Abstract

We consider a multi-dimensional screening problem of selling a product with multiple quality levels. We show that selling only the highest quality product is optimal if higher valued consumers are less sensitive to quality. To prove this result we generalize a standard approach of optimizing with relaxed incentive compatibility constraints. With multi-dimensional preferences, binding incentive compatibility constraints are endogenous to the mechanism, and thus it is not clear which ones can be relaxed. Our methodology identifies the appropriate relaxation. This approach also allows us to identify conditions for the optimality of grand bundling in a multi-product setting 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 and Becker Friedman Institute, 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 2015, ACM Conference on Economics and Computation 2015, and the North American Econometric Society Meeting 2016.
How should a monopolist sell its differentiated products? While the optimal mechanism for selling a non-differentiated product to an agent with linear utility is to simply post a price, optimal mechanisms for differentiated products 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. The identified conditions are easy to interpret and generalize the previous known conditions.

Consider a monopolist who can sell a product in high or low quality to an agent whose values for the two qualities are drawn from a distribution (alternatively, the monopolist is facing a population of consumers and can not discriminate based on consumers’ identities). With arbitrary distributions, optimal mechanisms for this multi-dimensional screening problem can be quite complex. For instance, [Thanassoulis 2004] shows that the optimal mechanism may involve randomization. To explain our result, it will be convenient to write the agent’s value for these two versions of the product as a base value for the high-quality product and the same base value times the value ratio (value for low quality divided by value for high quality) for the low-quality product. We prove the optimality of selling only the high-quality product when the base value and value ratio are positively correlated.\footnote{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. Our extensions remove regularity by requiring a slightly stronger correlation condition.}

Our characterization is intuitive. Price discrimination can be effective when higher valued consumers are more sensitive to quality than lower valued consumers (Figure 1). These higher 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 value ratio eliminates this possibility. It implies that higher valued consumers (measured by base value) are less sensitive to quality (measured by the value ratio) than lower valued consumers.\footnote{We show optimality even when randomization is allowed.}

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 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 examples of such practice). We show that this simplification generalizes to the design of auctions. 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 condition generalizes the previously known conditions which assumed that the value
ratio is either known to be a constant, or is independently drawn from the base value. It
is a standard result of Stokey (1979), Riley and Zeckhauser (1983), and Myerson (1981)
more generally, that when the base value is private (is drawn at random) but the value ratio
is known (is a constant), 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 product is optimal
(and it is done by posting a price). The analysis of Armstrong (1996), applied our setting,
shows that the same result holds when the base value and value ratio are independently
distributed but both private to the agent.

Our sufficient conditions become necessary when base value and value ratio are perfectly
correlated, i.e., the value ratio is a function of the base value. Such an instance can be
represented by a curve on which values for the differentiated products lie. A corollary of our
main result is that if the curve only crosses lines from the origin from below, i.e., the value
ratio is monotonically non-decreasing in the base value, then selling only the high-quality
product is optimal. On the other hand, if the value ratio is not monotone in the base value
then we show that there exists a distribution for the base value for which selling only the
high-quality product is not optimal.

In this application the value ratio has a natural interpretation as a type’s discount factor for receiving
the item via standard delivery.
From the analysis of the perfectly correlated case, we see that the analyses of Armstrong (1996) where the value ratio is independent of the base value, and Stokey (1979) and Riley and Zeckhauser (1983) where the value ratio 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 value ratio 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.

**Bundling.** Our result applies more generally to a setting with multiple alternatives that are not necessarily vertically differentiated, and identifies conditions for optimality of a mechanism that simply posts a uniform price for all alternatives (i.e., each type will either choose its favorite alternative or an outside option). This result, applied to a setting with a multi-product seller where the consumer can buy multiple items (i.e., each bundle of items is an alternative), 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). This result does not require any structural assumptions on preferences over bundles, such as additivity or sub-additivity (other than free disposal). We also apply our framework directly to a special case of multi-product pricing where the consumer’s values are additive. Using a more specialized analysis we obtain optimality conditions for grand bundling that are more general than the aforementioned corollary of our main result. 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 an intuition that previously was provided through examples (Stigler, 1963; Adams and Yellen, 1976) or numerical results (Chu et al., 2011). 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 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 where types

---

4Our correlation condition does not directly restrict the support of the distribution.
can be ordered, it is sufficient to 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 envelope analysis to express expected profit in terms of the expected virtual surplus of allocation (i.e., virtual value minus the cost of allocation). A type’s virtual value is simply its value minus the revenue loss from the incentive constraints of higher types. Similarly, given a relaxation of a multi-dimensional problem, types can be assigned virtual values that capture the revenue loss from “higher” types specified by the relaxation. The main challenge of multi-dimensional mechanism design is that the paths (in the agent’s type space) on which the incentive constraints bind vary. Thus a straightforward attempt to generalize single-dimensional virtual values to multi-dimensional agents by selecting an arbitrary relaxation of the problem fails because the optimal mechanism of the relaxed problem will not generally be incentive compatible. We develop a methodology to obtain an appropriate relaxation. Importantly, our framework leaves the paths that parameterize the relaxation as a variable and solves for them.

Our methodology instantiated to show optimality of selling only high quality is the following. We need to show the existence of a virtual value function for which (a) pointwise 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. Since our theorem reduces the problem to selling only a single product, then it must be that the virtual value of the high-quality product is equal to the single-dimensional virtual value specified from 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; 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. We show that given our correlation condition, the allocation of selling only the high quality product indeed optimizes virtual surplus with respect to the identified virtual values.

**Related Work.** Our approach is related to several others in multi-dimensional screening. Wilson (1993) and Armstrong (1996) relax all incentive compatibility constraints other than local radial constraints (along straight lines to the zero type) and use the envelope analysis to identify optimal mechanisms. Carroll (2016) considers a robust version of screening an agent across multiple components (e.g., a multi-product seller who only knows the marginal

---

5Applications of multi-dimensional screening goes well beyond the settings studied in this work, to agents with non-quasilinear utilities (Pai and Vohra 2014), optimal taxation (Rothschild and Scheuer 2014), insurance (Azevedo and Gottlieb 2015), and dynamic mechanism design (Battaglini and Lamba 2015).
distribution for each item but not the joint distribution), relaxes all incentive compatibility constraints other than misreports along single components, and uses a duality approach to prove optimality of separate screening. In contrast to these works, we do not fix the relaxation a priori and identify an appropriate one for a given distribution of types. Rochet and Chone (1998) and Daskalakis et al. (2014) provide general characterizations of optimal mechanisms. These characterizations require multi-dimensional ironing operators whereas our main results use either simple single-dimensional ironing or no ironing at all. The conditions in Rochet and Chone (1998) have similarities to ours but their analysis requires a strict concavity condition that is not satisfied in our setting (see Appendix A for an in depth comparison).

A number of papers study optimal mechanisms for an agent with additive preferences. Vincent and Manelli (2007) characterize mechanisms that are optimal for some distribution of preferences. McAfee et al. (1989) show that with independent distributions, selling separately is never optimal. McAfee and McMillan (1988) and Manelli and Vincent (2006) find sufficient conditions for optimality of deterministic mechanisms (without identifying the optimal mechanism). Hart and Nisan (2012), Daskalakis et al. (2014), and Pavlov (2011) show optimality of grand bundling for certain independent distributions. Our conditions apply to correlated distributions and is not comparable to these results. For instance, Pavlov (2011) shows that when values for two items are independently and uniformly distributed on \([a, a + 1]\) for sufficiently large \(a\), the grand bundling mechanism is optimal. In comparison, a simple corollary of our theorem states that grand bundling is optimal for a uniform distribution over a triangle with vertices \((a, a), (a, b),\) and \((b, a)\), for any \(a \leq b\).

1 Preliminaries

We consider a single-agent mechanism design problem with allocation space \(X \subseteq \mathbb{R}^m\) for a finite \(m\), where the cost to the seller for producing allocation \(x \in X\) is \(c(x)\). The agent has a convex, continuous, and bounded type space \(T \subseteq \mathbb{R}^m\) with Lipschitz continuous boundary. The type \(t \in T\) of the agent is drawn at random from a distribution with continuous density \(f > 0\). The utility of the agent with type \(t \in T\) for allocation \(x \in X\) and payment \(p \in \mathbb{R}\) is \(t \cdot x - p\).
We use the revelation principle and focus on direct mechanisms. A single-agent mechanism is a pair of functions, the allocation function $x : T \to X$ 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,

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

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

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

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

$$E_{t \sim f} [p(t) - c(x(t))].$$

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

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

**Lemma 1** ([Rochet 1985]). Function $u$ is the utility function of an agent in an incentive-compatible mechanism if and only if $u$ is convex and non-decreasing. The allocation is $x(t) = \nabla u(t)$, wherever the gradient $\nabla u(t)$ is defined.\(^8\)

Even though the above lemma completely characterizes incentive compatibility in our setting, we will mainly use the envelope equality $x(t) = \nabla u(t)$ as a necessary (but not sufficient) condition for incentive compatibility of a mechanism.

## 2 The Framework: 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.

\(^8\)The gradient $\nabla u(t)$ is a vector $(\partial_1 u(t), \ldots, \partial_m u(t))$, where $\partial_i u(t)$ is the partial derivative of the function $u$ with respect to the $i$’th variable $t_i$. If $u$ is convex, $\nabla u(t)$ 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 $\nabla u(t)$ is defined everywhere.
A standard approach to understanding optimal mechanisms is as follows: in the single-dimensional special case of the problem, e.g., \( m = 1 \) alternative, 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 removed. The relaxed problem is relatively easy to solve: the envelope equality (\( \nabla u(t) = x(t) \) in our setting) can be used to express profit in terms of expected virtual surplus of the allocation. The allocation of the solution to the relaxed problem optimizes virtual surplus pointwise, and its payment is obtained from the un-relaxed conditions. Assuming that a regularity condition on the distribution of types holds, the identified solution to the relaxed problem satisfies the global IC and IR constraints, and thus it is a solution to the original problem.\(^9\)

The challenge with multi-dimensional preferences is that the paths on which incentive compatibility binds vary depending on the mechanism. Therefore, a mechanism that satisfies an arbitrary set of relaxed IC conditions may not be incentive compatible. Nonetheless, there may exist a relaxation whose optimal solution is incentive compatible. An example is a single-dimensional setting discussed above: even though downward local IC conditions alone are not sufficient to guarantee incentive compatibility of all mechanisms, when a regularity condition on the distribution holds the solution to the relaxed problem with only downward local IC constraints is an IC and IR mechanism. We extend this idea to our setting with multi-dimensional preferences as follows: consider a partial ordering on the set of types 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 (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:

\(^9\)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 $t$ reporting $t'$, and an IR condition on the lowest type on each path, e.g., type $t'$.

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 the appropriate relaxation for a given mechanism. Our main technical contribution is to identify a set of conditions, in Lemma 3, that imply that a candidate function is indeed a virtual surplus function induced by some relaxation. If the given mechanism optimizes a virtual surplus function satisfying the conditions of Lemma 3, then the mechanism is the solution to a relaxation of problem, and therefore must be optimal.

2.1 Illustration: An Instance with Two Types

To make the above discussion more concrete, we analyze a simple scenario with two types. We apply the analysis to show optimality of only selling the high quality product even though vertical differentiation is possible. A reader eager to see our general treatment can skip directly to Section 2.2.

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

$$
\mathbb{E}[p(t)] = p(t)f(t) + p(t')f(t'),
$$

(1)

\footnote{Only in this subsection we consider a distribution with discrete support and let $f$ denote the probability of types.}
subject to two IC constraints (one for each ordered pair of types) and two IR constraints (one for each type). Consider the IR constraint for type \( t' \),

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

and the IC constraint for type \( t \),

\[
t \cdot x(t) - p(t) \geq t \cdot x(t') - p(t'). \tag{3}
\]

Together, inequalities \( (2) \) and \( (3) \) imply an upper bound on the expected revenue \( (1) \) as explained below. In particular, \( (2) \) implies that \( p(t') \leq t' \cdot x(t') \), and \( (3) \) implies that \( p(t) - p(t') \leq t \cdot (x(t) - x(t')) \). We can now bound the expected revenue in terms of the allocation function given \( f(t) + f(t') = 1 \) and the above two inequalities,

\[
E[p(t)] = p(t)f(t) + p(t')f(t') = (p(t) - p(t'))f(t) + p(t') \\
\leq [t \cdot (x(t) - x(t'))]f(t) + t' \cdot x(t').
\]

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 side of the above inequality to get

\[
E[p(t)] \leq t \cdot x(t)f(t) + [t' - tf(t)] \cdot x(t') \\
= t \cdot x(t)f(t) + [t'f(t') - (t - t')f(t)] \cdot x(t') \\
= t \cdot x(t)f(t) + [t' - (t - t')f(t)/f(t')] \cdot x(t')f(t'),
\]

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

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

\[
E[p(t)] \leq E[\tilde{\phi}(t) \cdot x(t)], \tag{4}
\]

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. In particular, we show that the mechanism is the optimal solution to a
relaxation with constraints (2) and (3). The example below is a special case of our main result Theorem 6.

**Example 1.** Consider the problem of selling an item that can be offered in high or low quality to a unit-demand agent. In particular, let the allocation $x = (x_1, x_2)$, where $0 \leq x_1, x_2$ and $x_1 + x_2 \leq 1$ denote the probabilities $x_1$ and $x_2$ of assigning the product in high or low quality respectively. Assume without loss of generality that $t_1 \geq t_1'$. We show that the optimal mechanism only sells the high quality product (and not offer a discount for low quality product or randomized allocations) if $t'$ is below the ray that connects $t$ to the origin, that is, $t_2'/t_1' \leq t_2/t_1$ (see Figure 3).

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

$$\bar{\phi}(t) \cdot x(t) = t \cdot x(t) \leq t_1 = \bar{\phi}_1(t).$$  \hspace{1cm} (5)$$

The assumption that $t_2'/t_1' \leq t_2/t_1$ implies that

$$\frac{t_2'}{t_1'} \bar{\phi}_1(t') = \frac{t_2}{t_1} t_1' - (t_1 - t_1') \frac{f(t)}{f(t')} = t_2' - (t_1 \frac{t_2'}{t_1'} - t_2') \frac{f(t)}{f(t')} \geq t_2' - (t_2 - t_2') \frac{f(t)}{f(t')} = \bar{\phi}_2(t').$$ \hspace{1cm} (6)$$

Now given the above inequality, we show that a mechanism that only offers the high quality product is an optimal mechanism by arguing that it is the optimal solution of the relaxed problem (i.e., maintain (4) and (5) and relax the other two IC and IR constraints). Consider two cases:
Let us now summarize the above analysis to obtain a general approach for proving optimality of a given IC and IR mechanism \((x, p)\). 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 \(E[\overline{\phi}(t) \cdot x(t) - c(x(t))]\) as in \((4)\) (recall that the above analysis assumed that costs are zero and thus expected virtual surplus simplified to \(E[\overline{\phi}(t) \cdot x(t)]\)). Third, argue that \((x, p)\) is the optimum solution to the relaxed problem and therefore the optimum solution to the original problem by arguing that \(x\) optimizes virtual surplus subject to the feasibility condition \(x(t) \in X\) for all \(t\), and that the revenue of \((x, p)\) is indeed equal to \(E[\overline{\phi}(t) \cdot x(t) - c(x(t))]\) 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. 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.
2.2 Definitions: Amortizations and Virtual Values

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

Definition 1. A vector field $\bar{\phi} : T \to \mathbb{R}^m$ is an amortization of revenue if expected virtual surplus of any individually-rational incentive-compatible mechanism $(\hat{x}, \hat{p})$ upper bounds the mechanism’s expected profit, that is, $E[\phi(t) \cdot \hat{x}(t) - c(\hat{x}(t))] \geq E[\hat{p}(t) - c(\hat{x}(t))]$, or equivalently $E[\phi(t) \cdot \hat{x}(t)] \geq E[\hat{p}(t)]$. An amortization of revenue $\bar{\phi}$ is tight for a mechanism $(x, p)$ if the inequality is tight, i.e., $E[\phi(t) \cdot x(t)] = E[p(t)]$.

Definition 2. An amortization of revenue $\bar{\phi} : T \to \mathbb{R}^m$ is a virtual value function for a mechanism $(x, p)$ if $x$ pointwise maximizes virtual surplus, i.e., $x(t) \in \arg \max_{\hat{x} \in X} \hat{x} \cdot \phi(t) - c(\hat{x}), \forall t \in T$, and $\bar{\phi}$ is tight for the mechanism $(x, p)$. We alternatively say that the mechanism $(x, p)$ admits a virtual value $\bar{\phi}$.

Proposition 2. An incentive compatible and individually rational mechanism $(x, p)$ is optimal if it admits a virtual value function $\bar{\phi}$.

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

$$E[p(t) - c(x(t))] = E[\phi(t) \cdot x(t) - c(x(t))]$$

$$\geq E[\phi(t) \cdot \hat{x}(t) - c(\hat{x}(t))] \geq E[\hat{p}(t) - c(\hat{x}(t))].$$

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 version of our problem with value $v$ in type space $T = [\underline{v}, \bar{v}]$, the function $\phi$ 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 $v$. 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 $v^* \geq v$ such that virtual surplus $\phi(x) - cx$ 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 implying tightness. Therefore $\phi$ is a virtual value function for the posted price mechanism with price $v^*$ and the mechanism is optimal.

### 2.3 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 $\phi$ (instead of a parameterization by paths) that guarantees that it is indeed an amortization. In particular, we argue that $\phi$ defined as $\phi(t) = t - \alpha(t)/f(t)$ is an amortization if vector field $\alpha$ satisfies two properties:

- **divergence density equality**: $\nabla \cdot \alpha(t) = -f(t)$ for all types $t \in T$\(^{13}\)
- **boundary inflow**: $\alpha(t) \cdot \eta(t) \leq 0$ for all types $t \in \partial T$ where $\partial T$ denotes the boundary

\(^{13}\)These conditions resemble optimality conditions of [Rochet and Chone (1998)](https://doi.org/10.1016/S0095-1197(97)00037-2). However, their analysis does not apply in our setting since it requires the cost function to be strictly convex. A more in depth comparison is provided in [Appendix A](#).

\(^{14}\) $\nabla \cdot \alpha(t)$ is the divergence of $\alpha$ and is defined as $\nabla \cdot \alpha(t) = \partial_1 \alpha_1(t) + \ldots + \partial_k \alpha_k(t)$ (not to be confused with $\nabla \alpha(t)$ which is the gradient of $\alpha$).
Figure 4: The vector $\alpha$ specifies the direction of paths. For each type $t$, $\alpha(t)/f(t)$ is the revenue loss from the incentive compatibility constraints of higher types on the path.

of type space $T$, and $\eta(t)$ is the outward pointing normal vector to the boundary at type $t \in \partial T$.

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

**Lemma 3.** For any vector field $\alpha : T \to \mathbb{R}^m$ satisfying the divergence density equality and boundary inflow, the vector field $\phi(t) = t - \alpha(t)/f(t)$ is an amortization of revenue; moreover, $\phi$ is tight for any incentive compatible mechanism for which the participation constraint is binding for all boundary types with strict inflow, i.e., $u(t) = 0$ for $t \in \partial T$ with $\alpha(t) \cdot \eta(t) < 0$.

**Proof.** We apply integration by parts to rewrite the revenue of an IC and IR mechanism.\(^{15}\)

In particular, integration by parts allows expected utility $E[u(t)]$ to be rewritten in terms of

\[^{15}\text{Integration by parts for functions } u : \mathbb{R}^k \to \mathbb{R} \text{ and } \alpha : \mathbb{R}^k \to \mathbb{R}^k \text{ over a set } T \text{ with Lipschitz continuous boundary is as follows}
\]

$$\int_{t \in T} \nabla u(t) \cdot \alpha(t) \, dt = - \int_{t \in T} u(t)(\nabla \cdot \alpha(t)) \, dt + \int_{t \in \partial T} u(t)\alpha(t) \cdot \eta(t) \, dt,$$

where $\nabla \cdot \alpha(t)$ is the divergence of $\alpha$.\[^{15}\]
gradient $\nabla u$

$$
\int_{t \in T} \nabla u(t) \cdot \alpha(t) \, dt = -\int_{t \in T} u(t) (\nabla \cdot \alpha(t)) \, dt + \int_{t \in \partial T} u(t) \alpha(t) \cdot \eta(t) \, dt
$$

where the second equation followed from divergence density equality. The envelope equality $(x(t) = \nabla u(t), \text{Lemma 1})$, implies that $\nabla u(t) \cdot \alpha(t) = x(t) \cdot \alpha(t)$ (we only use the envelope equality in the direction of $\alpha$). Thus

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

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

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

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

$$
E \left[ p(t) \right] = E \left[ t \cdot x(t) - u(t) \right] \tag{8}
\leq E \left[ (t - \frac{\alpha(t)}{f(t)}) \cdot x(t) \right] = E \left[ \phi(t) \cdot x(t) \right].
$$

Finally, if the last term of the right-hand side of equation (7) is zero, then the inequalities above are equalities and the amortization is tight.

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

As discussed above, given $\phi$, $E[\phi(t) \cdot x(t)]$ is an upper bound on revenue obtained by only invoking incentive compatibility conditions along paths induced by $\alpha$, 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 $\phi$ is a virtual value function for an IC and IR mechanism $(x, p)$, then the mechanism is the solution to the relaxation of the problem induced by $\alpha$. 


For a single-dimensional agent with value $v$ in type space $T = [\bar{v}, \bar{v}]$, the canonical amortization of revenue that is tight for any mechanism with binding IR is unique and given by $\phi(v) = v - \frac{1-F(v)}{f(v)}$. When $\phi$ is monotone, the allocation of posting a price $\phi^{-1}(0)$ optimizes virtual surplus pointwise, and thus the canonical amortization $\phi$ is a virtual value function for the mechanism that posts a price for a deterministic allocation.

2.4 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 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 1). 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 3 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 $\phi$, another (non-canonical) amortization $\bar{\phi}$ 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 $\phi$ to obtain $\bar{\phi}$ 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).

---

16Divergence density equality implies that $\alpha(v) = \alpha(v) - F(v)$. Tightness requires that $\alpha(\bar{v})u(\bar{v}) = (\alpha(v) - 1)u(\bar{v}) = 0$. Since $u(\bar{v}) > 0$ for any non-trivial mechanism, we must have $\alpha(\bar{v}) = 1$ and thus $\alpha(v) = 1 - F(v)$. Tightness also requires that $\alpha(v)u(v) = 0$, which is satisfied for any mechanism with binding participation constraint $u(v) = 0$. 

---

16
Proposition 5. A vector field $\bar{\phi}$ is an amortization of revenue if, for all incentive compatible individually rational mechanisms $(\hat{x}, \hat{p})$ and some other amortization of revenue $\phi$, it satisfies $E[\bar{\phi}(t) \cdot \hat{x}(t)] \geq E[\phi(t) \cdot \hat{x}(t)]$.

We adopt the sweeping approach in Section 4 (and Theorem 14 which extends the main result of Section 3) to obtain a virtual value function $\bar{\phi}$ from a canonical amortization $\phi$. 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.

3 Optimality of Uniform Pricing

In this section we apply our framework to a multi-alternative setting where $X = \{x \in [0,1]^m | \sum_i x_i \leq 1\}$. That is, an allocation $x \in X$ is a distribution over $m$ alternatives, where $x_i$ is the probability of selecting alternative $i$ ($1 - \sum_i x_i$ 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. Another example is a multi-product seller that can offer different bundles of products. In this case, $x \in X$ is a probability distribution over bundles, with $x_i$ the probability of assigning bundle $i$, and $m$ is be the number of possible bundles of the products. As discussed in Section 2.4, we use a class of cost functions to restrict the admissible amortizations. Throughout this section we assume uniform constant marginal costs, that is, $c(x) = c \sum_i x_i$ for some constant service cost $c \geq 0$.

Our main result is to identify sufficient conditions that imply optimality of a uniform pricing mechanism in which the same price $p$ is posted for all non-trivial alternatives. In a direct representation of such a mechanism, each type is assigned a deterministic allocation of its favorite alternative (an alternative in $\arg \max_i t_i$) and pays $p$ if its value for the favorite alternative is at least $p$, and otherwise is assigned a zero allocation and price. For simplicity we analyze the special case with two alternatives, and provide extensions to multiple alternatives in Appendix B.2 (which also includes an extension to multi-agent auctions.).

---

17 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.

18 Any instance with constant but non-uniform marginal costs can be converted to an instance with zero cost by redefining value as value minus cost.

19 Tie-breaking among alternatives with maximum value can be arbitrary, as the seller obtains the same revenue since the posted price is uniform.
In Section 3.2 we identify a class of instances for which our sufficient conditions are also necessary.

### 3.1 Sufficient Conditions

We first define the conditions used in the statement of the main result. A distribution over \( T \subset \mathbb{R}^2 \) is max-symmetric if the distribution of the value for favorite alternative \( v = \max(t_1, t_2) \) is identical conditioned on either \( t_1 \geq t_2 \) or \( t_2 \geq t_1 \). For instance, any symmetric distribution over \( \mathbb{R}^2 \) is max-symmetric. Also, an instance with vertically differentiated alternatives, i.e., \( t_1 \geq t_2 \) for all \( t \), is max-symmetric.\(^{[20]}\)

Let \( F_{\text{max}}(v) \) and \( f_{\text{max}}(v) \) be the cumulative distribution and the density function of the value for favorite alternative. As described in Section 2, the amortization of revenue for a single-dimensional agent is

\[
\phi_{\text{max}}(v) = v - \frac{1 - F_{\text{max}}(v)}{f_{\text{max}}(v)}.
\]

Distribution \( F_{\text{max}} \) is regular if \( \phi_{\text{max}} \) is monotone non-decreasing in \( v \). Each type \( t \) has a value ratio \( \theta(t) := \min(t_1, t_2) / \max(t_1, t_2) \) which can be thought of as a measure of the type’s indifference between alternatives. Let \( F(\theta|v,i) \) be the conditional distribution of the value ratio on \( v = t_i \geq t_{-i} \), that is, \( F(\theta|v,i) = \Pr[\theta(t) \leq \theta|v = t_i \geq t_{-i}] \).

**Theorem 6.** Uniform pricing is optimal with \( m = 2 \) alternatives and uniform constant marginal costs \( c(x) = c \sum x_i \) for any \( c \geq 0 \) and max-symmetric distribution where (a) the favorite-alternative distribution \( F_{\text{max}} \) is regular and (b) the conditional distribution \( F(\theta|v,i) \) is monotone non-increasing in \( v \) for all \( \theta \) and \( i \).

Conditions of Theorem 6 on \( F_{\text{max}} \) and \( F(\theta|v,i) \) are orthogonal. To see this, notice that the conditional distributions \( F(\theta|v,i) \) jointly with \( F_{\text{max}} \) uniquely represent any max-symmetric distribution as follows: with probability \( \Pr[t_1 \geq t_2] \), draw \( t_1 \) from \( F_{\text{max}} \), \( \theta \) from \( F(\cdot|t_1, 1) \), and set \( t_2 = t_1 \theta \) (otherwise assign favorite value to \( t_2 \) and draw \( \theta \) from \( F(\cdot|t_2, 2) \)). Appendix B.4 relaxes the regularity condition of the above theorem by adopting a simple single-dimensional sweeping operator.

Monotonicity of \( F(\theta|v,i) \) in \( v \) is correlation of \( \theta \) and \( v \) in first order stochastic dominance sense.\(^{[21]}\) It states that as \( v \) increases, more mass should be packed between a ray parameterized by \( \theta \), 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 \( \theta \), more likely.

\(^{[20]}\)Since the distribution conditioned on the zero probability event \( t_2 \geq t_1 \) is arbitrary.

\(^{[21]}\)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).
Let us identify some examples and non-examples of the distributions satisfying the conditions of the theorem. A function $C$ is *ratio-monotone* if $C(v)/v$ is monotone non-decreasing in $v$. Fix a regular $F_{\max}$ and a ratio-monotone $C$ satisfying $C(v) \leq v$. The following two distributions satisfy the conditions of the theorem: draw $t_1$ from $F_{\max}$, and either define $t_2 = C(t_1)$ or draw $t_2$ uniformly from the interval $[C(t_1), t_1]$.

These instances are *vertically differentiated* (that is $t_2 \leq t_1$) by the assumption that $C(v) \leq v$, but other distributions can be obtained by mirroring the construction to obtain the distribution conditioned on $t_2 \geq t_1$ (see Figure 5 for an example of a symmetric distribution). As another example, if $t_1$ and $t_2$ are drawn independently from a distribution with density proportional to $e^{\max \log(x)}$ for any monotone non-decreasing convex function $h$, then the distribution satisfies the conditions of the theorem (see Appendix B.3). On the other hand, a distribution where values for alternatives are uniformly and independently drawn from $[v, \bar{v}]$, with $v > 0$, does not satisfy the conditions (when $v = 5, \bar{v} = 6$, Thanassoulis 2004 showed that uniform pricing is not optimal).

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. Therefore for the rest of this

---

In the first case, by monotonicity of $C(v)/v$ for any $\theta$ there exists $v^*$ such that $\theta(t) \leq \theta$ if $t_1 \leq v$, and $\theta(t) > \theta$ otherwise. As a result, $F(\theta|t_1, 1) = 1$ if $t_1 \leq v^*$, and $F(\theta|t_1, 1) = 0$ otherwise, which is monotone in $t_1$. In the second case, $F(\theta|t_1, 1) = 1$ is monotone in $t_1$ by monotonicity of $C(t_1)/t_1$ in $t_1$.

If a single mechanism, namely uniform pricing $(x, \bar{p})$, is optimal for each event (of alternative 1 or alternative 2 being the favorite alternative), the mechanism is optimal for any probability distribution over the two events. In particular, the revenue of any mechanism $(x, \bar{p})$ is $E[\bar{p}(t)] = E[\bar{p}(t)|t_1 \geq t_2]Pr[t_1 \geq t_2] + E[\bar{p}(t)|t_2 \geq t_1]Pr[t_2 \geq t_1]$, which is at most $E[\bar{p}(t)|t_1 \geq t_2]Pr[t_1 \geq t_2] + E[\bar{p}(t)|t_2 \geq t_1]Pr[t_2 \geq t_1] = E[\bar{p}(t)]$. 

---

22In the first case, by monotonicity of $C(v)/v$ for any $\theta$ there exists $v^*$ such that $\theta(t) \leq \theta$ if $t_1 \leq v$, and $\theta(t) > \theta$ otherwise. As a result, $F(\theta|t_1, 1) = 1$ if $t_1 \leq v^*$, and $F(\theta|t_1, 1) = 0$ otherwise, which is monotone in $t_1$. In the second case, $F(\theta|t_1, 1) = 1$ is monotone in $t_1$ by monotonicity of $C(t_1)/t_1$ in $t_1$.

23If a single mechanism, namely uniform pricing $(x, \bar{p})$, is optimal for each event (of alternative 1 or alternative 2 being the favorite alternative), the mechanism is optimal for any probability distribution over the two events. In particular, the revenue of any mechanism $(x, \bar{p})$ is $E[\bar{p}(t)] = E[\bar{p}(t)|t_1 \geq t_2]Pr[t_1 \geq t_2] + E[\bar{p}(t)|t_2 \geq t_1]Pr[t_2 \geq t_1]$, which is at most $E[\bar{p}(t)|t_1 \geq t_2]Pr[t_1 \geq t_2] + E[\bar{p}(t)|t_2 \geq t_1]Pr[t_2 \geq t_1] = E[\bar{p}(t)]$.
section we work with the distribution conditioned on \( t_1 \geq t_2 \). In particular, we assume that \( T \) is a bounded subset of \( \mathbb{R}^2 \) specified by an interval \([t_1, \bar{t}_1] \) of values \( t_1 \) and bottom and top boundaries \( t_2(t_1) \) and \( \bar{t}_2(t_1) \) satisfying \( t_2(t_1) \leq \bar{t}_2(t_1) \leq t_1 \). The proof follows the framework of Section 2. In Definition 4 we define \( \phi \) and \( \alpha \) from the properties they must satisfy to prove optimality of uniform pricing. Lemma 7 shows that \( \phi \) 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 a uniform pricing maximizes virtual surplus pointwise with respect to \( \phi \). The theorem follows from Proposition 2.

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

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

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

(b) Let \( \alpha_1(t) = (t_1 - \phi_1(t))f(t) = \frac{1-F_{\text{max}}(t_1)}{f_{\text{max}}(t_1)}f(t) \).

\[ \Delta \alpha_2 = \int \partial_2 \alpha_2 \]

by optimality of \((x, p)\) conditioned on each event.

\[ p_{\text{max}}(v) = v - \frac{1-F_{\text{max}}(v)}{f_{\text{max}}(v)} \]
(c) Define $\alpha_2(t)$ uniquely to satisfy divergence density equality $\partial_2\alpha_2 = -f - \partial_1\alpha_1$ and boundary orthogonality of the bottom boundary.

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

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

(a) First, $\phi_1(t)$ may only be a function of $t_1$; otherwise, if $\phi_1(t) > \phi_1(t')$ with $t_1 = t'_1$, maximizing virtual surplus pointwise with cost $c$ satisfying $\phi_1(t) > c > \phi_1(t')$ implies $x_1(t') = 0$, and either $x_1(t) > 0$ or $x_2(t) > 0$ (if $\phi_2(t) > \phi_1(t > c)$. Such an allocation $x$ is not the allocation of a uniform pricing mechanism. Second, given the first point, the expected virtual surplus of uniform pricing $p$ is $\int_{t_1 \geq p}[\phi_1(t_1) f_{\max}(t_1) - c] dt_1$, which by tightness we need to be equal to $(p - c)(1 - F_{\max}(p))$. Solving this equation for all $p$ gives $\phi_1(t) = \phi_{\max}(t_1)$.

(b) We obtain $\alpha_1$ from $\phi_1$ by Definition 3.

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

(d) We obtain $\phi_2$ from $\alpha_2$ by Definition 3.

For $\phi$ to prove optimality of uniform pricing, we need the allocation of uniform pricing to optimize virtual surplus pointwise with respect to $\phi$. This additional requirement demands that $\phi_1(t) \geq \phi_2(t)$ for any type $t \in T$ for which either $\phi_1(t)$ or $\phi_2(t)$ is positive. Straightforward algebra shows that this condition is implied by the angle of $\alpha(t)$ being at most the angle of $t$ with respect to the horizontal $t_1$ axis, that is, $t_2\alpha_1(t) \leq t_1\alpha_2(t)$ (Lemma 8 below). The direction of $\alpha$ corresponds to the paths on which incentive compatibility constraints are un-relaxed (Corollary 4). The following lemma is proven by the divergence theorem, and specifies the direction of $\alpha$.

---

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.

This argument applies only if $\phi_1(t) > 0$. Nevertheless, we impose the requirement that $\phi_1(t) = \phi_1(t_1)$ everywhere as it allows us to uniquely solve for $\phi$.
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 \( \{T, R, B, L\}(t_1, q) \), for top, right, bottom, and left boundaries respectively. 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

\[
T_q = \{ t | \Pr_{t'} [t_2' \leq t_2 | t_1' = t_1, t_1' \geq t_2'] = \frac{\int_{t_2' \leq t_2} f(t_1, t_2') dt_2'}{\int_{t_2' \leq t_1} f(t_1, t_2') dt_2'} \leq q \}.
\]

For example, notice that for the *perfectly correlated* class, that is when \( t_2 = C_{cor}(t_1) \), the equi-quantile curves \( C_q \) are identical to \( C_{cor} \).

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

**Proof.** Tightness follows directly from the definition of \( \phi_1 \) (see the justification for Step [a] of the construction). The divergence density equality and bottom boundary orthogonality of \( \alpha \) are automatically satisfied by Step [c] of the construction. Orthogonality of the right boundary \( (t_1 = \bar{t}_1) \) requires that \( \alpha(\bar{t}_1, t_2) \cdot (1, 0) = 0 \), 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, \( \alpha \cdot \eta \leq 0 \) 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 \( \alpha \) to \( C_1 \) implies that \( \alpha \) 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 \( \alpha \) on the subspace of type space to the right of \( t_1 \) and
More formally, divergence theorem is applied to the set of types \( T(t_1, q) = \{ t' \in T \mid t'_1 \geq t_1; F(t_2 | t_1) \leq q \} \) (see Figure 7). The divergence theorem equates the integral of the orthogonal magnitude of vector field \( \alpha \) 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 \( \alpha(t') \) with the upward orthogonal vector to \( C_q \) at \( t' \).

Differentiating this integral with respect to \( t_1 \) gives the desired quantity. More specifically, by the divergence theorem,

\[
\int_{t' \in T(t_1, q)} \eta(t') \cdot \alpha(t') \, dt' = \int_{t' \in T(t_1, q)} \nabla \cdot \alpha(t') \, dt' - \int_{t' \in \{R, B, L\}(t_1, q)} \eta(t') \cdot \alpha(t') \, dt'.
\]  

(9)

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

\[
= - \int_{t' \in T(t_1, q)} f(t') \, dt' - \int_{t' \in L(t_1, q)} \eta(t') \cdot \alpha(t') \, dt' \\
= -q(1 - F_{\max}(t_1)) - \int_{t' \in L(t_1, q)} \eta(t') \cdot \alpha(t') \, dt'
\]

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

\[
\int_{t' \in L(t_1, q)} \eta(t') \cdot \alpha(t') \, dt' = -\frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} \int_{t'_2 \leq C_q(t_1)} f(t_1, t'_2) \, dt'_2 \\
= -\frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} q f_{\max}(t_1) = -(1 - F_{\max}(t_1)) q.
\]

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

\[
\int_{t' \in T(t_1, q)} \eta(t') \cdot \alpha(t') \, dt' = 0.
\]

Since the above equation holds for all \( t_1 \) and \( q \), we conclude that \( \alpha \) is tangent to the equi-quantile curve at any type. \( \square \)

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 \( \phi_1(t) \geq \)

\[\text{The divergence theorem for vector field } \alpha \text{ is } \int_{t \in T} (\nabla \cdot \alpha)(t) \, dt = \int_{t \in \partial T} (\alpha \cdot \eta)(t) \, dt.\]
\( c \) then \( \phi_1(t) \geq \phi_2(t) \), and that \( \phi_1(t) \geq c \) if and only if \( t_1 \) is greater than a certain threshold (implied by monotonicity of \( \phi_1(t) \geq c \)).

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

*Proof.* Tangency of \( \alpha \) to the equi-quantile curves (Lemma 7) implies that 
\[
\frac{t_2}{t_1} \phi_1(t) = t_2 \left( t_1 - \frac{\alpha_1(t)}{f(t)} \right) = t_2 - \frac{t_2}{t_1} \cdot \frac{\alpha_1(t)}{f(t)} \geq t_2 - \frac{\alpha_2(t)}{f(t)} = \phi_2(t).
\]
Thus, for \( t \) with \( \phi_1(t) \geq c \), \( \phi_1(t) \geq \phi_2(t) \) and pointwise virtual surplus maximization serves the agent alternative 1. Since \( \phi_1(t) \) 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 \( \phi_1(t) \leq c \), again the above calculation implies that \( \phi_2(t) \leq c \) and therefore the type is not served. This allocation is the allocation of a uniform pricing. \( \square \)

We next prove Theorem 6 by showing that \( \phi = t - \alpha / f \) of Definition 4 is a virtual value function for a uniform pricing and invoke Proposition 2. Recall from Lemma 7 that \( \alpha \) is tangent to the equi-quantile curves. By Corollary 4, under the conditions of the theorem, uniform pricing is the optimal solution to the relaxation induced by equi-quantile paths.

*Proof of Theorem 6.* Lemma 7 showed that \( \phi \) is a tight amortization for any uniform pricing.** Lemma 8 showed that the allocation of a uniform pricing pointwise maximizes virtual surplus with respect to \( \phi \). \( \square \)

---

Special attention is needed in case that the left boundary is a singleton, since in that case \( f_{\text{max}}(t_1) = 0 \) and \( \alpha_1 \) is unbounded. In this case our analysis showed that \( \alpha \cdot \eta = 0 \) everywhere except possibly at \( (t_1, t_2(t_1)) \). Divergence theorem states that 
\[
\int_{t \in \partial T} (\alpha \cdot \eta)(t) \, dt = -\int_{t \in T} f(t) \, dt = -1,
\]
which implies that \( \alpha \cdot \eta \) is a negative Dirac delta at \( (t_1, t_2(t_1)) \). The integral of \( u(\alpha \cdot \eta) \) over the boundary is thus \(-u(t_1, t_2(t_1)) = 0\).
3.2 Perfect Correlations and Necessary Conditions

We next complement our main theorem by providing necessary conditions for optimality of uniform pricing for the class of perfectly correlated distributions below. Recall that a distribution is perfectly correlated if \( t_2 = C_{\text{cor}}(t_1) \) for some function \( C_{\text{cor}} \). Under the assumptions of the theorem, we argue that another mechanism (in particular, a mechanism that offers different prices for the two alternative) obtains more revenue than the optimal uniform pricing mechanism. 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 \( F_{\text{max}} \) such that uniform pricing is not optimal for the perfectly correlated instance jointly defined by \( F_{\text{max}} \) and \( C_{\text{cor}} \). Furthermore, if \( C_{\text{cor}}(t_1)/t_1 \) is strictly decreasing everywhere, then uniform pricing is not optimal for any distribution \( F_{\text{max}} \).

Note the contrast with Theorem 9, which proves that if \( C_{\text{cor}} \) satisfying \( C_{\text{cor}}(t_1) \leq t_1 \) is ratio-monotone and \( F_{\text{max}} \) is regular then uniform pricing is optimal.

4 Grand Bundle Pricing for Additive Preferences

A corollary of proving optimality of uniform pricing for the multi-alternative setting (definition in Section 3) is optimality of grand bundling for a multi-product seller with a single agent. A multi-product problem can be thought of as a special case of a multi-alternative setting, with an alternative for each bundle of products. A uniform pricing mechanism posts the same price for all bundles of products. Assuming free disposal and given the uniform price, any type of the agent will either choose the grand bundle of products or no bundle at all. Thus, the uniform pricing mechanism is in essence a grand bundle pricing mechanism in a multi-product setting. In this section we provide sufficient conditions for optimality of grand bundle pricing that are more permissive than conditions of Section 3 but only apply to more structured additive preferences.

We model a multi-product seller with \( k \) items and a buyer with additive preferences by defining the allocation space to be \( X = [0, 1]^k \) where \( x_i \) denotes the probability of assigning item \( i \). As discussed in Section 2.4 and similar to Section 3, we use a class of cost functions to restrict the admissible amortizations. In particular, we assume that the cost of an allocation \( x \in [0, 1]^2 \) is \( c(x) = c \max(x_1, x_2) \) for a \( c \geq 0 \). A symmetric distribution is defined by a
marginal distribution $F_{\text{sum}}$ of value for the bundle $s$ as well as a conditional distribution $F(\theta|s)$ of the ratio $\theta(t) = \max(t_1, t_2)/\min(t_1, t_2)$ conditioned on value for the bundle $s$.

**Theorem 10.** 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) $F_{\text{sum}}$ has monotone amortization $\phi_{\text{sum}}$ and (b) the conditional distribution $F(\theta|s)$ is monotone non-decreasing in $s$.

The proof of the theorem and some examples are in Appendix C.1. Similar to Section 3, we show that if the distribution is perfectly correlated, then the sufficient conditions are also necessary (Appendix C.2). The main result of Section 3, projected to the additive setting studied here, requires independence of $\theta$ and $s$ which is relaxed to negative correlation in the above theorem.

### References


---

28In 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 12 (the extension of Theorem 6 to more than two alternatives) identifies a sufficient positive correlation condition for optimality of grand bundle pricing. If the preference is indeed additive, we have $\delta_2 = 1 - \delta_1$, and Theorem 12 requires that $\delta_1$ be both positively and negatively correlated with $v$. The only admissible case is independence. 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., $\Pr[\delta \leq \delta(q, v)|v] = q$. Theorem 12 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 $q, \theta_2$. The only possible case is independence of $v$ and $\delta$, that is, $\delta(q, v)$ is a constant.


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 \( \nabla u(t) = x(t) \), but the convexity condition on the utility function is relaxed (see Lemma 1). In this section we refer to this problem as the relaxed problem. Recall from Section 2 (Proposition 2) that a mechanism \((x, p)\) is optimal if there exists a tight canonical amortization \( \phi \) (satisfying divergence density and boundary inflow of Lemma 3) such that the allocation function \( x \) optimizes virtual surplus pointwise, that is,

\[
x(t) \in \arg \max_{\tilde{x} \in X} \tilde{x} \cdot \phi(t) - c(\tilde{x}), \forall t.
\]

(11)
Now assume that the cost function \( c \) is strictly convex. As a result, for any \( t, x(t) \) is the virtual surplus maximizer if and only if it satisfies the first order conditions of optimality of (11), that is
\[
\phi(t) - \nabla c(x(t)) = 0.
\]

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

**Corollary 11.** If the cost function \( c \) is strictly convex, a mechanism \((x, p)\) with utility function \( u \) is the optimal solution to the relaxed problem if \( \alpha(t) := (t + \nabla c(x(t))) f(t) \) satisfies the divergence density and boundary inflow conditions and \( \phi \) is tight for the mechanism, that is,
\[
\begin{align*}
\bullet & \quad \nabla \cdot \alpha(t) = -f(t) \text{ for all types } t \in T. \\
\bullet & \quad \alpha(t) \cdot \eta(t) \leq 0 \text{ for all types } t \in \partial T, \text{ with } u(t) = 0 \text{ if } \alpha(t) \cdot \eta(t) < 0.
\end{align*}
\]

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 \((x, p)\) is the optimal solution to the relaxed problem if and only if \( \alpha(t) := (t + \nabla c(x(t))) f(t) \) satisfies
\[
\begin{align*}
\bullet & \quad \nabla \cdot \alpha(t) \geq -f(t) \text{ for all types } t \in T, \text{ with equality if } u(t) > 0. \\
\bullet & \quad \alpha(t) \cdot \eta(t) \leq 0 \text{ for all types } t \in \partial T, \text{ with } u(t) = 0 \text{ if } \alpha(t) \cdot \eta(t) < 0.
\end{align*}
\]

Note the differences with Corollary 11: 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 11 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 3

This section includes proofs from Section 3.
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 $F_{\text{max}}$ such that uniform pricing is not optimal for the perfectly correlated instance jointly defined by $F_{\text{max}}$ and $C_{\text{cor}}$. Furthermore, if $C_{\text{cor}}(t_1)/t_1$ is strictly decreasing everywhere, then uniform pricing is not optimal for any distribution $F_{\text{max}}$.

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: 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 $C_{\text{cor}}$ 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 8)

$$\int_{t_1=p-\frac{c_{\text{cor}}(p)}{C_{\text{cor}}(p)}}^{p} f_{\text{max}}(t_1) \, dt_1.$$ 

Therefore, the positive effect is at least

$$\Delta_+(\epsilon) := (C_{\text{cor}}(p) - \epsilon) \times \int_{t_1=p-\frac{c_{\text{cor}}(p)}{C_{\text{cor}}(p)}}^{p} f_{\text{max}}(t_1) \, dt_1.$$ 

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 8)

$$\int_{t_1=p}^{p+\frac{c_{\text{cor}}(p)}{p-C_{\text{cor}}(p)}} f_{\text{max}}(t_1) \, dt_1.$$ 

Therefore, the negative effect is at most

$$\Delta_-(\epsilon) := (p - C_{\text{cor}}(p) + \epsilon) \times \int_{t_1=p}^{p+\frac{c_{\text{cor}}(p)}{p-C_{\text{cor}}(p)}} f_{\text{max}}(t_1) \, dt_1.$$
Figure 8: (a) As a result of adding an offer with price $C_{\text{cor}}(p) - \epsilon$ for alternative 2 to the existing offer of price $p$ for alternative 1, the types in darker shaded part of curve $C_{\text{cor}}$ 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 $\epsilon p / C_{\text{cor}}(p)$ and $p - \epsilon p / C_{\text{cor}}(p)$, respectively. (b) For any $\theta$, $t_1$, $F(\theta|t_1,1) = 1$ if $C_{\text{cor}}(t_1)/t_1 \leq \theta$, and $F(\theta|t_1,1) = 0$ otherwise. Therefore, ratio-monotonicity is equivalent to monotonicity of $F(\theta|t_1,1)$ in $t_1$.

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 $\epsilon$ by showing that $\Delta'_+(0) \geq \Delta'_-(0)$ (since $C_{\text{cor}}(t_1)/t_1$ 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

$$\Delta'_+(\epsilon) = -\int_{t_1 = p - \frac{\epsilon p}{C_{\text{cor}}(p)}}^{p} f_{\text{max}}(t_1) \, dt_1 + \left( C_{\text{cor}}(p) - \epsilon \right) \times \frac{p}{C_{\text{cor}}(p)} f_{\text{max}}(p - \frac{\epsilon p}{C_{\text{cor}}(p)}),$$

and therefore,

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

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

We will now complete the proof of both statements. For the first statement, consider $p$ where $C_{\text{cor}}(t_1)/t_1$ is decreasing at $t_1 = p$, and any regular distribution $F_{\text{max}}$ such that $p$ maximizes $p(1 - F_{\text{max}}(p))$. 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 $C_{\text{cor}}(t_1)/t_1$ 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 Extension to Multiple Alternatives and Agents

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 \( q_j^i(t) \) to be the quantile of the distribution of \( t_j \) conditioned on \( i \) being the favorite outcome, and conditioned on the values \( t_{<j} = (t_1, \ldots, t_{j-1}) \) of the lower indexed outcomes. Formally, 

\[
q_j^i(t) = \Pr_{t'}[t_j' \leq t_j | t_{<j}' = t_{<j}, t_i' = t_i = \max_k t_k]
\]

Define \( F(\theta_j | t_i, i, q_{<j}) = \Pr_{t'}[t_j' / t_i' \leq \theta_j | q_{<j} = q_{<j}(t'), t_i' = t_i = \max_k t_k] \) to be the distribution of the value ratio of \( j \)th to favorite outcome, conditioned on \( i \) being the favorite outcome and given vector \( q_{<j} \) 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.\(^{29}\)

**Theorem 12.** A favorite-outcome projection mechanism is optimal for an item with \( m \geq 1 \) configurations, multiple independent agents, and uniform constant marginal costs \( c(x) = c \sum_i x_i \) with any \( c \geq 0 \), if the distribution of each agent is max-symmetric and (a) the favorite-outcome projection has monotone non-decreasing amortization \( \phi_{\text{max}}(v) = v - \frac{1 - F_{\text{max}}(v)}{F_{\text{max}}(v)} \) and (b) \( F(\theta_j | v, i, q_{<j}) \) is monotone non-increasing in \( v \) for all \( i, j, \theta_j, \) and \( q_{<j} \).

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 \( \phi_{\text{max}}(v) \) (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 \( v \).

---

\(^{29}\)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.
Proof. The construction extends the construction of Theorem 6. Let outcome 1 be the favorite outcome. For \( q \), let \( C^q(t_1) \) be a function that maps \( t_1 \) to \((t_2, \ldots, t_m)\) such that \( q(t) = q \). Define \( \alpha \) by integrating by parts along the curves \( C^q(t_1) \). This defines \( \alpha_1(t) = \frac{1 - F_{\max}(t_1)}{f_{\max}(t_1)} f(t) \), and \( \alpha_i(t) = \alpha_1(t) \partial_{t_1} C^q_i(t_1) \). The assumptions of the theorem also implies that \( \alpha_i(t) - (t_i/t_1)\alpha_1(t) \leq 0 \). As a result, \( \phi_i(t) \leq (t_i/t_1)\phi_1(t) \).

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 13.** 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 12. In order to show that \( F(\theta|v) \) is monotone in \( v \), we show that the joint distribution of \( \theta \) and \( v \) satisfies the stronger property of affiliation. That is,

\[
f^{MR}(t_1, \theta) \times f^{MR}(t_1', \theta') \geq f^{MR}(t_1, \theta') \times f^{MR}(t_1', \theta), \quad \forall t_1 \leq t_1', \theta \leq \theta',
\]

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

\[
f_1(t_1)f_2(t_1 \theta) \times f_1(t_1')f_2(t_1' \theta') \geq f_1(t_1)f_2(t_1 \theta') \times f_1(t_1')f_2(t_1' \theta),
\]

33
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 \( F_{\text{max}} \) is regular. This is the case because \( f_{\text{max}}(v) = F(v) f(v) \), \( f(v) = \eta \cdot e^{h(\log(v))} \) is monotone in \( v \) by monotonicity of \( h \).

**B.4 Relaxing the Regularity Condition and Ironing**

The second extension removes the regularity assumption of **Theorem 12** 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 \( \phi \) along the equi-quantile curves.

**Theorem 14.** A favorite-outcome projection mechanism is optimal for an item with \( m = 2 \) configurations, multiple independent agents, and uniform constant marginal costs \( c(x) = c \sum_i x_i \) with any \( c \geq 0 \), 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 9** 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 9**).

We will design a virtual value function \( \tilde{\phi} \) from the canonical amortization \( \phi \) satisfying conditions of **Lemma 3**. Importantly, \( \tilde{\phi} \) satisfies the monotonicity of \( \tilde{\phi}_1 \) 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 \( \tilde{\phi}_1 \) of the amortization. The second component \( \tilde{\phi}_2 \) is then defined such that the expected virtual surplus with respect \( \tilde{\phi} \) 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 \( \tilde{\phi} - \phi \) may have for \( \tilde{\phi} \) to be an upper bound on revenue. We use this identity to solve for \( \tilde{\phi}_2 \), and finally identify conditions such that optimization of \( \tilde{\phi} \) gives uniform pricing.
We first transform the value space to quantile space using following mappings. Recall from Section 3 that $F_{\text{max}}$ and $f_{\text{max}}$ are the distribution and the density functions of the favorite item projection. Define the first quantile mapping

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

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

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

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

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

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

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

Figure 9: The connection between convexity and ratio-monotonicity of equi-quantile curves. (a) Convexity implies ratio-monotonicity. (b) Ratio-monotonicity does not imply convexity.
where \( x^Q \) and \( \phi^Q \) are representations of \( x \) and \( \phi \) in quantile space. In particular, \( \phi^Q_1(q) = \phi_{\max}(t_1(q_1)) \) might not be monotone in \( q_1 \). In what follows we design the amortization \( \bar{\phi}^Q \) using \( \phi^Q \).

We now derive \( \bar{\phi}^Q \) from the properties it must satisfy. In particular, we require \( \bar{\phi}^Q_1(q_1) = \phi^Q_1(q_1) \) to be a monotone non-decreasing function of \( q_1 \), and that \( \bar{\phi}^Q_1(q_1) \geq \bar{\phi}^Q_2(q_1) \) whenever either is positive. These properties will imply that a pointwise optimization of \( \bar{\phi}^Q \) will result in an incentive compatible allocation of only the favorite item, such that \( x^Q_1(q_1) = x^Q_1(q_1) \), and \( x^Q_2(q_1) = 0 \) (which is the case for the allocation of uniform pricing). Note that for any such allocation,

\[
\int_{q_1=0}^{1} \int_{q_2=0}^{1} x^Q(q) \cdot \phi^Q(q) \, dq = \int_{q_1}^{1} x^Q_1(q_1) \phi^Q_1(q_1) \, dq_1.
\]

Similarly, for any such allocation,

\[
\int_{q_1=0}^{1} \int_{q_2=0}^{1} x^Q(q) \cdot \phi^Q(q) \, dq = \int_{q_1}^{1} x^Q_1(q_1) \phi^Q_1(q_1) \, dq_1.
\]

We can therefore use Myerson’s ironing and define \( \bar{\phi}^Q_1 \) to be the derivative of the convex hull of the integral of \( \phi^Q_1 \). This will imply that \( \bar{\phi}^Q \) upper bounds revenue for any allocation that satisfies \( x^Q_1(q_1) = x^Q_1(q_1) \), and \( x^Q_2(q_1) = 0 \), with equality for the allocation that optimizes \( \bar{\phi}^Q \) pointwise.

We will next define \( \bar{\phi}^Q_2 \) such that \( \bar{\phi}^Q \) upper bounds revenue for all incentive compatible allocations. That is, we require that for all incentive compatible \( x \),

\[
\int \int x^Q(q) \cdot (\bar{\phi}^Q - \phi^Q)(q) \, dq \geq 0.
\]

Using integration by parts we can write

\[
\int \int x^Q(q) \cdot (\bar{\phi}^Q - \phi^Q)(q) \, dq = \int_{q_2} \int_{q_1} \frac{d}{dq_1} x^Q(q) \cdot \int_{q_1}^{q_2} (\bar{\phi}^Q - \phi^Q)(q_1, q_2) \, dq_1 \, dq_1 \, dq_2.
\]

Incentive compatibility implies that the dot product of any vector and the change in allocation rule in the direction of that vector is non-negative [Lemma 1]. In particular this must be true for the tangent vector to equi-quantile curve parameterized by \( q_2 \). Thus incentive compatibility of \( x \) implies that the above expression is positive if the vector that is multiplied by \( \frac{d}{dq_1} x^Q(q) \) 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
$q_1 = q_1,$

$$
\int_{q_1 \geq q_1} (\bar{\phi}_2^Q - \phi_2^Q)(q'_1, q_2) \, dq'_1
\quad = \quad \frac{1}{dq_1} t_2(q).
$$

We will set $\bar{\phi}_2^Q$ to satisfy the above equality. In particular, define for simplicity $\mu(q) = \frac{1}{dq_1} t_1(q)$ and take derivative of the above equality with respect to $q_1$

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

As a result, $\bar{\phi}_2^Q$ defined above is a tight amortization if its optimization indeed gives uniform pricing. The next lemma formally states the above discussion.

**Lemma 15.** The virtual surplus, with respect to $\bar{\phi}_2^Q$ of any incentive compatible allocation $x$ upper bounds its revenue. If $x_1$ is only a function of $q_1$ (equivalently $t_1$), $x'_1(q_1) = 0$ whenever $\int_{q_1 \geq q_1} (\bar{\phi}_1^Q - \phi_1^Q)(q'_1) \, dq'_1 > 0$, and $x_2(q) = 0$ for all $q$, the expected virtual surplus with respect to $\bar{\phi}_2^Q$ equals revenue.

We will finally need to verify that $\bar{\phi}_2^Q$ also satisfies the properties required for ex-post optimization. **Lemma 17** below identifies convexity of equi-quantile curves as a sufficient condition. The proof requires the following technical lemma.

**Lemma 16.** The amortization $\bar{\phi}$ satisfies $\bar{\phi}_1(t) \leq t_1$.

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

$$
\bar{\phi}_1^Q(q''_1) = \frac{q_1 t_1(q'_1) - qt_1(q_1)}{q'_1 - q_1} \leq \frac{q_1 t_1(q'_1) - qt_1(q'_1)}{q'_1 - q_1} = t_1(q'_1) \leq t_1(q''_1).
$$

**Lemma 17.** If the equi-quantile curves are convex for all $q_2$, the amortization $\bar{\phi}_2^Q$ defined above satisfies $\theta(q) \bar{\phi}_1^Q(q) \geq \bar{\phi}_2^Q(q)$. As a result, $\bar{\phi}_1^Q \geq \bar{\phi}_2^Q$ whenever either is positive.
Proof. [Lemma 7] showed that $\alpha$ is tangent to the equi-quantile curves. This implies that 
\[
\phi_1^Q(q)\mu(q) - \phi_2^Q(q) = t_1(q)\mu(q) - t_2(q).
\]
By rearranging the definition of $\phi_2$ we get 
\[
\bar{\phi}_1^Q(q)\mu(q) - \bar{\phi}_2^Q(q) = \phi_1^Q(q)\mu(q) - \phi_2(q) + \int_{q'_1 \geq q} (\bar{\phi}_1^Q(q'_1) - \phi_1^Q(q'_1, q_2)) \, dq'_1 \cdot \frac{d}{dq_1} \mu(q) \geq t_1(q)\mu(q) - t_2(q),
\]
where the inequality followed since by definition of $\bar{\phi}_1^Q$, we have \[
\int_{q'_1 \geq q} (\bar{\phi}_1^Q - \phi_1^Q(q'_1, q_2)) \, dq'_1 \geq 0, \text{ and } \frac{d}{dq_1} \mu(q) \geq 0 \text{ by the assumption of the lemma.}
\]
We can now rearrange the above inequality and write 
\[
t_2(q) - \bar{\phi}_2^Q(q) \geq \mu(q)(t_1(q) - \bar{\phi}_1^Q(q)) \geq \theta(q)(t_1(q) - \bar{\phi}_1^Q(q)),
\]
where the inequality followed since convexity of equi-quantile curves imply that $\mu(q) \geq \theta(q)$, and by Lemma 16, $t_1(q) - \bar{\phi}_1^Q(q) \geq 0$.

We can now use the above inequality to write 
\[
\theta(q)\bar{\phi}_1^Q(q) = \theta(q)(t_1(q) + (\bar{\phi}_1^Q(q) - t_1(q))) = t_2(q) + \theta(q)(\bar{\phi}_1^Q(q) - t_1(q)) \geq t_2(q) - \phi_2^Q(q) - t_2(q) = \phi_2^Q(q). \]

Proof of [Theorem 14] Combining Lemma 15 and Lemma 17 proves the theorem. 

C Proofs from Section 4

C.1 Proof of [Theorem 10]

This section contains the proof of [Theorem 10], restated below.

[Theorem 10] For a single agent with additive preferences over two items, bundle pricing
Figure 10: 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 $\{t|t_1 + t_2 = s, \min(t_1, t_2)/\max(t_1, t_2) \geq \theta(s)\}$. 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 $t_1, t_2 \in [\delta, \bar{s} - \delta], t_1 + t_2 \leq \bar{s}$.

is optimal for any costs $c\max(x_1, x_2), c \geq 0$, and any symmetric distribution where (a) $F_{\text{sum}}$ has monotone amortization $\phi_{\text{sum}}$ 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 10. Draw the value for the bundle $s$ from a regular distribution $F_{\text{sum}}$, 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 10).

We now turn to the proof of Theorem 10. Similar to Section 3, it is sufficient to prove the statement assuming $t_1 \geq t_2$. As in Section 3, the sum-of-values projection, via the divergence density equality (of Lemma 3), pins down an amortization $\phi$ 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 $\phi$ is not pointwise optimized by a grand bundle pricing. For this reason, we directly define $\bar{\phi}$ and then prove that it is a virtual value function for the grand bundle pricing mechanism by comparing the virtual surplus with respect to $\bar{\phi}$ and $\phi$.

**Definition 6.** The two-dimensional extension $\bar{\phi}$ 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:

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

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

The following lemma provides conditions on vector field $\bar{\phi}$ such that bundle pricing maximizes virtual surplus pointwise with respect to $\bar{\phi}$. These conditions are satisfied for $\bar{\phi}$.
Lemma 18. The allocation of a bundle pricing mechanism pointwise optimizes virtual surplus with respect to vector field $\bar{\phi}$ for all costs $c \max(x_1, x_2)$ if and only if: $\bar{\phi}_1(t)$ and $\bar{\phi}_2(t)$ have the same sign, $\bar{\phi}_1(t) + \bar{\phi}_2(t)$ 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 $x$ of posting a price $p$ for the bundle optimizes $\phi$ pointwise. Pointwise optimization of $x \cdot \bar{\phi}$ will result in $x = (1, 1)$ whenever $\bar{\phi}_1 + \bar{\phi}_2 \geq c$, and $x = (0, 0)$ otherwise. \hfill \square

Given Lemma 18, the remaining steps in proving that $\bar{\phi}$ is a virtual value function is showing that it is a tight amortization for grand bundle pricing. The following lemma proves tightness.

Lemma 19. The expected revenue of a bundle pricing is equal to its expected virtual surplus with respect to the two-dimensional extension $\bar{\phi}$ of the sum-of-values projection (Definition 6).

Proof. Let $x^p$ be the allocation corresponding to posting price $p$ for the bundle, that is $x^p_1(t) = x^p_2(t) = 1$ if $t_1 + t_2 \geq p$, and $x^p_1(t) = x^p_2(t) = 0$ otherwise. We will show that the virtual surplus of $x^p$ is equal to the revenue of posting price $p$, $R(p) = p(1 - F_{\text{sum}}(p))$. The virtual surplus is

$$\int_{t \in T} (x^p \cdot \phi f)(t) \, dt = \int_{t \in T} x^p(t_1, t_2) \cdot \phi(t_1, t_2) f(t_1, t_2) \, dt$$

$$= \int_{t \in T \cap t_1 + t_2 \geq p} \phi_{\text{sum}}(t_1 + t_2) f(t_1, t_2) \, dt.$$ 

$$= -\int_{s \geq p} \frac{d}{ds}(s(1 - F_{\text{sum}}(s))) \, ds$$

$$= R(p) - R(1) = R(p).$$  

\hfill \square

The rest of this section shows that $\bar{\phi}$ provides an upper bound on revenue of any mechanism. For that, we study the existence of a tight canonical amortization $\phi$ such that the virtual surplus of any incentive compatible mechanism with respect to $\bar{\phi}$ upper bounds its virtual surplus with respect to $\phi$ (any such $\phi$ must be tight for any bundle pricing since $\bar{\phi}$
Define the \textit{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 20.** \textit{If the conditional distribution $F(\theta|s)$ is monotone non-decreasing in $s$, then there exists a canonical amortization $\phi(t) = t - \alpha(t)/f(t)$ such that $E[x(t) \cdot (\bar{\phi}(t) - \phi(t))] \geq 0$ for all incentive compatible mechanisms.} For any $t$, $\alpha(t)$ is tangent to the equi-quantile curve crossing $t$.

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

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

That is, we use a sweeping process in a single dimension and along lines with constant sum of values $s$ (see Section 2.4). Consider the amortization $\phi$ that, like $\bar{\phi}$, sets $\phi_1(t) + \phi_2(t) = \phi_{\text{sum}}(t_1 + t_2)$ but, unlike $\bar{\phi}$, 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 $x \cdot (\bar{\phi} - \phi) = (x_1 - x_2)(\bar{\phi}_1 - \phi_1)$. We will first show that to satisfy equation (12) for all incentive compatible $x$ it is sufficient for $\phi$, relative to $\bar{\phi}$, 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 $\bar{\phi}_1 t_2/t_1 = \bar{\phi}_2$, the condition $\phi_1 \leq \bar{\phi}_1$ is equivalent to the condition $\phi_1 t_2/t_1 \leq \phi_2$.

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

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

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 $\theta$ 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 $\theta$ (Lemma 21). Second, if $\phi$ shifts value from coordinate one to coordinate two relative to the vector field $\bar{\phi}$, then, it also shifts expected value from coordinate one to coordinate two, conditioned 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 $\phi$ relative to $\bar{\phi}$ and equation (12) is satisfied (by Lemma 22).
Later in the section we will describe sufficient conditions on the distribution to guarantee existence of $\phi$ where this sufficient condition that $\frac{t_2}{t_1} \leq \phi_2$, is satisfied (Lemma 23).

**Lemma 21.** The allocation of any differentiable incentive compatible mechanism satisfies

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

*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 $x$ in the direction of that vector, here $\frac{d}{d\theta} x^{SR}(s, \theta)$, is positive.*

**Lemma 22.** The two-dimensional extension of the sum-of-values projection $\tilde{\phi}$ is an amortization if there exists an amortization $\phi$ with $\phi_1(t) + \phi_2(t) = \phi_{\text{sum}}(t_1 + t_2)$ that satisfies $\phi_1(t) \frac{t_2}{t_1} \leq \phi_2(t)$.

*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 $t_1$).*

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

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

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

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

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

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

---

\[30\] 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.
Write the left-hand side of equation (12) as:

\[
E \left[ x(t) \cdot (\bar{\phi}(t) - \phi(t)) \mid t_1 + t_2 = s \right]
\]

\[
= \int_{\theta=0}^{1} x^{SR}(s, \theta) \cdot \left[ \bar{\phi} - \phi \right]^{SR}(s, \theta) f^{SR}(s, \theta) \frac{s}{1 + \theta} \, d\theta
\]

\[
= \int_{\theta=0}^{1} x^{SR}(s, \theta) \cdot \frac{d}{d\theta} \int_{\theta'=0}^{\theta} \left[ \bar{\phi} - \phi \right]^{SR}(s, \theta') f^{SR}(s, \theta') \frac{s}{1 + \theta'} \, d\theta' \, d\theta.
\]

Substituting \( \Gamma \) into the integral above, we have

\[
= \int_{\theta=0}^{1} x^{SR}(s, \theta) \cdot \frac{d}{d\theta} \Gamma(s, \theta) \, d\theta
\]

\[
= x^{SR}(s, \theta) \cdot \Gamma(s, \theta) \bigg|_{\theta=0}^{1} - \int_{\theta=0}^{1} \frac{d}{d\theta} x^{SR}(s, \theta) \cdot \Gamma(s, \theta) \, d\theta.
\]

\[
= - \int_{\theta=0}^{1} \frac{d}{d\theta} x^{SR}(s, \theta) \cdot \Gamma(s, \theta) \, d\theta \, ds
\]

\[
\geq 0.
\]

The second equality is integration by parts. The third equality follows because the first term on the left-hand side is zero: For \( \theta = 0, \Gamma(s, \theta) = 0 \) by property \( \text{(b)} \); 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 \( \text{(a)} \). The final inequality follows from \( -\frac{d}{d\theta} x^{SR}(s, \theta) \cdot (1, -1) \geq 0 \) (Lemma 21) and properties \( \text{(a)} \) and \( \text{(b)} \). \( \square \)

To identify sufficient conditions for \( \bar{\phi} \) to be an amortization it now suffices to derive conditions under which there exists a canonical amortization \( \phi \) satisfying \( \phi_1(t) + \phi_2(t) = \phi_{\text{sum}}(t_1 + t_2) \) and the condition of Lemma 22, i.e., \( \phi_1(t) \frac{t_2}{t_1} \leq \phi_2(t) \). 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

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

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

The following constructs the canonical amortization \( \phi \) and specifies the direction of \( \alpha \). Similar to Section 3, \( \alpha \) 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 23.** A canonical amortization \( \phi = t - \alpha / f \) satisfying \( \phi_1(t) + \phi_2(t) = \phi_{\text{sum}}(t_1 + t_2) \) exists and is unique, where \( \alpha(t) \) is tangent to the equi-quantile curve crossing \( t \).
Proof. We assume that \( \phi \) satisfying the requirements of the lemma exists, derive the closed form suggested in the lemma, and then verify that the derived \( \phi \) indeed satisfies all the required properties. We fix \( s \) and \( q \) and apply the divergence theorem to \( \alpha \) 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 \( T(s, q) = \{ t' \in T | t'_1 + t'_2 \geq s; F(t_2|s) \leq q \} \). The divergence theorem equates the integral of the orthogonal magnitude of vector field \( \alpha \) 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 \( \alpha(t') \) with the upward orthogonal vector to \( C_q \) at \( t' \). Differentiating this integral with respect to \( t_1 \) gives the desired quantity.

\[
\int_{t' \in \text{TOP}(s, q)} \eta(t') \cdot \alpha(t') \, dt' = \int_{t' \in T(s, q)} \nabla \cdot \alpha(t') \, dt' - \int_{t' \in \{ \text{RIGHT,BOTTOM,LEFT} \}(s, q)} \eta(t') \cdot \alpha(t') \, dt'. \tag{13}
\]

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

\[
= -\int_{t' \in T(s, q)} f(t') \, dt' - \int_{t' \in \{ \text{LEFT} \}(s, q)} \eta(t') \cdot \alpha(t') \, dt' \\
= -q(1 - F_{\text{sum}}(s)) - \int_{t' \in \{ \text{LEFT} \}(q)} \eta(t') \cdot \alpha(t') \, dt'
\]

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

\[
\int_{t' \in \{ \text{LEFT} \}(s, q)} \eta(t') \cdot \alpha(t') \, dt' = -\frac{1 - F_{\text{sum}}(s)}{f_{\text{sum}}(s)} \int_{t'_2 \leq C_q(t_1)} f(t_1, t'_2) \, dt'_2 \\
= -\frac{1 - F_{\text{sum}}(s)}{f_{\text{sum}}(s)} - qf_{\text{sum}}(s) \\
= -(1 - F_{\text{sum}}(s))q
\]

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

\[
\int_{t' \in \text{TOP}(s, q)} \eta(t') \cdot \alpha(t') \, dt' = 0.
\]

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

We now complete the proof of Theorem 10.

Proof of Lemma 20. The assumption that $F(\theta|s)$ is monotone implies that the equi-quantile curves are ratio-monotone. The tangency property of Lemma 23 implies that $\alpha_1 t_2 / t_1 \geq \alpha_2$ and subsequently $\phi_1(t) t_2 / t_1 \leq \phi_2(t)$. Lemma 22 then implies that $\phi$ is an amortization.

Proof of Theorem 10. Lemma 20 showed that $\tilde{\phi}$ is an amortization. Lemma 18 showed that the allocation of bundle pricing maximizes virtual surplus with respect to $\tilde{\phi}$, and Lemma 19 showed that $\tilde{\phi}$ is tight for bundle pricing. Invoking Proposition 2 completes the proof.

C.2 Necessary Conditions for Optimality of Grand Bundle Pricing

Similar to Section 3, we first study a family of instances with perfect correlation to obtain necessary conditions of optimality. In particular, let $F_{\text{sum}}$ 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 is, value for item 1 is $t_1 = s / (1 + \theta)$, and value for item 2 is $t_2 = \theta s / (1 + \theta)$.

The following theorem shows that if $\theta(s)$ is not monotone non-increasing in $s$, then bundling is not optimal for some distribution $F_{\text{sum}}$. The proof is similar to Theorem 9 and is omitted.

Theorem 24. If $\theta(s)$ is not monotone non-increasing in $s$, then there exists a regular distribution $F_{\text{sum}}$ over $s$ such that grand bundle pricing is not optimal for the perfectly correlated instance jointly defined by $F_{\text{sum}}$ and $\theta(\cdot)$ and with zero costs.

---

31 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 $\theta$ define the value for individual items.