtual values for single-dimensional agents can be seen as making such a charging argument where a low type, if served, is charged for the loss in revenue from all higher types.

[10] A vector field is simply a vector-valued function. Here, $\bar{\boldsymbol{\phi}}(\boldsymbol{t})$ is an m-dimensional vector.

equivalently $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \hat{\boldsymbol{x}}(\boldsymbol{t})] \geq \mathbf{E}[\hat{p}(\boldsymbol{t})]$; an amortization of revenue $\bar{\boldsymbol{\phi}}$ is *tight for a mechanism* $(\boldsymbol{x}, p)$ if the inequality above is tight, i.e., $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t})] = \mathbf{E}[p(\boldsymbol{t})]$.

Definition 2. An amortization of revenue $\bar{\boldsymbol{\phi}} : T \to \mathbb{R}^m$ is a *virtual value function* if a *pointwise virtual surplus maximizer* $\boldsymbol{x}$, i.e., $\boldsymbol{x}(\boldsymbol{t}) \in \arg\max_{\hat{\boldsymbol{x}} \in X} \hat{\boldsymbol{x}} \cdot \bar{\boldsymbol{\phi}}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}), \forall \boldsymbol{t} \in T$,[11] is incentive compatible and tight for $\bar{\boldsymbol{\phi}}$, i.e., there exists a payment rule p such that the mechanism $(\boldsymbol{x}, p)$ is incentive compatible, individually rational, and tight for $\bar{\boldsymbol{\phi}}$.

Proposition 2. *For any mechanism design problem that admits a virtual value function, its virtual surplus maximizer is the optimal mechanism.*

Proof. Denote the virtual surplus maximizer of Definition 2 by $(\boldsymbol{x}, p)$ and any alternative IC and IR mechanism by $(\hat{\boldsymbol{x}}, \hat{p})$; then,

$$\begin{aligned} \mathbf{E}\left[p(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))\right] &= \mathbf{E}\left[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \boldsymbol{x}(\boldsymbol{t}) - c(\boldsymbol{x}(\boldsymbol{t}))\right] \\ &\geq \mathbf{E}\left[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \hat{\boldsymbol{x}}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}(\boldsymbol{t}))\right] \geq \mathbf{E}\left[\hat{p}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}(\boldsymbol{t}))\right]. \end{aligned}$$

The expected profit of the mechanism is equal to its expected virtual surplus (first equality, by tightness). This expected virtual surplus is at least the virtual surplus of any alternate mechanism (first inequality, by pointwise optimality). (Note that this step does not require uniqueness of virtual surplus optimizers.) The expected virtual surplus of the alternative mechanism is an upper bound on its expected profit (second inequality, by definition of amortization). □

For instance, as the standard analysis goes and we will argue formally shortly, for a single-dimensional agent with value v in type space $T = [\underline{v}, \bar{v}]$, the function ϕ defined as $\phi(v) = v - \frac{1-F(v)}{f(v)}$ is an amortization of revenue and is tight for any mechanism with binding IR condition on $\underline{v}$. If in addition $\phi(v)$ is monotone non-decreasing and with constant marginal costs ($c(x) = cx$ for some non-negative c), there exists a threshold $v^* \geq \underline{v}$ such that virtual surplus $\phi(v)x - cx$ is at least zero if $v \geq v^*$, and at most zero otherwise (v^* need not be unique). Now consider a mechanism that posts a price v^* for the alternative ($x(v) = 1$ and $p(v) = v^*$ if $v \geq v^*$, and $x(v) = p(v) = 0$ otherwise): the allocation function of the mechanism optimizes virtual surplus pointwise, and the IR condition binds for the lowest type. Therefore ϕ is a virtual value function for the posted price mechanism with price v^* and the mechanism is optimal.

[11] Often this virtual surplus maximizer is unique up to measure zero events, when it is not then these conditions must hold for one of the virtual surplus maximizers and we refer to this one as *the* virtual surplus maximizer.

3.3 Canonical Amortizations

We now complement our definitions with a construction of a family of *canonical amortizations*. As outlined in the beginning of this section, there is a family of amortizations of revenue given by partitioning the type space into a family of paths and applying the envelope equality along each path to obtain an upper bound on profit in terms of virtual surplus of allocation. Since our goal is to find among possible amortizations one that additionally is maximized by the allocation of an incentive compatible mechanism, we obtain a set of conditions *directly on* $\boldsymbol{\phi}$ that guarantees that it is indeed an amortization. In particular, we argue that $\boldsymbol{\phi}$ defined as $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ is an amortization if vector field $\boldsymbol{\alpha}$ satisfies two properties:[12]

- *divergence density equality*: $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = -f(\boldsymbol{t})$ for all types $\boldsymbol{t} \in T$.[13]
- *boundary inflow*: $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) \leq 0$ for all types $\boldsymbol{t} \in \partial T$ where ∂T denotes the boundary of type space T, and $\boldsymbol{\eta}(\boldsymbol{t})$ is the outward pointing normal vector at type $\boldsymbol{t} \in \partial T$.

It is best to think of $\boldsymbol{\phi}$ as an amortization of revenue obtained by invoking downward local IC conditions along paths specified by $\boldsymbol{\alpha}$ (see Figure 4). In particular, for each type $\boldsymbol{t}$, the local incentive compatibility condition in the direction of $-\boldsymbol{\alpha}(\boldsymbol{t})$ is applied. The vector $\boldsymbol{\alpha}(\boldsymbol{t})$ corresponds to the revenue loss as a result of allocation to a type $\boldsymbol{t}$: to allocate to $\boldsymbol{t}$, the types that are "higher" than $\boldsymbol{t}$ on the path that crosses $\boldsymbol{t}$ must be offered discounts to maintain incentive compatibility. The divergence density equality condition requires $\boldsymbol{\alpha}$ to correspond to distributing the required density f on paths. Roughly speaking, it states that as one moves along a path, the marginal change in revenue loss $\boldsymbol{\alpha}$ at a type $\boldsymbol{t}$ is equal to its density $f(\boldsymbol{t})$. The boundary inflow condition ensures that the "highest" type on a path is assigned no revenue loss $\boldsymbol{\alpha} = 0$. In addition, the lemma below identifies conditions for $\boldsymbol{\phi}$ to be tight for a mechanism $(\boldsymbol{x}, p)$: the individual rationality condition must be binding at the start of a path.

Lemma 3. *For a vector field $\boldsymbol{\alpha} : T \to \mathbb{R}^m$ satisfying the divergence density equality and boundary inflow, the vector field $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ is an amortization of revenue; moreover, it is tight for any incentive compatible mechanism for which the participation con-*

[12] These conditions resemble optimality conditions of Rochet and Chone (1998). However, their analysis does not apply in our setting since it critically requires the cost function to be *strictly convex* (roughly speaking, to ensure that the profit maximization problem is strictly concave, and its unique locally optimal solution is also globally optimal). A more in depth comparison is provided in Appendix A.

[13] $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t})$ is the divergence of $\boldsymbol{\alpha}$ and is defined as $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = \partial_1 \alpha_1(\boldsymbol{t}) + \ldots + \partial_k \alpha_k(\boldsymbol{t})$ (not to be confused with $\nabla \boldsymbol{\alpha}(\boldsymbol{t})$ which is the gradient of $\boldsymbol{\alpha}$).

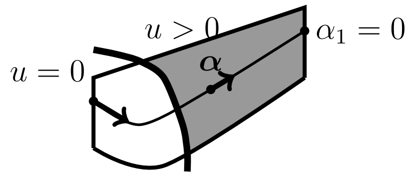


Figure 4: The vector $\boldsymbol{\alpha}$ specifies the direction of paths. For each type $\boldsymbol{t}$, $\boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ is the amortized loss in revenue imposed by incentive compatibility constraints from higher types on the path.

straint is binding for all boundary types with strict inflow, i.e., $u(\boldsymbol{t}) = 0$ *for* $\boldsymbol{t} \in \partial T$ *with* $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) < 0$.

Proof. The following holds for any incentive compatible mechanism. Integration by parts allows expected utility $\mathbf{E}[u(\boldsymbol{t})]$ to be rewritten in terms of gradient ∇u and vector field $\boldsymbol{\alpha}$ satisfying the divergence density equality.[14]

$$\begin{aligned}\int_{\boldsymbol{t}\in T} \nabla u(\boldsymbol{t}) \cdot \boldsymbol{\alpha}(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t} &= -\int_{\boldsymbol{t}\in T} u(\boldsymbol{t})\,(\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}))\, \mathrm{d}\boldsymbol{t} + \int_{\boldsymbol{t}\in\partial T} u(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t} \\ &= \int_{\boldsymbol{t}\in T} u(\boldsymbol{t})\, f(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t} + \int_{\boldsymbol{t}\in\partial T} u(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t}.\end{aligned}$$

The envelope equality ($\boldsymbol{x}(\boldsymbol{t}) = \nabla u(\boldsymbol{t})$, Lemma 1), implies that $\nabla u(\boldsymbol{t}) \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = \boldsymbol{x}(\boldsymbol{t}) \cdot \boldsymbol{\alpha}(\boldsymbol{t})$ and thus

$$\mathbf{E}\left[\tfrac{\boldsymbol{\alpha}(\boldsymbol{t})}{f(\boldsymbol{t})} \cdot \boldsymbol{x}(\boldsymbol{t})\right] = \mathbf{E}\left[u(\boldsymbol{t})\right] + \int_{\boldsymbol{t}\in\partial T} u(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t}. \tag{7}$$

Individual rationality implies that $u(\boldsymbol{t}) \geq 0$ for all $\boldsymbol{t} \in T$; combined with the assumed boundary inflow condition, the last term on the right-hand side is non-positive. Thus,

$$\mathbf{E}\left[\tfrac{\boldsymbol{\alpha}(\boldsymbol{t})}{f(\boldsymbol{t})} \cdot \boldsymbol{x}(\boldsymbol{t})\right] \leq \mathbf{E}\left[u(\boldsymbol{t})\right].$$

[14]Integration by parts for functions $h : \mathbb{R}^k \to \mathbb{R}$ and $\boldsymbol{\alpha} : \mathbb{R}^k \to \mathbb{R}^k$ over a set T with Lipschitz continuous boundary is as follows

$$\int_{\boldsymbol{t}\in T} (\nabla h \cdot \boldsymbol{\alpha})(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t} = -\int_{\boldsymbol{t}\in T} h(\boldsymbol{t})(\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}))\, \mathrm{d}\boldsymbol{t} + \int_{\boldsymbol{t}\in\partial T} h(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t},$$

where $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t})$ is the divergence of $\boldsymbol{\alpha}$ and $\boldsymbol{\eta}(\boldsymbol{t})$ is the outward pointing normal to the boundary at $\boldsymbol{t}$.

Revenue is surplus less utility; thus, $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ is an amortization of revenue, i.e.,

$$\begin{aligned} \mathbf{E}\left[p(\boldsymbol{t})\right] &= \mathbf{E}\left[\boldsymbol{t}\cdot\boldsymbol{x}(\boldsymbol{t}) - u(\boldsymbol{t})\right] \\ &\leq \mathbf{E}\left[(\boldsymbol{t} - \tfrac{\boldsymbol{\alpha}(\boldsymbol{t})}{f(\boldsymbol{t})})\cdot\boldsymbol{x}(\boldsymbol{t})\right] = \mathbf{E}\left[\boldsymbol{\phi}(\boldsymbol{t})\cdot\boldsymbol{x}(\boldsymbol{t})\right]. \end{aligned} \tag{8}$$

Finally, notice that if the last term of the right-hand side of equation (7) is zero, which holds for all mechanisms for which the individual rationality constraint is binding for types $\boldsymbol{t}$ on the boundary at which the paths specified by $\boldsymbol{\alpha}$ originate, then the inequalities above are equalities and the amortization is tight. □

Definition 3. A *canonical amortization of revenue* is $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ with $\boldsymbol{\alpha}$ satisfying the divergence density inequality and boundary inflow.

As discussed above, given $\boldsymbol{\phi}$, $\mathbf{E}[\boldsymbol{\phi}(\boldsymbol{t})\cdot\boldsymbol{x}(\boldsymbol{t})]$ is an upper bound on revenue obtained by only invoking incentive compatibility conditions along paths induced by α, and individual rationality conditions where the paths begin. As a result, we obtain the following corollary to Proposition 2 and Lemma 3.

Corollary 4. *If a canonical amortization $\boldsymbol{\phi}$ is a virtual value function for an IC and IR mechanism $(\boldsymbol{x}, p)$, then the mechanism is the solution to the relaxation of the problem induced by $\boldsymbol{\alpha}$.*

For a single-dimensional agent with value v in type space $T = [\underline{v}, \bar{v}]$, the canonical amortization of revenue that is tight for any non-trivial mechanism is unique and given by $\phi(v) = v - \frac{1-F(v)}{f(v)}$.[15] When it is monotone, pointwise virtual surplus maximization is incentive compatible, and thus the canonical amortization $\boldsymbol{\phi}$ is a virtual value function.

3.4 Reverse Engineering Virtual Value Functions

Multi-dimensional amortizations of revenue, themselves, do not greatly simplify the problem of identifying the optimal mechanism as they are not unique and in general virtual surplus maximization for such an amortization is not incentive compatible. The main approach of this paper is to consider a family of mechanisms (e.g., the family of uniform pricing mechanisms) and to add constraints imposed by tightness and virtual surplus maximization

[15] Divergence density equality implies that $\alpha(v) = \alpha(\underline{v}) - F(v)$. Tightness requires that $\alpha(\bar{v})u(\bar{v}) = (\alpha(\underline{v}) - 1)u(\bar{v}) = 0$. Since $u(\bar{v}) > 0$ for any non-trivial mechanism, we must have $\alpha(\underline{v}) = 1$ and thus $\alpha(v) = 1 - F(v)$. Tightness also requires that $\alpha(\underline{v})u(\underline{v}) = 0$, which is satisfied for any mechanism with binding participation constraint $u(\underline{v}) = 0$.

of this family of mechanisms to obtain a unique amortization. First, we will search for a single amortization that is tight for all mechanisms in the family. Second, we will consider virtual surplus maximization with a class of cost functions and require that a mechanism in the family be a virtual surplus maximizer for each cost (see Section 2). These two constraints pin down a degree of freedom in an amortization of revenue and allow us to solve for the amortization uniquely. The remaining task is to identify the sufficient conditions on the distribution such that a mechanism in the family is a virtual surplus maximizer. We will use this approach in Section 4, where we identify sufficient conditions on the distribution of types for the family of uniform pricing mechanisms to be optimal.

Our framework also allows for proving the optimality of mechanisms when no canonical amortization of revenue is a virtual value function. In the single-dimensional case, the ironing method of Mussa and Rosen (1978) and Myerson (1981), can be employed to construct, from the canonical amortization $\boldsymbol{\phi}$, another (non-canonical) amortization $\bar{\boldsymbol{\phi}}$ that is a virtual value function. The multi-dimensional generalization of ironing, termed sweeping by Rochet and Chone (1998), can similarly be applied to multi-dimensional amortizations of revenue. The goal of sweeping is to reshuffle the amortized values in $\boldsymbol{\phi}$ to obtain $\bar{\boldsymbol{\phi}}$ that remains an amortization, but additionally its virtual surplus maximizer is incentive compatible and tight. Our approach is to invoke the following proposition, which follows directly from the definition of amortization (Definition 2).

Proposition 5. *A vector field $\bar{\boldsymbol{\phi}}$ is an amortization of revenue if, for all incentive compatible individually rational mechanisms $(\hat{\boldsymbol{x}}, \hat{p})$ and some other amortization of revenue $\boldsymbol{\phi}$, it satisfies* $\mathbf{E}[\bar{\boldsymbol{\phi}}(\boldsymbol{t}) \cdot \hat{\boldsymbol{x}}(\boldsymbol{t})] \geq \mathbf{E}[\boldsymbol{\phi}(\boldsymbol{t}) \cdot \hat{\boldsymbol{x}}(\boldsymbol{t})]$.

We adopt the sweeping approach in Section 5 (and Theorem 11 which extends the main result of Section 4) to obtain a virtual value function $\bar{\boldsymbol{\phi}}$ from a canonical amortization $\boldsymbol{\phi}$. Just as there are many paths in multi-dimensional settings, there are many possibilities for the multi-dimensional sweeping of Rochet and Chone (1998). Our positive results using this approach will be based on very simple single-dimensional sweeping arguments.

4 Optimality of Favorite-alternative Projection

In this section we study conditions that imply a favorite-alternative projection mechanism is optimal in the multi-alternative setting (see Section 2). In that case, the problem collapses to a monopoly problem with a single parameter (the value for the favorite alternative), where we know from Riley and Zeckhauser (1983) that the optimum mechanism is *uniform pricing*: all nontrivial alternatives are deterministically and uniformly priced.

As discussed in Section 3.4, we use a class of cost functions to restrict the admissible amortizations. Throughout this section we assume uniform constant marginal costs, that is, $c(\boldsymbol{x}) = c\sum_i x_i$ for some constant service cost $c \geq 0$.[16] For simplicity we focus on the case of two alternatives (extension in Section 4.3).

4.1 General Distributions and Sufficient Conditions

We will now state the main theorem of this section which identifies sufficient conditions for optimality of uniform pricing for general distributions. We say that a distribution over $T \subset \mathbb{R}^2$ is *max-symmetric* if the distribution of maximum value $v = \max(t_1, t_2)$, conditioned on either $t_1 \geq t_2$ or $t_2 \geq t_1$, is identical.[17] Let $F_{\max}(v)$ and $f_{\max}(v)$ be the cumulative distribution and the density function of the value for favorite alternative. As described in Section 3, the amortization of revenue for a single-dimensional agent is $\phi_{\max}(v) = v - \frac{1-F_{\max}(v)}{f_{\max}(v)}$. Let $F(\theta|v,i)$ be the conditional distribution of the *value ratio* $\theta(\boldsymbol{t}) := \min(t_1,t_2)/\max(t_1,t_2)$ on $v = t_i \geq t_{-i}$, that is, $F(\theta|v,i) = \mathbf{Pr}_{\boldsymbol{t}}[\theta(\boldsymbol{t}) \leq \theta | v = t_i \geq t_{-i}]$.

Theorem 6. *Uniform pricing is optimal with $m = 2$ alternatives and uniform constant marginal costs $c(\boldsymbol{x}) = c\sum_i x_i$ for any $c \geq 0$ and max-symmetric distribution where (a) the favorite-alternative projection has monotone non-decreasing amortization of revenue $\phi_{\max}(v) = v - \frac{1-F_{\max}(v)}{f_{\max}(v)}$ and (b) the conditional distribution $F(\theta|v,i)$ is monotone non-increasing in v for all θ and i.*

Monotonicity of $F(\theta|v,i)$ in v is correlation of θ and v in first order stochastic dominance sense.[18] It states that as v increases, more mass should be packed between a ray parameterized by θ, and the 45 degree line connecting $(0,0)$ and $(1,1)$ (Figure 5). In other words, a higher favorite value makes relative indifference between alternatives, measured by θ, more likely.

As a class of distributions satisfying the conditions of Theorem 6, one can draw the maximum value v from a regular distribution $F_{\max}$, and the minimum value uniformly from

[16] Any instance with non-uniform marginal costs can be converted to an instance with zero cost by redefining value as value minus cost.

[17] As examples, any distribution with a domain $\boldsymbol{t} \in \mathbb{R}^2, t_1 \geq t_2$ is max-symmetric (since the distribution conditioned on $t_2 \geq t_1$ is arbitrary), as is any symmetric distribution over $\mathbb{R}^2$.

[18] Stronger correlation conditions, such as *Inverse Hazard Rate Monotonicity*, *affiliation*, and independence of favorite value v and the non-favorite to favorite ratio are also sufficient (Milgrom and Weber, 1982; Castro, 2007).

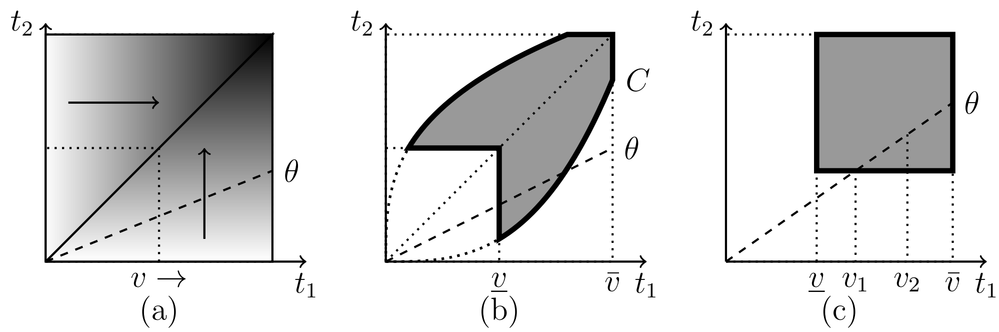
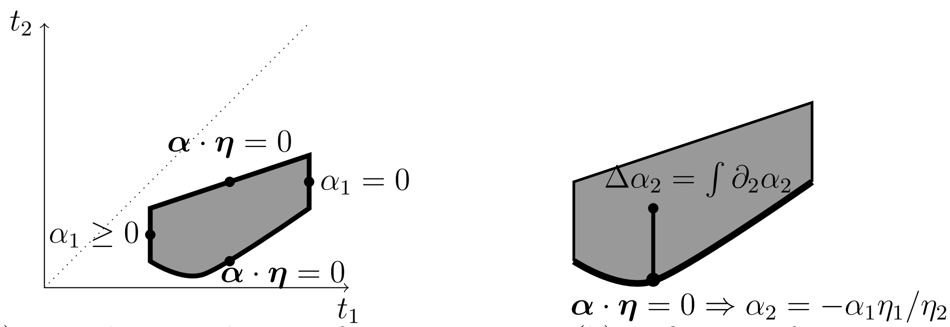


Figure 5: (a) As v increases, relatively more mass is packed towards the 45 degree line. (b) A class of distribution satisfying the correlation condition of Theorem 6. When $t_1 \geq t_2$, mass is distributed uniformly above a ratio-monotone curve C. (c) A class of distributions not satisfying the correlation condition of Theorem 6, since $0 = F(\theta|v_1, 1) < F(\theta|v_2, 1)$.

$[C(v), v]$, for a ratio-monotone function C satisfying $C(v) \leq v$ (Figure 5). On the other hand, a distribution where values for alternatives are uniformly and independently drawn from $[\underline{v}, \bar{v}]$, with $\underline{v} > 0$, does not satisfy the conditions (when $\underline{v} = 5, \bar{v} = 6$, Thanassoulis, 2004, showed that uniform pricing is not optimal). As another example, if t_1 and t_2 are drawn independently from a distribution with density proportional to $e^{h(\log(x))}$ for any monotone non-decreasing convex function h, then the distribution satisfies the conditions of the theorem (see Appendix B.3).

Notice that the conditional distributions $F(\theta|v, i)$ jointly with $F_{\max}$ are alternative representations of any max-symmetric distribution as follows: with probability $\mathbf{Pr}[t_1 \geq t_2]$, draw t_1 from $F_{\max}$, θ from $F(\cdot|t_1, 1)$, and set $t_2 = t_1\theta$ (otherwise assign favorite value to t_2 and draw θ from $F(\cdot|t_2, 2)$). As a result, the requirements of Theorem 6 on $F_{\max}$ and $F(\theta|v, i)$ are orthogonal. This view is particularly useful since it is natural to define several instances of the problem in terms of distributions over parameters v and θ. For example, in the pricing with delay model discussed in the introduction, θ has a natural interpretation as the discount factor for receiving an item with delay. We will revisit the conditions of Theorem 6 in Section 4.3.

The rest of this section proves the above theorem by constructing the appropriate virtual value functions. Notice that max-symmetry allows us to focus on only the conditional distribution when the favorite alternative is alternative 1. If a single mechanism, namely uniform pricing, is optimal for each case (of alternative 1 or alternative 2 being the favorite alternative), the mechanism is optimal for any probability distribution over the two cases. Therefore for the rest of this section we work with the distribution conditioned on $t_1 \geq t_2$.



(a) Boundary conditions of $\boldsymbol{\alpha}$ (b) Definition of α_2

Figure 6: (a) In addition to the divergence density equality $\nabla\cdot\boldsymbol{\alpha}=-f$, $\boldsymbol{\alpha}$ must be boundary orthogonal at all boundaries except possibly the left boundary, and an inflow at the left boundary. (b) Given α_1, we solve for α_2 to satisfy boundary orthogonality at the bottom and divergence density equality. Boundary orthogonality uniquely defines α on the bottom boundary. Integrating the divergence density equality $\partial_2\alpha_2=-f-\partial_1\alpha_1$ defines α_2 everywhere.

In particular, T is a bounded subset of $\mathbb{R}^2$ specified by an interval $[\underline{t}_1,\bar{t}_1]$ of values t_1 and bottom and top boundaries $\underline{t}_2(t_1)$ and $\bar{t}_2(t_1)$ satisfying $\underline{t}_2(t_1)\leq\bar{t}_2(t_1)\leq t_1$. The proof follows the framework of Section 3. In Definition 4 we define $\boldsymbol{\phi}$ and $\boldsymbol{\alpha}$ from the properties they must satisfy to prove optimality of uniform pricing. Lemma 7 shows that $\boldsymbol{\phi}$ is a canonical amortization and is tight for any uniform pricing. Lemma 8 shows that given the conditions of Theorem 6 on the distribution, the allocation of uniform pricing maximizes virtual surplus pointwise with respect to $\boldsymbol{\phi}$. The theorem follows from Proposition 2.

A uniform pricing $p\in[\underline{t}_1,\bar{t}_1]$ implies $\boldsymbol{x}(\boldsymbol{t})=0, u(\boldsymbol{t})=0$ if $t_1\leq p$, and $\boldsymbol{x}(\boldsymbol{t})=(1,0), u(\boldsymbol{t})>0$ otherwise (recall the assumption that $t_1\geq t_2$). Therefore, in order to satisfy the requirement of Lemma 3 that $u(\boldsymbol{t})(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})=0$ everywhere on the boundary and for all uniform pricings $p\in[\underline{t}_1,\bar{t}_1]$, $\boldsymbol{\alpha}$ must be *boundary orthogonal*, $(\boldsymbol{\alpha}\cdot\boldsymbol{\eta})(\boldsymbol{t})=0$, except possibly at the left boundary, where $u(\boldsymbol{t})=0$ (Figure 6). With this refinement of Lemma 3 of the boundary conditions of $\boldsymbol{\alpha}$ we now define $\boldsymbol{\alpha}$ and $\boldsymbol{\phi}$.

Definition 4. The *two-dimensional extension* $\boldsymbol{\phi}$ of the amortization for the favorite-alternative projection $\phi_{\max}(v)=v-\frac{1-F_{\max}(v)}{f_{\max}(v)}$ is constructed as follows:

(a) Set $\phi_1(\boldsymbol{t})=\phi_{\max}(t_1)$ for all $\boldsymbol{t}\in T$.

(b) Let $\alpha_1(\boldsymbol{t})=(t_1-\phi_1(\boldsymbol{t}))\,f(\boldsymbol{t})=\frac{1-F_{\max}(t_1)}{f_{\max}(t_1)}\,f(\boldsymbol{t})$.[19]

[19]Our assumption that $f>0$ and the regularity assumptions on T imply that $f_{\max}>0$ everywhere except potentially at the left boundary if the left boundary is a singleton. We treat this case separately in the upcoming proof of the theorem.

(c) Define $\alpha_2(\boldsymbol{t})$ uniquely to satisfy divergence density equality $\partial_2\alpha_2 = -f - \partial_1\alpha_1$ and boundary orthogonality of the bottom boundary.

(d) Set $\phi_2(\boldsymbol{t}) = t_2 - \alpha_2(\boldsymbol{t})/f(\boldsymbol{t})$.

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

(a) First, $\phi_1(\boldsymbol{t})$ may only be a function of t_1; otherwise, if $\phi_1(\boldsymbol{t}) > \phi_1(\boldsymbol{t}')$ with $t_1 = t_1'$, maximizing virtual surplus pointwise with cost c satisfying $\phi_1(\boldsymbol{t}) > c > \phi_1(\boldsymbol{t}')$ implies $x_1(\boldsymbol{t}') = 0$, and either $x_1(\boldsymbol{t}) > 0$ or $x_2(\boldsymbol{t}) > 0$ (if $\phi_2(\boldsymbol{t}) > \phi_1(\boldsymbol{t}) > c$). Such an allocation $\boldsymbol{x}$ is not the allocation of uniform pricing.[20] Second, given the first point, the expected virtual surplus of uniform pricing p is $\int_{t_1 \geq p} [\phi_1(t_1) f_{\max}(t_1) - c] \, \mathrm{d}t_1$, which by tightness we need to be equal to $(p-c)(1-F_{\max}(p))$. Solving this equation for all p gives $\phi_1(\boldsymbol{t}) = \phi_{\max}(t_1)$.

(b) We obtain α_1 from ϕ_1 by Definition 3.

(c) Given α_1, α_2 is defined to satisfy divergence density equality, $\partial_2\alpha(\boldsymbol{t}) = -f(\boldsymbol{t}) - \partial_1\alpha(\boldsymbol{t})$, and boundary orthogonality at the bottom boundary (i.e., $t_2 = \underline{t}_2(t_1)$). Integrating and employing boundary orthogonality on the bottom boundary of the type space, which requires that $\boldsymbol{\alpha} \cdot \boldsymbol{\eta} = 0$, gives the formula (Figure 6). For example, if $\underline{t}_2(t_1) = 0$, boundary orthogonality requires that $\alpha_2(t_1, 0) = 0$, and thus $\alpha_2(\boldsymbol{t}) = -\int_{y=0}^{t_2} \big(f(t_1, y) + \partial_1\alpha_1(t_1, y)\big) \, \mathrm{d}y$.

(d) We obtain ϕ_2 from α_2 by Definition 3.

For $\boldsymbol{\phi}$ to prove optimality of uniform pricing, we need the allocation of uniform pricing to optimize virtual surplus pointwise with respect to $\boldsymbol{\phi}$. This additional requirement demands that $\phi_1(\boldsymbol{t}) \geq \phi_2(\boldsymbol{t})$ for any type $\boldsymbol{t} \in T$ for which either $\phi_1(\boldsymbol{t})$ or $\phi_2(\boldsymbol{t})$ is positive. A little algebra shows that this condition is implied by the angle of $\boldsymbol{\alpha}(\boldsymbol{t})$ being at most the angle of $\boldsymbol{t}$ with respect to the horizontal t_1 axis, that is, $t_2\alpha_1(\boldsymbol{t}) \leq t_1\alpha_2(\boldsymbol{t})$ (Lemma 8, below). The direction of $\boldsymbol{\alpha}$ corresponds to the paths on which incentive compatibility constraints are considered. Importantly, our approach does not fix the direction and allows any direction that satisfies the above constraint on angles. The following lemma is proved by the divergence theorem, and specifies the direction of $\boldsymbol{\alpha}$.

[20] This argument applies only if $\phi_1(\boldsymbol{t}) > 0$. Nevertheless, we impose the requirement that $\phi_1(\boldsymbol{t}) = \phi_1(t_1)$ everywhere as it allows us to uniquely solve for $\boldsymbol{\phi}$.

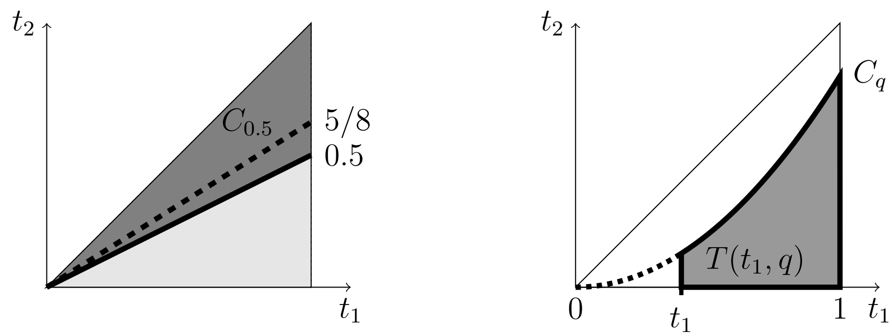


Figure 7: (a) The density in the darker region is twice the density in lighter region. For example, $C_{0.5}(t_1) = 5t_1/8$, meaning given t_1, the probability that $t_2 \leq 5t_1/8$ is 1/2. (b) $T(t_1, q)$ is the set of types below C_q and to the right of t_1. The four curves that define the boundary of $T(t_1, q)$ are $\{$TOP, RIGHT, BOTTOM, LEFT$\}(t_1, q)$. For simplicity the picture assumes T is the triangle defined on (0,0), (1,0), and (1,1).

Definition 5. For any $q \in [0, 1]$, define the *equi-quantile* function $C_q(t_1)$ such that conditioned on t_1, the probability that $t_2 \leq C_q(t_1)$ is equal to q (see Figure 7). More formally, C_q is the upper boundary of T_q, where

$$T_q = \{\boldsymbol{t} | \, \mathbf{Pr}_{\boldsymbol{t}'} \left[t'_2 \leq t_2 | t'_1 = t_1, t'_1 \geq t'_2 \right] := \frac{\int_{t'_2 \leq t_2} f(t_1, t'_2) \mathrm{d}t'_2}{\int_{t'_2 \leq t_1} f(t_1, t'_2) \mathrm{d}t'_2} \leq q\}.$$

For example, notice that for the perfectly correlated class, the equi-quantile curves C_q are identical to C_{cor}.

Lemma 7. *The vector field $\boldsymbol{\phi}$ of Definition 4 is a tight canonical amortization for any uniform pricing. At any $\boldsymbol{t}$, $\boldsymbol{\alpha}(\boldsymbol{t})$ is tangent to the equi-quantile curve crossing $\boldsymbol{t}$.*

Proof. Tightness follows directly from the definition of ϕ_1 (see the justification for Step (a) of the construction). The divergence density equality and bottom boundary orthogonality of $\boldsymbol{\alpha}$ are automatically satisfied by Step (c) of the construction. Orthogonality of the right boundary ($t_1 = \bar{t}_1$) requires that $\boldsymbol{\alpha}(\bar{t}_1, t_2) \cdot (1, 0) = 0$, which is $\alpha_1(\bar{t}_1, t_2) = 0$. This property follows directly from the definitions since $\phi_1(\bar{t}_1, t_2) = \phi_{\max}(\bar{t}_1) = \bar{t}_1$, and therefore $\alpha_1(\bar{t}_1, t_2) = (\bar{t}_1 - \phi_1(\bar{t}_1, t_2)) \, f(\bar{t}_1, t_2) = 0$. At the left boundary, $\boldsymbol{\alpha} \cdot \boldsymbol{\eta} \leq 0$ since $\alpha_1 \geq 0$ from definition and the normal vector is $(-1, 0)$. The only remaining condition, the top boundary orthogonality, is implied by the tangency property of the lemma as follows. The top boundary is C_1. Tangency of $\boldsymbol{\alpha}$ to C_1 implies that $\boldsymbol{\alpha}$ is orthogonal to the normal, which is the top boundary orthogonality requirement. It only remains to prove the tangency property.

The strategy for the proof of the tangency property is as follows. We fix t_1 and q and apply the divergence theorem to $\boldsymbol{\alpha}$ on the subspace of type space to the right of t_1 and below C_q.[21] More formally, divergence theorem is applied to the set of types $T(t_1, q) = \{\boldsymbol{t}' \in T | t'_1 \geq t_1; F(t_2|t_1) \leq q\}$ (see Figure 7). The divergence theorem equates the integral of the orthogonal magnitude of vector field $\boldsymbol{\alpha}$ on the boundary of the subspace to the integral of its divergence within the subspace. As the upper boundary of this subspace is C_q, one term in this equality is the integral of $\boldsymbol{\alpha}(\boldsymbol{t}')$ with the upward orthogonal vector to C_q at $\boldsymbol{t}'$. Differentiating this integral with respect to t_1 gives the desired quantity.

$$\begin{aligned}
&\int_{\boldsymbol{t}' \in \mathrm{TOP}(t_1,q)} \boldsymbol{\eta}(\boldsymbol{t}') \cdot \boldsymbol{\alpha}(\boldsymbol{t}') \, \mathrm{d}\boldsymbol{t}' \\
&\qquad = \int_{\boldsymbol{t}' \in T(t_1,q)} \nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}') \, \mathrm{d}\boldsymbol{t}' - \int_{\boldsymbol{t}' \in \{\mathrm{RIGHT,BOTTOM,LEFT}\}(t_1,q)} \boldsymbol{\eta}(\boldsymbol{t}') \cdot \boldsymbol{\alpha}(\boldsymbol{t}') \, \mathrm{d}\boldsymbol{t}'. \qquad (9)
\end{aligned}$$

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

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

where the last equality followed directly from definition of $T(t_1, q)$. By definition of $\boldsymbol{\alpha}$, and since normal $\boldsymbol{\eta}$ at the left boundary is $(-1, 0)$,

$$\begin{aligned}
\int_{\boldsymbol{t}' \in \{\mathrm{LEFT}\}(t_1,q)} \boldsymbol{\eta}(\boldsymbol{t}') \cdot \boldsymbol{\alpha}(\boldsymbol{t}') \, \mathrm{d}\boldsymbol{t}' &= -\frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \int_{t'_2 \leq C_q(t_1)} f(t_1, t'_2) \, \mathrm{d}t'_2 \\
&= -\frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} q f_{\max}(t_1) \\
&= -(1 - F_{\max}(t_1))q.
\end{aligned}$$

As a result, the right-hand side of equation (9) sums to zero, and we have

$$\int_{\boldsymbol{t}' \in \mathrm{TOP}(t_1,q)} \boldsymbol{\eta}(\boldsymbol{t}') \cdot \boldsymbol{\alpha}(\boldsymbol{t}') \, \mathrm{d}\boldsymbol{t}' = 0.$$

Since the above equation must hold for all t_1 and q, we conclude that $\boldsymbol{\alpha}$ is tangent to the equi-quantile curve at any type. □

[21] The divergence theorem for vector field $\boldsymbol{\alpha}$ is $\int_{\boldsymbol{t} \in T} (\nabla \cdot \boldsymbol{\alpha})(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t} = \int_{\boldsymbol{t} \in \partial T} (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}$.

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

Lemma 8. *The allocation of a uniform pricing mechanism optimizes virtual surplus pointwise with respect to $\boldsymbol{\phi} = \boldsymbol{t} - \boldsymbol{\alpha}/f$ of Definition 4 and any non-negative service cost c if the equi-quantile curves are ratio-monotone and $\phi_1(\boldsymbol{t})$ is monotone non-decreasing in $\boldsymbol{t}$.*

Proof. Tangency of α to the equi-quantile curves (Lemma 7) implies that $\frac{t_2}{t_1}\alpha_1(t_1, t_2) - \alpha_2(t_1, t_2) \leq 0$ if all equi-quantile curves are ratio-monotone. From the assumption $\frac{t_2}{t_1}\alpha_1(t_1, t_2) - \alpha_2(t_1, t_2) \leq 0$ and Definition 4 we have

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

Thus, for $\boldsymbol{t}$ with $\phi_1(\boldsymbol{t}) \geq c$, $\phi_1(\boldsymbol{t}) \geq \phi_2(\boldsymbol{t})$ and pointwise virtual surplus maximization serves the agent alternative 1. Since $\phi_1(\boldsymbol{t})$ is a function only of t_1 (Definition 4), its monotonicity implies that there is a smallest t_1 such that all greater types are served. Also, if $\phi_1(\boldsymbol{t}) \leq c$, again the above calculation implies that $\phi_2(\boldsymbol{t}) \leq c$ and therefore the type is not served. This allocation is the allocation of a uniform pricing. □

Proof of Theorem 6. We show that $\boldsymbol{\phi} = \boldsymbol{t} - \boldsymbol{\alpha}/f$ of Definition 4 is a virtual value function for a uniform pricing and invoke Proposition 2. Lemma 7 showed that $\boldsymbol{\phi}$ is a tight amortization for any uniform pricing.[22] Lemma 8 showed that the allocation of a uniform pricing pointwise maximizes virtual surplus with respect to $\boldsymbol{\phi}$. □

4.2 Perfect Correlations and Necessary Conditions

Consider a simple class of *perfectly correlated* instances where the value t_1 for alternative 1 pins down the value for alternative 2, $t_2 = C_{\text{cor}}(t_1)$. Assume $C_{\text{cor}}(t_1) \leq t_1$, that is, al-

[22]Special attention is needed in case that the left boundary is a singleton, since in that case $f_{\max}(\underline{t}_1)) = 0$ and α_1 is unbounded. In this case our analysis showed that $\boldsymbol{\alpha} \cdot \boldsymbol{\eta} = 0$ everywhere except possibly at $(\underline{t}_1, t_2(\underline{t}_1))$. Divergence theorem states that

$$\int_{\boldsymbol{t} \in \partial T} (\boldsymbol{\alpha} \cdot \boldsymbol{\eta})(\boldsymbol{t}) \,\mathrm{d}\boldsymbol{t} = -\int_{\boldsymbol{t} \in T} f(\boldsymbol{t}) \,\mathrm{d}\boldsymbol{t} = -1, \tag{10}$$

which implies that $\boldsymbol{\alpha} \cdot \boldsymbol{\eta}$ is a negative Dirac delta at $(\underline{t}_1, t_2(\underline{t}_1))$. The integral of $u(\boldsymbol{\alpha} \cdot \boldsymbol{\eta})$ over the boundary is thus $-u(\underline{t}_1, t_2(\underline{t}_1)) = 0$.

ternative 1 is favored to alternative 2 for all types. We say that a curve C_{cor} is *ratio-monotone* if $C_{\text{cor}}(t_1)/t_1$ is monotone increasing in t_1. Let $F_{\max}$ be the distribution of value for alternative 1. A distribution $F_{\max}$ is *regular* if its (canonical) amortization of revenue $\phi_{\max}(t_1) = t_1 - \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)}$ is monotone non-decreasing in t_1 (see the discussion of amortizations of revenue for single-dimensional agents in Section 3). We investigate optimality of uniform pricing for this class by comparing the profit from a uniform price with the profit from other mechanisms (discounted prices for the less favored alternative or distributions over alternatives).[23] The proof is deferred to Appendix B.1.

Theorem 9. *For any value continuous mapping function $C_{\text{cor}}, C_{\text{cor}}(t_1) \leq t_1$ that is not ratio-monotone, there exists a regular distribution $F_{\max}$ such that uniform pricing is not optimal for the perfectly correlated instance jointly defined by $F_{\max}$ and C_{cor}. Furthermore, if $C_{\text{cor}}(t_1)/t_1$ is strictly decreasing everywhere, then uniform pricing is not optimal for* any *distribution $F_{\max}$.*

4.3 Extensions

This section contains extensions of Theorem 6 to $m \geq 2$ outcomes, $n \geq 1$ agents, and distributions where the favorite-outcome projection may not be regular.

First, Theorem 6 can be extended to the case of more than two outcomes and more than one agent. The positive correlation property becomes a sequential positive correlation where the ratio of the value of any outcome to the favorite outcome is positively correlated with the value of favorite outcome, conditioned on the draws of the lower indexed outcomes. A distribution over types $[0,1]^m$ is max-symmetric if the distribution of $v = \max_i t_i$ stays the same conditioned on any outcome having the highest value. For $j \neq i$, define $q_j^i(\boldsymbol{t})$ to be the quantile of the distribution of t_j conditioned on i being the favorite outcome, and conditioned on the values $\boldsymbol{t}_{<j} = (t_1, \ldots, t_{j-1})$ of the lower indexed outcomes. Formally, $q_j^i(\boldsymbol{t}) = \mathbf{Pr}_{\boldsymbol{t}'}[t'_j \leq t_j | \boldsymbol{t}'_{<j} = \boldsymbol{t}_{<j}, t'_i = t_i = \max_k t'_k]$. Define $F(\theta_j | t_i, i, \boldsymbol{q}_{<j}) = \mathbf{Pr}_{\boldsymbol{t}'}[t'_j/t'_i \leq \theta_j | \boldsymbol{q}_{<j} = \boldsymbol{q}^i_{<j}(\boldsymbol{t}'), t'_i = t_i = \max_k t'_k]$ to be the distribution of the value ratio of jth to favorite outcome, conditioned on i being the favorite outcome and given vector $\boldsymbol{q}_{<j}$ of the quantiles of the lower indexed outcomes. In the multi-agent problem with a configurable item, a single item with m configurations is to be allocated to at most one of the agents.[24]

[23] With a uniform price when alternative 1 is favored to alternative 2 for all types, the offer for alternative 2 will not be taken.

[24] We assume that the item has the same possible configurations for each agent. This can be achieved by defining the set of configurations to be the union over the configurations of all agents.

Theorem 10. *A favorite-outcome projection mechanism is optimal for an item with $m \geq 1$ configurations, multiple independent agents, and any service cost $c \geq 0$, if the distribution of each agent is max-symmetric and (a) the favorite-outcome projection has monotone non-decreasing amortization $\phi_{\max}(v) = v - \frac{1-F_{\max}(v)}{f_{\max}(v)}$ and (b) $F(\theta_j|v, i, \boldsymbol{q}_{<j})$ is monotone non-increasing in v for all i, j, θ_j, and $\boldsymbol{q}_{<j}$.*

The proof of the above theorem is in Appendix B.2. From Myerson (1981) we know that if a favorite-outcome projection mechanism is optimal, the optimum mechanism is to allocate the item to the agent with highest $\phi_{\max}(v)$ (no ironing is required as we are assuming regularity), and let the agent choose its favorite configuration. With a single agent, the configurable item setting is identical to the original model with multiple outcomes. The above theorem implies it is optimal to offer a single agent a price for its choice of outcome, generalizing Theorem 6 to $m \geq 2$ outcomes. A special case of the correlation above is when the ratios are independent of each other conditioned on the value of the favorite outcome, that is, each $\theta_j = t_j/v$ for $j \neq i$ is drawn independently of others from a conditional distribution $F(\theta|v, i)$ that is monotone in v.

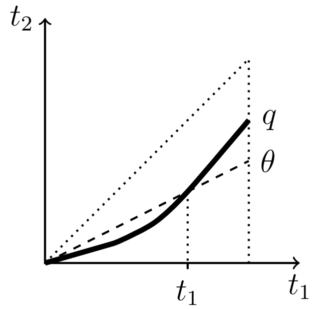
The second extension removes the regularity assumption of Theorem 10 by assuming a slightly stronger correlation assumption, and designs a virtual value function with a simple sweeping procedure in a single dimension (proof in Appendix B.4). In particular, we only iron the canonical amortization $\boldsymbol{\phi}$ along the equi-quantile curves.

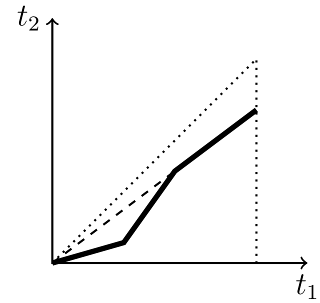
Theorem 11. *A favorite-outcome projection mechanism is optimal for an item with $m = 2$ configurations, multiple independent agents, and any service cost $c \geq 0$, if the distribution of each agent is max-symmetric with convex equi-quantile curves.*

From Myerson (1981), optimality of a favorite-outcome projection mechanism implies optimality of allocating to the agent with highest ironed virtual value. Figure 8 depicts how convexity of equi-quantile curves is stronger than the stochastic dominance requirement of Theorem 6. Convexity states that the line connecting any two points, namely $(0, 0)$ and $(t_1, t_1\theta)$, 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, for any $t'_1 \geq t_1$, $F(\theta|t'_1) \leq F(\theta|t_1)$, and the other direction holds for $t'_1 \leq t_1$ (see Figure 8).

5 Grand Bundle Pricing for Additive Preferences

In single-agent multi-product settings with free disposal (i.e., value for a set of items does not decrease as more items are added), optimality of a favorite-outcome projection mechanism





(a) convex equi-quantile curve(a) non-convex monotone equi-quantile curves

Figure 8: The connection between convexity and ratio-monotonicity of equi-quantile curves. (a) Convexity implies ratio-monotonicity. (b) ratio-monotonicity does not imply convexity.

is equivalent to optimality of posting a single price for the grand bundle of items. Thus, Theorem 10 can be used to obtain conditions for optimality of grand bundle pricing. For example, in the case of two items, when the value for the bundle is v and value for individual items are $v\delta_1$ and $v\delta_2$, Theorem 10 identifies a sufficient positive correlation condition. Note that the theorem does not require any structure on values, such as additivity (value for a bundle is the sum of the values of items in it) or super- or sub-additivity, other than free disposal. If the preference is indeed additive, we have $\delta_2 = 1 - \delta_1$, and Theorem 10 requires that δ_1 be both positively and negatively correlated with v. The only admissible case is independence.[25] In this section we apply the framework of Section 3 to prove optimality of grand bundle pricing for additive preferences, and obtain conditions of optimality that are more permissive than independence by constructing a virtual value function $\bar{\boldsymbol{\phi}}$ from a canonical amortization $\boldsymbol{\phi}$ that is tight for any grand bundle pricing and is constructed to satisfy conditions of Lemma 3. In this section we consider a single agent, $m = 2$ items. As discussed in Section 3.4 and similar to Section 4, we use a class of cost functions to restrict the admissible amortizations. In particular, we assume that the cost of an allocation $\boldsymbol{x} \in [0,1]^2$ is $c(\boldsymbol{x}) = c \max(x_1, x_2)$ for a $c \geq 0$.

Similar to Section 4, we first study a family of instances with perfect correlation to obtain necessary conditions of optimality. In particular, let F_{sum} be a distribution over value s for the bundle (in the case of two items we refer to the grand bundle simply as the bundle), and $\theta(s)$ be the ratio of the value of item 2 to item 1 when value for the bundle is s, that

[25]Let $t_1 = v$ be the value for the bundle, and $t_2 = \delta v$ and $t_3 = (1-\delta)v$ the values for the two items. Let $\delta(q, v)$ be the inverse of the quantile mapping, i.e., $\mathbf{Pr}[\delta \leq \delta(q,v) | v] = q$. Theorem 10 demands that $\delta(q, v)$ be monotone non-decreasing and $F(1 - \delta \leq \theta_2 | v, \delta = \delta(q, v))$ be monotone non-increasing in v for all q, θ_2. The only possible case is independence of v and δ, that is, $\delta(q, v)$ is a constant.

is, value for item 1 is $t_1 = s/(1+\theta)$, and value for item 2 is $t_2 = \theta s/(1+\theta)$.[26] The following theorem shows that if $\theta(s)$ is not monotone non-increasing in s, then bundling is not optimal for some distribution F_{sum}. The proof is similar to Theorem 9 and is omitted.

Theorem 12. *If $\theta(s)$ is not monotone non-increasing in s, then there exists a regular distribution F_{sum} over s such that grand bundle pricing is not optimal for the perfectly correlated instance jointly defined by F_{sum} and $\theta(\cdot)$ and with zero costs.*

The main theorem of this section states sufficient conditions for optimality of pricing the bundle. A symmetric distribution is identified by a marginal distribution F_{sum} of value for the bundle s as well as a conditional distribution $F(\theta|s)$ of the ratio $\theta(\boldsymbol{t}) = \max(t_1, t_2)/\min(t_1, t_2)$ conditioned on value for the bundle s. The main theorem of this section states that regularity of F_{sum} and negative correlation of s and θ in the first order stochastic dominance sense is sufficient for optimality of bundling.

Theorem 13. *For a single agent with additive preferences over two items, bundle pricing is optimal for any costs $c\max(x_1, x_2)$, $c \geq 0$, and any symmetric distribution where (a) F_{sum} has monotone amortization ϕ_{sum} and (b) the conditional distribution $F(\theta|s)$ is monotone non-decreasing in s.*

The following is an example class of distributions satisfying the conditions of Theorem 13. Draw the value for the bundle s from a regular distribution F_{sum}, and value for the items t_1 and t_2 uniformly such that $t_1 + t_2 = s$, $\max(t_1, t_2)/\min(t_1, t_2) \geq \theta(s)$, for any monotone non-increasing function $\theta(s)$ (see Figure 9).

References

Adams, W. J. and Yellen, J. L. (1976). Commodity bundling and the burden of monopoly. *The quarterly journal of economics*, pages 475–498.

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.

[26]Because of the additivity structure imposed on preferences, two parameters are sufficient to define values for three outcomes. For example, t_1 and t_2 define the value for the bundle $s = t_1 + t_2$. Alternatively, s and θ define the value for individual items.

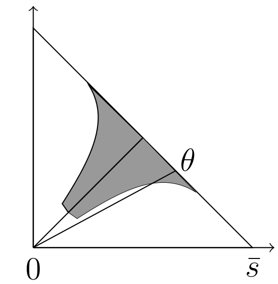


Figure 9: The conditional distribution $F(\theta|s)$ is monotone for a monotone non-increasing $\theta(s)$ where conditioned on s, the values are uniform from the set $\{\boldsymbol{t}|t_1 + t_2 = s, \min(t_1, t_2)/\max(t_1, t_2) \geq \theta(s)\}$. For example, for any $\delta \leq \bar{s}/2$, setting $\theta(s) = \delta(1+s)/s$ defines the set of types to be the triangle $t_1, t_2 \in [\delta, \bar{s}-\delta]$, $t_1 + t_2 \leq \bar{s}$.

Borodin, A. and El-Yaniv, R. (1998). *Online Computation and Competitive Analysis*. Cambridge University Press, New York, NY, USA.

Castro, L. I. d. (2007). Affiliation, equilibrium existence and the revenue ranking of auctions.

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.

Deneckere, R. J. and McAfee, P. (1996). Damaged goods. *Journal of Economics & Management Strategy*, 5(2):149–174.

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.

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.

Milgrom, P. R. and Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica: Journal of the Econometric Society*, pages 1089–1122.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

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.

Rochet, J.-C. and Chone, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, 66(4):783–826.

Rochet, J.-C. and Stole, L. A. (2003). The economics of multidimensional screening. *Econometric Society Monographs*, 35:150–197.

Stokey, N. L. (1979). Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371.

Thanassoulis, J. (2004). Haggling 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*.

Wilson, R. B. (1993). *Nonlinear pricing*. Oxford University Press.

A Comparison with Theorem 2 of Rochet and Chone (1998)

This section compares Lemma 3 with Theorem 2 of Rochet and Chone (1998), which considers the optimization problem over mechanisms that satisfy the envelope equality $\nabla u(\boldsymbol{t}) = \boldsymbol{x}(\boldsymbol{t})$, but the convexity condition on the utility function is relaxed (seeLemma 1). In this section we refer to this problem as *the relaxed problem.* Recall from Section 3 (Proposition 2) that a mechanism $(\boldsymbol{x}, p)$ is optimal if there exists a tight canonical amortization $\boldsymbol{\phi}$ (satisfying divergence density and boundary inflow of Lemma 3) such that the allocation function $\boldsymbol{x}$ optimizes virtual surplus pointwise, that is,

$$\boldsymbol{x}(\boldsymbol{t}) \in \arg\max_{\hat{\boldsymbol{x}} \in X} \hat{\boldsymbol{x}} \cdot \boldsymbol{\phi}(\boldsymbol{t}) - c(\hat{\boldsymbol{x}}), \forall \boldsymbol{t}. \tag{11}$$

Now assume that the cost function c is strictly convex. As a result, for any $\boldsymbol{t}$, $\boldsymbol{x}(\boldsymbol{t})$ is the virtual surplus maximizer if and only if it satisfies the first order conditions of optimality of (11), that is

$$\boldsymbol{\phi}(\boldsymbol{t}) - \nabla c(\boldsymbol{x}(\boldsymbol{t})) = 0.$$

We conclude that a mechanism $(\boldsymbol{x}, p)$ is the optimal mechanism if $\boldsymbol{\phi}$ defined as $\boldsymbol{\phi}(\boldsymbol{t}) = \nabla c(\boldsymbol{x}(\boldsymbol{t}))$ is a canonical amortization of revenue, and is tight for $(\boldsymbol{x}, p)$. To make comparison easier, let us summarize the above discussion as a corollary of Lemma 3 and Proposition 2.

Corollary 14. *If the cost function c is strictly convex, a mechanism $(\boldsymbol{x}, p)$ with utility function u is the optimal solution to the relaxed problem if $\boldsymbol{\alpha}(\boldsymbol{t}) := (\boldsymbol{t} + \nabla c(\boldsymbol{x}(\boldsymbol{t})))f(\boldsymbol{t})$ satisfies the divergence density and boundary inflow conditions and $\boldsymbol{\phi}$ is tight for the mechanism, that is,*

- $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) = -f(\boldsymbol{t})$ *for all types* $\boldsymbol{t} \in T$.
- $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) \leq 0$ *for all types* $\boldsymbol{t} \in \partial T$, *with* $u(\boldsymbol{t}) = 0$ *if* $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) < 0$.

Cast in our setting, Theorem 2 of Rochet and Chone (1998) states that if the cost function c is *strictly convex* (and some extra assumptions), then a mechanism $(\boldsymbol{x}, p)$ is the optimal solution to the relaxed problem *if and only if* $\boldsymbol{\alpha}(\boldsymbol{t}) := (\boldsymbol{t} + \nabla c(\boldsymbol{t}))f(\boldsymbol{t})$ satisfies

- $\nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t}) \geq -f(\boldsymbol{t})$ for all types $\boldsymbol{t} \in T$, with equality if $u(\boldsymbol{t}) > 0$.

- $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) \leq 0$ for all types $\boldsymbol{t} \in \partial T$, with $u(\boldsymbol{t}) = 0$ if $\boldsymbol{\alpha}(\boldsymbol{t}) \cdot \boldsymbol{\eta}(\boldsymbol{t}) < 0$.

Note the differences with Corollary 14: a) Our analysis applies to dot-product utility functions whereas Theorem 2 of Rochet and Chone (1998) applies to settings with non-linear utilites as in this sense is more general. b) The conditions of optimality in Theorem 2 of Rochet and Chone (1998) are more permissive and are *necessary* and sufficient for optimality, whereas the conditions of Corollary 14 are only sufficient but not necessary. c) Theorem 2 of Rochet and Chone (1998) *requires strict convexity* of the cost function so that the relaxed problem admits a unique local optima, whereas Lemma 3 makes no such assumption on the cost function and in this sense is more general. In particular, since our main theorems apply to linear cost functions, Theorem 2 of Rochet and Chone (1998) is not applicable.

B Proofs from Section 4

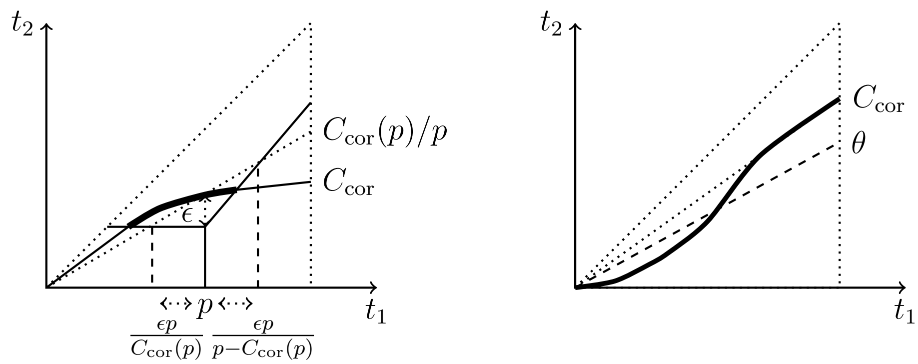
This section includes proofs from Section 4.

B.1 Proof of Theorem 9

Theorem 9. For any value continuous mapping function $C_{\text{cor}}, C_{\text{cor}}(t_1) \leq t_1$ that is not ratio-monotone, there exists a regular distribution $F_{\max}$ such that uniform pricing is not optimal for the perfectly correlated instance jointly defined by $F_{\max}$ and C_{cor}. Furthermore, if $C_{\text{cor}}(t_1)/t_1$ is strictly decreasing everywhere, then uniform pricing is not optimal for *any* distribution $F_{\max}$.

Proof. Let the cost $c = 0$. At the core of the proof of both statements is the analysis of revenue change as a result of offering a small discount for alternative 2. In particular, consider the change in revenue as a result of supplementing a price p for the alternative 1 with a price $C_{\text{cor}}(p) - \epsilon$ for alternative 2. Assume that $C_{\text{cor}}(t_1)/t_1$ is strictly decreasing at p. The results of this change are twofold (Figure 10): On one hand, a set of types with value slightly less than p for alternative 1 will pay $C_{\text{cor}}(p) - \epsilon$ for this new discounted offer. Since $C_{\text{cor}}(t_1)/t_1$ is strictly decreasing at p and C_{cor} is continuous, this set lies above the ray connecting $(0, 0)$ to $C_{\text{cor}}(p)/p$. As a result, the measure of this set of types is at least (see Figure 10)

$$\int_{t_1 = p - \frac{\epsilon p}{C_{\text{cor}}(p)}}^{p} f_{\max}(t_1) \, \mathrm{d}t_1.$$



(a) Discount price for alternative 2 (b) ratio-monotonicity

Figure 10: (a) As a result of adding an offer with price $C_{\rm cor}(p) - \epsilon$ for alternative 2 to the existing offer of price p for alternative 1, the types in darker shaded part of curve $C_{\rm cor}$ will change decisions and contribute to a change in revenue. The lengths of the projected intervals on the t_1 axis of the types contributing to loss and gain in revenue are lower- and upper-bounded by $\frac{\epsilon p}{C_{\rm cor}(p)}$ and $\frac{\epsilon p}{p-C_{\rm cor}(p)}$, respectively. (b) For any θ, t_1, $F(\theta|t_1, 1) = 1$ if $C_{\rm cor}(t_1)/t_1 \leq \theta$, and $F(\theta|t_1, 1) = 0$ otherwise. Therefore, ratio-monotonicity is equivalent to monotonicity of $F(\theta|t_1, 1)$ in t_1.

Therefore, the positive effect is at least

$$\Delta_+(\epsilon) := (C_{\rm cor}(p) - \epsilon) \times \int_{t_1 = p - \frac{\epsilon p}{C_{\rm cor}(p)}}^{p} f_{\max}(t_1)\ \mathrm{d}t_1.$$

On the other hand, a set of types with value slightly higher than p for alternative 1 will change their decision from selecting alternative 1 to alternative 2. The measure of this set of types is at most (see Figure 10)

$$\int_{t_1 = p}^{p + \frac{\epsilon p}{p - C_{\rm cor}(p)}} f_{\max}(t_1)\ \mathrm{d}t_1.$$

Therefore, the negative effect is at most

$$\Delta_-(\epsilon) := (p - C_{\rm cor}(p) + \epsilon) \times \int_{t_1 = p}^{p + \frac{\epsilon p}{p - C_{\rm cor}(p)}} f_{\max}(t_1)\ \mathrm{d}t_1.$$

Note that $\Delta_+(0) = \Delta_-(0) = 0$. We now show that the positive effect of adding the discount is strictly more than its negative effect for small enough ϵ by showing that $\Delta'_+(0) \geq \Delta'_-(0)$ (since $C_{\rm cor}(t_1)/t_1$ is strictly decreasing at p, either the positive effect is strictly more than

$\Delta'_+(\epsilon)$, or the negative effect is strictly less than $\Delta'_-(0)$). Note that

$$\Delta'_+(\epsilon) = -\int_{t_1=p-\frac{\epsilon p}{C_{\mathrm{cor}}(p)}}^{p} f_{\max}(t_1)\;\mathrm{d}t_1 + (C_{\mathrm{cor}}(p)-\epsilon)\times\frac{p}{C_{\mathrm{cor}}(p)} f_{\max}(p-\frac{\epsilon p}{C_{\mathrm{cor}}(p)}),$$

and therefore,

$$\Delta'_+(0) = p f_{\max}(p).$$

Similar analysis shows that $\Delta'_-(0) = p f_{\max}(p)$. It follows that offering a discount for the less favored alternative strictly improves revenue for small enough ϵ.

We will now complete the proof of both statements. For the first statement, consider p where $C_{\mathrm{cor}}(t_1)/t_1$ is decreasing at $t_1 = p$, and any regular distribution $F_{\max}$ such that p maximizes $p(1-F_{\max}(p))$. The above argument shows that the revenue of the optimum uniform price p is strictly less than an alternative mechanism that offers alternative 2 for a small discount. For the second statement, consider the optimum uniform price p. Since $C_{\mathrm{cor}}(t_1)/t_1$ is strictly decreasing everywhere, it is strictly decreasing at p, and the above argument shows that the uniform price p is not the optimal mechanism. □

B.2 Proof of Theorem 10

Theorem 10. A favorite-outcome projection mechanism is optimal for an item with $m \geq 1$ configurations, multiple independent agents, and any service cost $c \geq 0$, if the distribution of each agent is max-symmetric and (a) the favorite outcome projection has monotone non-decreasing amortization $\phi_{\max}(v) = v - \frac{1-F_{\max}(v)}{f_{\max}(v)}$ and (b) $F(\theta_j|v,i,\boldsymbol{q}_{<j})$ is monotone non-increasing in v for all i, j, θ_j, and $\boldsymbol{q}_{<j}$.

Proof. The construction extends the construction of Theorem 6. Let outcome 1 be the favorite outcome. For $\boldsymbol{q}$, let $C^{\boldsymbol{q}}(t_1)$ be a function that maps t_1 to $(t_2,\ldots,t_m)$ such that $\boldsymbol{q}(\boldsymbol{t}) = \boldsymbol{q}$. Define $\boldsymbol{\alpha}$ by integrating by parts along the curves $C^{\boldsymbol{q}}(t_1)$. This defines $\alpha_1(\boldsymbol{t}) = \frac{1-F_{\max}(t_1)}{f_{\max}(t_1)} f(\boldsymbol{t})$, and $\alpha_i(\boldsymbol{t}) = \alpha_1(\boldsymbol{t})\partial_{t_1} C_i^{\boldsymbol{q}}(t_1)$. The assumptions of the theorem also implies that $\alpha_i(\boldsymbol{t}) - (t_i/t_1)\alpha_1(\boldsymbol{t}) \leq 0$. As a result, $\phi_i(\boldsymbol{t}) \leq (t_i/t_1)\phi_1(\boldsymbol{t})$.

With multiple agents, $m \geq 1$, and uniform service cost c, ex-post optimization of virtual surplus allocates the agent with the highest positive virtual value. The argument above shows that the highest positive virtual value of any agent corresponds to the favorite outcome of that agent, and is equal to the virtual value of the single-dimensional projection. □

B.3 Product Distributions Over Values

In this section we derive conditions that prove optimality of the single-dimensional projection for product distributions over values.

Theorem 15. *Uniform pricing is optimal for any cost c for an instance with two outcomes where the value for each outcome is drawn independently from a distribution with density proportional to $e^{h(\log(x))}$.*

We will show that the distribution satisfies the conditions of Theorem 10. In order to show that $F(\theta|v)$ is monotone in v, we show that the joint distribution of θ and v satisfies the stronger property of affiliation. That is,

$$f^{MR}(t_1, \theta) \times f^{MR}(t_1', \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t_1', \theta), \qquad \forall t_1 \leq t_1', \theta \leq \theta',$$

where $f^{MR}(t_1, \theta) = f(t_1, t_1\theta)$ is the joint distribution of t_1 and v. 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, $t\theta \leq t'\theta, t\theta' \leq t'\theta'$. Since $f(x) = \eta \cdot e^{h(\log(x))}$,

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

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

$$f^{MR}(t_1, \theta) \times f^{MR}(t_1', \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t_1', \theta).$$

To complete the proof, we need to show that $F_{\max}$ is regular. This is the case because $f_{\max}(v) = F(v)f(v)$, $f(v) = \eta \cdot e^{h(\log(v))}$ is monotone in v by monotonicity of h.

B.4 Proof of Theorem 11

Theorem 11. A favorite-outcome projection mechanism is optimal for an item with $m = 2$ configurations, multiple independent agents, and any service cost $c \geq 0$, if the distribution of each agent is max-symmetric with convex equi-quantile curves.

We will design a virtual value function $\bar{\boldsymbol{\phi}}$ from the canonical amortization $\boldsymbol{\phi}$ satisfying conditions of Lemma 3. Importantly, $\bar{\boldsymbol{\phi}}$ 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 amortization. The second component $\bar{\phi}_2$ is then defined such that the expected virtual surplus with respect $\bar{\boldsymbol{\phi}}$ 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 $\bar{\boldsymbol{\phi}}-\boldsymbol{\phi}$ may have for $\bar{\boldsymbol{\phi}}$ 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 $\bar{\boldsymbol{\phi}}$ 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

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

where $f_{\max}(t_1) = \int_0^{t_1} f(t_1, t'_2)\, \mathrm{d}t'_2$ to the probability that a random draw $\boldsymbol{t}'$ 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

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

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

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

where $\boldsymbol{x}^Q$ and $\boldsymbol{\phi}^Q$ are representations of $\boldsymbol{x}$ and $\boldsymbol{\phi}$ in quantile space. In particular, $\phi_1^Q(\boldsymbol{q}) =$

$\phi_{\max}(t_1(q_1))$ might not be monotone in q_1. In what follows we design the amortization $\bar{\boldsymbol{\phi}}^Q$ using $\boldsymbol{\phi}^Q$.

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

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

Similarly, for any such allocation,

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

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 $\bar{\boldsymbol{\phi}}^Q$ upper bounds revenue for any allocation that satisfies $x_1^Q(\boldsymbol{q}) = x_1^Q(q_1)$, and $x_2^Q(\boldsymbol{q}) = 0$, with equality for the allocation that optimizes $\bar{\boldsymbol{\phi}}^Q$ pointwise.

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

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

Using integration by parts we can write

$$\int\int \boldsymbol{x}^Q(\boldsymbol{q}) \cdot (\bar{\boldsymbol{\phi}}^Q - \boldsymbol{\phi}^Q)(\boldsymbol{q}) \,\mathrm{d}\boldsymbol{q} = \int_{q_2}\int_{q_1} \frac{\mathrm{d}}{\mathrm{d}q_1}\boldsymbol{x}^Q(\boldsymbol{q}) \cdot \int_{q_1' \geq q_1} (\bar{\boldsymbol{\phi}}^Q - \boldsymbol{\phi}^Q)(q_1', q_2)\,\mathrm{d}q_1'\,\mathrm{d}q_1\,\mathrm{d}q_2.$$

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 1). In particular this must be true for the tangent vector to equi-quantile curve parameterized by q_2. Thus incentive compatibility of $\boldsymbol{x}$ implies that the above expression is positive if the vector that is multiplied by $\frac{\mathrm{d}}{\mathrm{d}q_1}\boldsymbol{x}^Q(\boldsymbol{q})$ 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

$q_1' = q_1$,

$$\frac{\int_{q_1' \ge q_1} (\bar{\phi}_2^Q - \phi_2^Q)(q_1', q_2)\, \mathrm{d}q_1'}{\int_{q_1' \ge q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q_1', q_2)\, \mathrm{d}q_1'} = \frac{\frac{\mathrm{d}}{\mathrm{d}q_1} t_2(\boldsymbol{q})}{\frac{\mathrm{d}}{\mathrm{d}q_1} t_1(\boldsymbol{q})}.$$

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

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

As a result, $\bar{\boldsymbol{\phi}}^Q$ defined above is a tight amortization if its optimization indeed gives uniform pricing. The next lemma formally states the above discussion.

Lemma 16. *The virtual surplus, with respect to $\bar{\boldsymbol{\phi}}^Q$ of any incentive compatible allocation $\boldsymbol{x}$ upper bounds its revenue. If x_1 is only a function of q_1 (equivalently t_1), $x_1'(q_1) = 0$ whenever $\int_{q_1' \ge q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q_1')\, \mathrm{d}q_1' > 0$, and $x_2(\boldsymbol{q}) = 0$ for all $\boldsymbol{q}$, the expected virtual surplus with respect to $\bar{\boldsymbol{\phi}}^Q$ equals revenue.*

We will finally need to verify that $\bar{\boldsymbol{\phi}}^Q$ also satisfies the properties required for ex-post optimization. Lemma 18 below identifies convexity of equi-quantile curves as a sufficient condition. The proof requires the following technical lemma.

Lemma 17. *The amortization $\bar{\phi}$ satisfies $\bar{\phi}_1(\boldsymbol{t}) \le t_1$.*

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

$$\begin{aligned}
\bar{\phi}_1^Q(q_1'') &= \frac{q_1' t_1'(q_1') - q_1 t_1(q_1)}{q_1' - q_1} \\
&\le \frac{q_1' t_1(q_1') - q_1 t_1(q_1')}{q_1' - q_1} \\
&= t_1(q_1') \le t_1(q_1'').
\end{aligned}$$

$\square$

Lemma 18. *If the equi-quantile curves are convex for all q_2, the amortization $\bar{\boldsymbol{\phi}}^Q$ defined above satisfies $\theta(\boldsymbol{q}) \bar{\phi}_1^Q(\boldsymbol{q}) \ge \bar{\phi}_2^Q(\boldsymbol{q})$. As a result, $\bar{\phi}_1^Q \ge \bar{\phi}_2^Q$ whenever either is positive.*

Proof. Lemma 7 showed that $\boldsymbol{\alpha}$ is tangent to the equi-quantile curves. This implies that $\phi_1^Q(\boldsymbol{q})\mu(\boldsymbol{q}) - \phi_2^Q(\boldsymbol{q}) = t_1(\boldsymbol{q})\mu(\boldsymbol{q}) - t_2(\boldsymbol{q})$. By rearranging the definition of ϕ_2 we get

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

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

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

where the inequality followed since convexity of equi-quantile curves imply that $\mu(\boldsymbol{q}) \geq \theta(\boldsymbol{q})$, and by Lemma 17, $t_1(\boldsymbol{q}) - \bar{\phi}_1^Q(\boldsymbol{q}) \geq 0$.

We can now use the above inequality to write

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

□

Proof of Theorem 11. Combining Lemma 16 and Lemma 18 proves the theorem. □

C Proof of Theorem 13

This section contains the proof of Theorem 13.

Similar to Section 4, it is sufficient to prove the statement assuming $t_1 \geq t_2$. As in Section 4 the sum-of-values projection, via the divergence density equality (of Lemma 3), pins down an amortization $\boldsymbol{\phi}$ that is tight for any grand bundle pricing. This tight amortization may fail to be a virtual value function because virtual surplus with respect to $\boldsymbol{\phi}$ is not pointwise optimized by a grand bundle pricing. For this reason, we directly define $\bar{\boldsymbol{\phi}}$ and then prove that it is a virtual value function for the grand bundle pricing mechanism by comparing the virtual surplus with respect to $\bar{\boldsymbol{\phi}}$ and $\boldsymbol{\phi}$.

Definition 6. The *two-dimensional extension* $\bar{\boldsymbol{\phi}}$ of the amortization of the sum-of-values projection $\phi_{\text{sum}}(s) = s - \frac{1-F_{\text{sum}}(s)}{f_{\text{sum}}(s)}$ is:

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

The following lemma provides conditions on vector field $\bar{\boldsymbol{\phi}}$ such that bundle pricing maximizes virtual surplus pointwise with respect to $\bar{\boldsymbol{\phi}}$. These conditions are satisfied for $\bar{\boldsymbol{\phi}}$ of Definition 6, if $\phi_{\text{sum}}(s)$ is monotone non-decreasing.

Lemma 19. *The allocation of a bundle pricing mechanism pointwise optimizes virtual surplus with respect to vector field $\bar{\boldsymbol{\phi}}$ for all costs $c\max(x_1,x_2)$ if and only if: $\bar{\phi}_1(\boldsymbol{t})$ and $\bar{\phi}_2(\boldsymbol{t})$ have the same sign, $\bar{\phi}_1(\boldsymbol{t})+\bar{\phi}_2(\boldsymbol{t})$ is only a function of t_1+t_2 and is monotone non-decreasing in t_1+t_2.*

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

Given Lemma 19, the remaining steps in proving that $\bar{\boldsymbol{\phi}}$ is a virtual value function is showing that it is a tight amortization for grand bundle pricing. The following lemma proves tightness.

Lemma 20. *The expected revenue of a bundle pricing is equal to its expected virtual surplus with respect to the two-dimensional extension $\bar{\boldsymbol{\phi}}$ of the sum-of-values projection (Definition 6).*

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

$$\begin{aligned}\int_{\boldsymbol{t}\in T}(\boldsymbol{x}^p\cdot\boldsymbol{\phi}f)(\boldsymbol{t})\ \mathrm{d}\boldsymbol{t} &= \int_{\boldsymbol{t}\in T}\boldsymbol{x}^p(t_1,t_2)\cdot\boldsymbol{\phi}(t_1,t_2)f(t_1,t_2)\ \mathrm{d}\boldsymbol{t}\\ &= \int_{\boldsymbol{t}\in T, t_1+t_2\geq p}\phi_{sum}(t_1+t_2)f(t_1,t_2)\ \mathrm{d}\boldsymbol{t}.\\ &= -\int_{s\geq p}\frac{\mathrm{d}}{\mathrm{d}s}(s(1-F_{sum}(s))\ \mathrm{d}s\\ &= R(p)-R(1)=R(p).\end{aligned}$$

□

The rest of this section shows that $\bar{\boldsymbol{\phi}}$ provides an upper bound on revenue of any mechanism. For that, we study the existence of a tight canonical amortization $\boldsymbol{\phi}$ such that the virtual surplus of any incentive compatible mechanism with respect to $\bar{\boldsymbol{\phi}}$ upper bounds its virtual surplus with respect to $\boldsymbol{\phi}$ (any such $\boldsymbol{\phi}$ must be tight for any bundle pricing since $\bar{\boldsymbol{\phi}}$ is) and invoke Proposition 5. Define the *equi-quantile* function $C_q(s)$ such that conditioned on s, the probability that $t_2 \leq C_q(s)$ is equal to q.

Lemma 21. *If the conditional distribution $F(\theta|s)$ is monotone non-decreasing in s, then there exists a canonical amortization $\boldsymbol{\phi}(\boldsymbol{t}) = \boldsymbol{t}-\boldsymbol{\alpha}(\boldsymbol{t})/f(\boldsymbol{t})$ such that $\mathbf{E}[\boldsymbol{x}(\boldsymbol{t})\cdot(\bar{\boldsymbol{\phi}}(\boldsymbol{t})-\boldsymbol{\phi}(\boldsymbol{t}))] \geq 0$ for all incentive compatible mechanisms. For any $\boldsymbol{t}$, $\boldsymbol{\alpha}(\boldsymbol{t})$ is tangent to the equi-quantile curve crossing $\boldsymbol{t}$.*

We show the following refinement of Proposition 5, for any incentive compatible allocation $\boldsymbol{x}$ and sum s,

$$\mathbf{E}\left[\boldsymbol{x}(\boldsymbol{t})\cdot(\bar{\boldsymbol{\phi}}(\boldsymbol{t})-\boldsymbol{\phi}(\boldsymbol{t}))\ \mid\ t_1+t_2=s\right]\geq 0. \tag{12}$$

That is, we use a sweeping process in a single dimension and along lines with constant sum of values s (see Section 3.4). Consider the amortization $\boldsymbol{\phi}$ that, like $\bar{\boldsymbol{\phi}}$, sets $\phi_1(\boldsymbol{t})+\phi_2(\boldsymbol{t}) = \phi_{\mathrm{sum}}(t_1+t_2)$ but, unlike $\bar{\boldsymbol{\phi}}$, splits this total amortized value across the two coordinates to satisfy the divergence density equality. Equation (12) can be expressed in terms of this relative difference $\bar{\phi}_1-\phi_1$ since $\boldsymbol{x}\cdot(\bar{\boldsymbol{\phi}}-\boldsymbol{\phi}) = (x_1-x_2)(\bar{\phi}_1-\phi_1)$. We will first show that to satisfy equation (12) for all incentive compatible $\boldsymbol{x}$ it is sufficient for $\boldsymbol{\phi}$, relative to $\bar{\boldsymbol{\phi}}$, to place less value on the favorite coordinate, i.e., $\phi_1 \leq \bar{\phi}_1$. Notice that since $\phi_1+\phi_2 = \bar{\phi}_1+\bar{\phi}_2$ and $\bar{\phi}_1\frac{t_2}{t_1} = \bar{\phi}_2$, the condition $\phi_1 \leq \bar{\phi}_1$ is equivalent to the condition $\phi_1\frac{t_2}{t_1} \leq \phi_2$.

To calculate the expectation in equation (12), 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

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

Our derivation of sufficient conditions for the two-dimensional extension of the sum-of-values projection to be an amortization exploits two properties. First, by convexity of utility (Lemma 1), the change in allocation probabilities of an incentive compatible mechanism, for a fixed sum s as the ratio θ increases, can not be more for coordinate one than coordinate two, that is, $x_1^{SR}(s,\theta) - x_2^{SR}(s,\theta)$ must be non-increasing in θ (Lemma 22). Second, if $\boldsymbol{\phi}$ shifts value from coordinate one to coordinate two relative to the vector field $\bar{\boldsymbol{\phi}}$, then, it also shifts expected value from coordinate one to coordinate two, conditioned on sum $t_1 + t_2 = s$ and ratio $t_2/t_1 \leq \theta$. We then use integration by parts to show that the shift in expected value only hurts the virtual surplus of $\boldsymbol{\phi}$ relative to $\bar{\boldsymbol{\phi}}$ and equation (12) is satisfied (by Lemma 23). Later in the section we will describe sufficient conditions on the distribution to guarantee existence of $\boldsymbol{\phi}$ where this sufficient condition that $\phi_1 \frac{t_2}{t_1} \leq \phi_2$, is satisfied (Lemma 24).

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

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

Proof. The proof follows directly from Lemma 1. In particular, convexity of the utility function implies that the dot product of any vector, here $(-1, 1)$, and the change in gradient of utility $\boldsymbol{x}$ in the direction of that vector, here $\frac{d}{d\theta}\boldsymbol{x}^{SR}(s,\theta)$, is positive. $\square$

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

Proof. Without loss of generality, in proving equation (12) we can assume that the allocation is symmetric. This is because by symmetry of the distribution, there exists an optimal mechanism that is also symmetric. Therefore, it is sufficient 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).[27]

[27] In general, when optimal mechanisms are known to satisfy a certain property, the inequality of amortization needs to be shown only for mechanisms satisfying that property.

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

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

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

(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 $\bar{\boldsymbol{\phi}}$ relative to $\boldsymbol{\phi}$. This follows from the fact that $\phi_1(\boldsymbol{t}) + \phi_2(\boldsymbol{t}) = \bar{\phi}_1(\boldsymbol{t}) + \bar{\phi}_2(\boldsymbol{t}) = \phi_{\text{sum}}(t_1 + t_2)$.

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

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

Write the left-hand side of equation (12) as:

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

Substituting $\boldsymbol{\Gamma}$ into the integral above, we have

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

The second equality is integration by parts. The third equality follows because the first term on the left-hand side is zero: For $\theta = 0$, $\boldsymbol{\Gamma}(s,\theta) = \boldsymbol{0}$ 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 $-\frac{d}{d\theta}\boldsymbol{x}^{SR}(s,\theta)\cdot(1,-1) \ge 0$ (Lemma 22) and properties (a) and (b). $\square$

To identify sufficient conditions for $\bar{\boldsymbol{\phi}}$ to be an amortization it now suffices to derive conditions under which there exists a canonical amortization $\boldsymbol{\phi}$ satisfying $\phi_1(\boldsymbol{t}) + \phi_2(\boldsymbol{t}) = \phi_{\text{sum}}(t_1 + t_2)$ and the condition of Lemma 23, i.e., $\phi_1(\boldsymbol{t})\frac{t_2}{t_1} \leq \phi_2(\boldsymbol{t})$. 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

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

Thus, it suffices to identify conditions under which $\alpha_1 \frac{t_2}{t_1} \geq \alpha_2$.

The following constructs the canonical amortization $\boldsymbol{\phi}$ and specifies the direction of $\boldsymbol{\alpha}$. Similar to Section 4, $\boldsymbol{\alpha}$ is tangent to the equi-quantile curve, that in the section are defined by conditioning on the value for bundle s. The proof is similar to the proof of Lemma 7.

Lemma 24. *A canonical amortization $\boldsymbol{\phi} = \boldsymbol{t} - \boldsymbol{\alpha}/f$ satisfying $\phi_1(\boldsymbol{t}) + \phi_2(\boldsymbol{t}) = \phi_{\text{sum}}(t_1 + t_2)$ exists and is unique, where $\boldsymbol{\alpha}(\boldsymbol{t})$ is tangent to the equi-quantile curve crossing $\boldsymbol{t}$.*

Proof. We assume that $\boldsymbol{\phi}$ satisfying the requirements of the lemma exists, derive the closed form suggested in the lemma, and then verify that the derived $\boldsymbol{\phi}$ indeed satisfies all the required properties. We fix s and q and apply the divergence theorem to $\boldsymbol{\alpha}$ on the subspace of type space to the right of $t_1 + t_2 = s$ and below C_q. More formally, divergence theorem is applied to the set of types $T(s,q) = \{\boldsymbol{t'} \in T | t'_1 + t'_2 \geq s; F(t_2|s) \leq q\}$. The divergence theorem equates the integral of the orthogonal magnitude of vector field $\boldsymbol{\alpha}$ on the boundary of the subspace to the integral of its divergence within the subspace. As the upper boundary of this subspace is C_q, one term in this equality is the integral of $\boldsymbol{\alpha}(\boldsymbol{t'})$ with the upward orthogonal vector to C_q at $\boldsymbol{t'}$. Differentiating this integral with respect to t_1 gives the desired quantity.

$$\begin{aligned}
&\int_{\boldsymbol{t'} \in \text{TOP}(s,q)} \boldsymbol{\eta}(\boldsymbol{t'}) \cdot \boldsymbol{\alpha}(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'} \\
&\qquad = \int_{\boldsymbol{t'} \in T(s,q)} \nabla \cdot \boldsymbol{\alpha}(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'} - \int_{\boldsymbol{t'} \in \{\text{RIGHT,BOTTOM,LEFT}\}(s,q)} \boldsymbol{\eta}(\boldsymbol{t'}) \cdot \boldsymbol{\alpha}(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'}.
\end{aligned} \tag{13}$$

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

$$\begin{aligned}
&= -\int_{\boldsymbol{t'} \in T(s,q)} f(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'} - \int_{\boldsymbol{t'} \in \{\text{LEFT}\}(s,q)} \boldsymbol{\eta}(\boldsymbol{t'}) \cdot \boldsymbol{\alpha}(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'} \\
&= -q(1 - F_{\text{sum}}(s)) - \int_{\boldsymbol{t'} \in \{\text{LEFT}\}(\boldsymbol{q})} \boldsymbol{\eta}(\boldsymbol{t'}) \cdot \boldsymbol{\alpha}(\boldsymbol{t'}) \, \mathrm{d}\boldsymbol{t'}
\end{aligned}$$

where the last equality followed directly from definition of $T(s,q)$. By definition of $\boldsymbol{\alpha}$, and since normal $\boldsymbol{\eta}$ at the left boundary is $(-1,-1)$,

$$\begin{aligned}\int_{\boldsymbol{t}'\in\{\text{LEFT}\}(s,q)} \boldsymbol{\eta}(\boldsymbol{t}')\cdot\boldsymbol{\alpha}(\boldsymbol{t}')\,\mathrm{d}\boldsymbol{t}' &= -\frac{1-F_{\text{sum}}(s)}{f_{\text{sum}}(s)}\int_{t_2'\le C_q(t_1)} f(t_1,t_2')\,\mathrm{d}t_2' \\ &= -\frac{1-F_{\text{sum}}(s)}{f_{\text{sum}}(s)} q f_{\text{sum}}(s) \\ &= -(1-F_{\text{sum}}(s))q\end{aligned}$$

As a result, the right hand side of equation (13) sums to zero, and we have

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

Since the above equation must hold for all s and q, we conclude that $\boldsymbol{\alpha}$ is tangent to the equi-quantile curve at any type.

□

We now complete the proof of Theorem 13.

Proof of Lemma 21. The assumption that $F(\theta|s)$ is monotone implies that the equi-quantile curves are ratio-monotone. The tangency property of Lemma 24 implies that $\alpha_1\frac{t_2}{t_1}\ge\alpha_2$ and subsequently $\phi_1(\boldsymbol{t})\frac{t_2}{t_1}\le\phi_2(\boldsymbol{t})$. Lemma 23 then implies that $\bar{\boldsymbol{\phi}}$ is an amortization. □

Proof of Theorem 13. Lemma 21 showed that $\bar{\boldsymbol{\phi}}$ is an amortization. Lemma 19 showed that the allocation of bundle pricing maximizes virtual surplus with respect to $\bar{\boldsymbol{\phi}}$, and Lemma 20 showed that $\bar{\boldsymbol{\phi}}$ is tight for bundle pricing. Invoking Proposition 2 completes the proof. □