# Screening Two Types

Nima Haghpanah and Ron Siegel*

May 22, 2025

**Abstract**

We characterize profit-maximizing menus in screening settings in which the agent has one of two privately-known types. We assume that utilities are quasi-linear but impose no other restrictions (such as increasing differences) on the agent's utility or the set of alternatives. Our characterization clarifies the role of increasing differences in the standard setting and shows when random menus are beneficial. We describe applications to vertical and horizontal differentiation and multi-product bundling.

## 1 Introduction

Screening settings, in which a profit-maximizing principal faces a privately-informed agent, have been studied extensively in the economics literature, beginning with the seminal contributions of Mussa and Rosen (1978), Riley and Zeckhauser (1983), and Maskin and Riley (1984). The principal designs a menu of (alternative, price) options, and the agent maximizes his (quasi-linear) utility by choosing either an optimal option from the menu or his outside option. The agent's private information (his type) determines his utility from the outside option and the various alternatives.

A leading example is second-degree price discrimination with alternatives that correspond to different quantities or qualities of a product. In this case, it is natural to order the alternatives from worst (lowest quantity/quality) to best (highest quantity/quality) and to assume that the agent's private information concerns his

*Department of Economics, the Pennsylvania State University, University Park, PA 16802 (e-mail: nuh47@psu.edu and rus41@psu.edu). We thank Ian Ball, Alessandro Bonatti, Laura Doval, Axel Niemeyer, Dan Quint, Phil Reny and seminar and conference participants at Brown, Caltech, Harvard-MIT, Johns Hopkins, Purdue, Tel-Aviv, USC, and Wisconsin.

marginal willingness to pay (for quantity/quality). This results in "increasing differences," which states that for any pair of alternatives, higher types are willing to pay more than lower types in order to obtain the better alternative instead of the worse one. Increasing differences is commonly assumed in the screening literature and plays a crucial role in characterizing the optimal (profit maximizing) menus.

Many natural settings, however, involve alternatives and types that are not naturally ordered from "best" to "worst" or from "low" to "high." One example, which we consider in Section 5.1, is product differentiation in which each product has a horizontal value and a vertical value. All types like higher vertical values, but types may differ in their ideal horizontal value.[1] Thus, one type is not necessarily "higher" or "lower" than another, and one product is not necessarily "better" or "worse" than another with a different horizontal value. Which products and at what prices should the principal offer to maximize his profit? And what distortions (relative to efficiency) will the optimal products entail? Another example, which we consider in Section 5.2, is bundling multiple products. Each type has a valuation for each product, and a type's valuation for a bundle is the sum of the type's valuations for the products in the bundle. We impose no restrictions on types' valuations for each product, so one type is not necessarily "higher" or "lower" than another, and one bundle is not necessarily "better" or "worse" than another. Which bundles should the principal offer and at what prices? This question has been considered in the literature, which we discuss in Section 1.2, but has proven difficult to analyze. We return to these examples below.

To analyze such settings, we investigate a screening model in which the agent has one of only two possible types, but his (quasi-linear) utility from the various alternatives is otherwise unrestricted. In particular, we do not assume that the alternatives or the types are ordered in a particular way, so neither type is necessarily "high" or "low." Our main results provide a complete, geometric characterization of the optimal menus.

The key to the characterization is our notion of an uncontested type. A type is uncontested if his valuation for his efficient alternative weakly exceeds the other type's valuation for this alternative. Otherwise, the type is contested. We show that

[1] Standard second-degree price discrimination corresponds to all types having the same ideal horizontal value.

the principal faces no tradeoff between maximizing an uncontested type's valuation for his allocated alternative and minimizing the resulting information rent for the other type. Thus, in any optimal menu an uncontested type obtains his efficient alternative and the other type's payoff is zero, regardless of the probabilities of the two types. In particular, the "first-best" is implementable if and only if both types are uncontested.[2]

We show that at least one of the types is uncontested. So, if the other type is contested, an optimal menu is pinned down by the contested type's alternative. If the probability of the contested type is high, then it is more important to assign him a high-valuation alternative and charge him his valuation than to minimize the uncontested type's resulting information rent. This generates information rent for the uncontested type, so the price he is charged for his efficient alternative is reduced to make him indifferent between his and the contested type's alternative (at their respective prices). As the probability of the uncontested type increases, decreasing his information rent becomes more important, even at the price of further distorting the contested type's allocation away from efficiency. Unlike in the standard settings with increasing differences, where the only possible distortion in the low type's allocation is a "downward" distortion, in our setting there may be many possible directions for the distortion in the contested type's allocation, and our characterization identifies the optimal distortion. If the probability of the uncontested type is high enough, it may be optimal to reduce his information rent to zero. This can be done by having the contested type obtain his most valuable alternative among those he values weakly more than the uncontested type. When this happens, both types are charged their valuations for their allocated alternatives and obtain zero utility, but unlike in the standard setting, neither type is necessarily excluded.[3]

An important special case of our setting is when the set of valuation pairs across alternatives is convex. This arises naturally in some applications, such as our combined vertical and horizontal product differentiation, and also when random alternatives are allowed, because randomization convexifies the set of valuation pairs. Our characterization of the optimal menus provides sharper predictions in this case. It also allows us

[2] The "first-best" refers to selling each type his efficient alternative for a price equal to his valuation for that alternative.

[3] Even though both types obtain zero utility this is not the first-best because the contested type does not obtain his efficient alternative.

to characterize when offering random alternatives in the menu strictly increases the principal's profit.[4] A sufficient condition is that the probability of the uncontested type is sufficiently high and that, for the contested type's most valuable alternative among those he values weakly more than the uncontested type, the contested type's valuation strictly exceeds that of the uncontested type.

We illustrate the potential usefulness of our results with the product differentiation and bundling applications mentioned above. The former combines vertical and horizontal product differentiation and assumes that the cost of production is linear in a product's vertical value (and independent of its horizontal value). Both types are uncontested, so first-best is implementable, if the types' ideal horizontal values are sufficiently different given the types' preference intensities. When this is not the case, we show that the horizontal value of the contested type's allocation is optimally distorted away from both types' ideal horizontal values, and the distortion increases, up to a point, in the probability of the uncontested type. The vertical value of the contested type's allocation is not distorted unless the contested type's preference intensity and probability are sufficiently low. In this case, the contested type is optimally excluded and only the uncontested type is served. When production is costless, the contested type is only excluded if his probability is sufficiently low and the two types' ideal horizontal values are identical. In this sense, the result in the standard price discrimination setting with costless production that the low type is excluded when his probability is sufficiently low is non-generic. Turning to the bundling application, we show that when one of the types is contested, his allocation can optimally be distorted in two ways. First, some products may be removed from the contested type's efficient bundle if they are valued sufficiently highly by the uncontested type. Second, and perhaps unexpectedly, some products that both types dislike may be added to the contested type's bundle if the uncontested type dislikes these products sufficiently more than the contested type. These distortions increase, up to a point, in the probability of the uncontested type.

The rest of the paper is organized as follows. We begin with an illustrative example in Section 1.1. Section 1.2 surveys the related literature. Section 2 describes the model and preliminaries. Section 2.1 reviews the standard setting with increasing differences.

[4] Thanassoulis (2004), Vincent and Manelli (2007), and Daskalakis et al. (2017) provide instances of randomization strictly increasing a multi-product monopolist's profit.

Section 3 analyzes the model, provides the main result, and describes the main parts of the proof. Section 4 studies the case in which the set of valuation pairs is convex, and investigates the optimal menus when the principal can offer random alternatives. Section 5 describes the product differentiation and bundling applications. Section 6 revisits the standard setting with increasing differences and concludes.

### 1.1 An Example

Consider a restaurant frequented by Theorists and Applied theorists, in proportions $1 - q$ and $q \in (0, 1)$, respectively. The chef knows how to prepare two dishes: steak and fish.[5] Applied theorists are willing to pay up to 60 for steak and up to 20 for fish. Theorists are willing to pay up to 30 for steak and up to $v > 0$ for fish.[6] Each customer can eat at most one dish. For different values of $v$, which dish or dishes should the restaurant offer and at what prices to maximize its profit? And how does the answer depend on the proportions of the two customer types? Our main result (Theorem 1) provides a complete answer. We describe the qualitative features here, and give the details of the derivation in Appendix B.

Regardless of the proportions of the two types, it is optimal for the restaurant to offer steak and price it so that Applied theorists order it, and to charge Theorists their entire willingness to pay for the dish they order (if any). Whether and which dish the restaurant should target at Theorists depends on the values of $v$ and $q$. If $v$ is low and $q$ is high, the restaurant should offer only steak at a price equal to Applied theorists' willingness to pay, thereby excluding Theorists. If both $v$ and $q$ are low, the restaurant should offer only steak at a price equal to Theorists' willingness to pay, so all customers order steak. Otherwise, the restaurant should offer both steak and fish, targeting fish at Theorists by setting its price equal to Theorists' willingness to pay for fish and adjusting the price of steak accordingly. Figure 1, (a) summarizes this by describing which dish (if any) is optimally targeted at Theorists for different values of $v$ and $q$.

[5] For simplicity only we assume that the dishes are costless to prepare.

[6] Thus, for $v < 20$, Applied theorists are willing to pay more than Theorists for fish and also to "upgrade" from fish to steak, so Applied theorists can be thought of as "high types" and Theorists can be thought of as "low types." But for $v > 20$, Theorists are willing to pay more than Applied theorists for fish, whereas Applied theorists are willing to pay more than Theorists customers to "upgrade" from fish to steak, so neither type is "high" or "low."

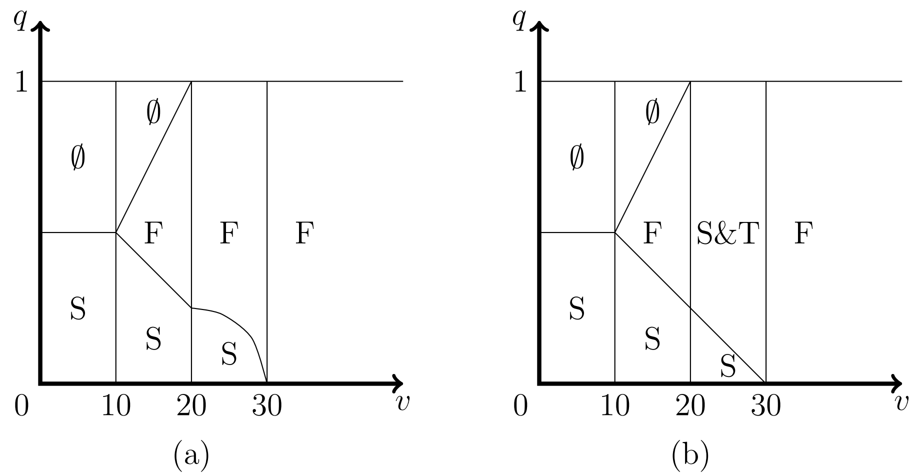



Figure 1: The dish optimally targeted at Theorists when (a) mixed dishes cannot be offered, and (b) mixed dishes can be offered. ∅ indicates exclusion, "S" indicates steak, "F" indicates Fish, and "S&T" indicates a "surf-and-turf" mixed dish.

Now suppose that in addition to steak and fish, the restaurant can offer "surf-and-turf" dishes, which combine steak and fish in proportions that sum up to 1.[7] Our results also identify the optimal menus in this case, and characterize precisely when the restaurant strictly benefits from offering mixed dishes. With mixed dishes, it remains optimal for the restaurant to offer steak and price it so that Applied theorists order it, and to charge Theorists their entire willingness to pay for the dish they order (if any). But for intermediate values of $v$, there is a range of values of $q$ (which increases in $v$), for which it is optimal to offer a "surf-and-turf" mixed dish targeted at Theorists. The optimal proportions of steak and fish in this dish depend only on $v$ and do not change with $q$. Figure 1, (b) describes which dish (if any) is optimally targeted at Theorists for different values of $v$ and $q$ when mixed dishes are allowed.

## 1.2 Related literature

The classical models of monopolistic price discrimination (Mussa and Rosen, 1978 and Maskin and Riley, 1984) allow only for vertically differentiated products and assume increasing differences. In these models, the only possible distortion in a

[7] We assume that customers' utility from a fractional steak or fish dish is linear in the fraction, so mixed dishes can be thought of as "random dishes."

type's allocation is a downward distortion. Our model makes it possible to study the *direction* of the distortion. Our application to vertical and horizontal differentiation shows that the horizontal value of the contested type's allocation may be distorted away from the ideal horizontal value of the uncontested type, and our application to bundling shows that products with a negative valuation might be *added* to the contested type's bundle.

The literature that studies horizontal differentiation of products mostly focuses on imperfect competition, where each firm sells one of the products. The classical models (Hotelling, 1929 and Salop, 1979) allow only for horizontal differentiation (and not vertical differentiation). Villas-Boas (1999) and Armstrong and Vickers (2001) add horizontal differentiation to the vertical differentiation model of Mussa and Rosen (1978) in discrete-choice models in which a random additive shock is added to the value of each product for each consumer. We are not aware of papers that combine horizontal and vertical differentiation in a monopoly setting, as in our application.

Our bundling application is related to the literature on multi-product bundling that started with Stigler (1963) and Adams and Yellen (1976). The theoretical literature highlights that optimal menus might be complex. Optimal menus may be random (Thanassoulis, 2004), include an infinite number of alternatives, (Vincent and Manelli, 2007), and be hard to compute (Daskalakis et al., 2014). As a result, large parts of the literature impose additional structure on the problem, and solutions are known only in very special cases. Even in what is considered the simplest case, two products with additive and independently drawn values, optimal menus can still be complex and are identified only for specific examples (Daskalakis et al., 2014, Thirumulanathan et al., 2019). We simplify the problem in another direction: instead of assuming that there are two products (with independent values), we assume that there are two types, allow for any number of products and any quasi-linear valuations, and fully solve the problem.[8]

Papers that study screening often focus (entirely or partially) on two-type versions of the problems they study. Bergemann et al. (2018) and Boleslavsky and Kolb

[8]Our main result imposes no structure on types' valuations and can be used to identify optimal menus without assuming additivity across products. We impose additivity in the bundling application for two reasons. First, additivity is a canonical assumption made by most of the literature (with exceptions that include (Haghpanah and Hartline, 2021, Ghili, 2023, and Yang, 2024). Second, with additivity we can describe the solution geometrically.

(2025) study selling information structures, Rochet and Stole (2002) study random participation constraints, Pycia (2006) studies bundling products with additive values, and Galperti (2015) studies screening an agent with self-control problems. Using our result, Bergemann et al. (2025) study a two-type version of the problem of designing large language models. These applications impose additional structure on types' utilities beyond quasi-linearity, and our paper clarifies which properties of optimal menus hold generally and (consequently) which ones rely on the additional structure.

# 2 Model and preliminaries

A profit-maximizing principal faces an agent whose privately-known type $t \in T = \{1, 2\}$ is 1 with probability $1 - q$ and 2 with probability $q \in (0, 1)$. There is a set $A$ of alternatives, where alternative $0 \in A$ represents the agent's outside option. An alternative can correspond to a quality/quantity of a product, to a bundle of products, or to a list of characteristics of a product. The agent's type $t$ specifies his valuation $v_t(a) \in \mathbb{R}$ for every alternative $a$. Type $t$'s utility from alternative $a$ and payment $p$ to the principal is $v_t(a) - p$. The principal's cost of producing alternative $a$ is $c(a) \in \mathbb{R}$.

Denote by $V = \{(v_1(a), v_2(a)) : a \in A\}$ the set of valuation pairs across the alternatives and assume that $V$ is closed. Let

$$a_t^{\mathrm{E}} = \arg\max_{a \in A} \big(v_t(a) - c(a)\big) \tag{E1-E2}$$

be the efficient alternative of type $t$, and assume that it exists and is unique.[9] Closedness of $V$ and existence of $a_t^{\mathrm{E}}$ are satisfied for example when $A$ is finite or, more generally, when $A$ is compact and $v_t$ and $c$ are continuous.

The principal designs a menu of $(a, p)$ pairs, where $a \in A$ is an alternative and $p$ is a price. The agent chooses either a pair from the menu or his outside option to maximize his utility. Notice that we require the principal to offer only deterministic alternatives. This is more general than allowing the principal to offer random alternatives, because we can consider random alternatives by convexifying $V$. We discuss random alternatives in Section 4.2.

[9] Existence is important for the existence of optimal menus. Uniqueness is for expositional simplicity only.

By the revelation principle (Myerson, 1979), and because the principal can include the outside option with a price of zero in the menu, it suffices to consider menus $\{(a_1, p_1), (a_2, p_2)\}$ from which each type $t$ optimally chooses the pair $(a_t, p_t)$.[10] Thus, the principal solves the following problem:

$$\begin{aligned}
\max_{a_1,a_2\in A \;\; p_1,p_2\in\mathbb{R}} \quad & (1-q)(p_1 - c(a_1)) + q(p_2 - c(a_2)) \\
\text{subject to} \quad & v_1(a_1) - p_1 \geq v_1(0), \\
& v_2(a_2) - p_2 \geq v_2(0), \\
& v_1(a_1) - p_1 \geq v_1(a_2) - p_2, \\
& v_2(a_2) - p_2 \geq v_2(a_1) - p_1.
\end{aligned}$$

Appendix C shows that we can normalize the production costs of all the alternatives and the valuations for the outside option to zero by changing the valuations for all the alternatives in a way that maintains the same solutions to the principal's problem.[11] Consequently, we will solve the following problem:

$$\begin{aligned}
\max_{a_1,a_2\in A \;\; p_1,p_2\in\mathbb{R}} \quad & (1-q)p_1 + qp_2 \\
\text{subject to} \quad & v_1(a_1) - p_1 \geq 0, && \text{(IR1)} \\
& v_2(a_2) - p_2 \geq 0, && \text{(IR2)} \\
& v_1(a_1) - p_1 \geq v_1(a_2) - p_2, && \text{(IC1)} \\
& v_2(a_2) - p_2 \geq v_2(a_1) - p_1. && \text{(IC2)}
\end{aligned}$$

Figure 2 depicts an example of the normalized set $V$, which may contain valuation pairs with positive or negative values, and highlights the valuation pairs corresponding to the efficient alternatives $a_1^{\mathrm{E}}$ and $a_2^{\mathrm{E}}$, which are the rightmost alternative and the topmost alternative, respectively. Throughout the paper, we sometimes identify an alternative $a$ with its associated valuation pairs $(v_1(a), v_2(a))$ when this should not cause confusion. For example, we can say that in Figure 2 alternatives $a_1^{\mathrm{E}}$ and $a_2^{\mathrm{E}}$ lie in the positive orthant.

[10] It may be that $a_1 = a_2$.

[11] We change type $t$'s valuation for every alternative $a$ to $v_t(a) - c(a) - (v_t(0) - c(0))$, which might be positive or negative. This does not change the efficient alternatives $a_1^{\mathrm{E}}$ and $a_2^{\mathrm{E}}$.

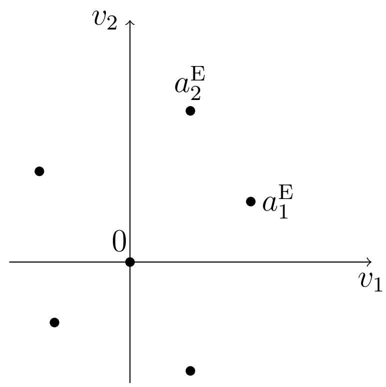



Figure 2: The normalized set $V$ may contain valuation pairs with positive or negative values.

If the principal could observe the agent's type, that is, if we ignored the constraints IC1 and IC2, then the solution would be the "first-best," which assigns to each type his efficient alternative at a price equal to his valuation for that alternative. If the first-best also satisfies IC1 and IC2, then this is the solution to the principal's problem with all the constraints. We then say that the optimal menu implements the first-best, or that the first-best is implementable.

**Definition 1** *A menu* $\{(a_1, p_1), (a_2, p_2)\}$ *implements the first-best if* $a_t = a_t^{\mathrm{E}}$ *and* $p_t = v_t(a_t^{\mathrm{E}})$ *for both types.*

Our main result solves the principal's problem with all the constraints and also shows when the solution implements the first-best.

## 2.1 Increasing Differences

The screening literature usually assumes that the alternatives are ordered, with the outside option being the lowest alternative, and that one of the types, say type 2, values any increase in his allocated alternative strictly more than the other type, type 1. This "increasing differences" condition facilitates the characterization of the optimal menus because it allows us to ignore IR2, which then makes it possible to repeatedly simplify the problem and obtain a formulation that is separable across $a_1$ and $a_2$. This separable formulation implies that type 2 (the "high type") is served efficiently ("efficiency at the top"), that IR1 and IC2 hold as equalities, and that the allocation of type 1 (the "low type") may be distorted downward to reduce type 2's information rent. As the proportion $q$ of type 2 approaches 1, type 1 is optimally excluded or "almost excluded" in that in the space $V$ of valuation pairs type 1's allocation approaches the outside option. In addition, the first-best is implementable if and only if it is efficient to exclude type 1. We formally review these results, as well as a weaker notion of increasing differences, in Appendix D, and highlight that these results crucially rely on the outside option being the lowest alternative.

Without increasing differences, there are no natural "high" and "low" types, distortions can take various forms, and no constraint can be ignored in order to simplify the problem and ultimately obtain a separable formulation, so it is not clear which constraints hold as equalities at the optimum. Thus, characterizing the optimal menus requires a different approach, which we now present.

# 3 Characterizing Optimal Menus

Our main result (Theorem 1 below) characterizes the optimal menus. As we will see, the characterization hinges on the location of the efficient alternatives $a_1^{\mathrm{E}}$ and $a_2^{\mathrm{E}}$ relative to the 45 degree line through the outside option $(0, 0)$. This line will feature prominently in our analysis, and we will henceforth refer to it as "the diagonal." Alternatives that lie above the diagonal are those for which the valuation of type 2 exceeds that of type 1, and the reverse holds for alternatives that lie below the diagonal.

Whether a type's efficient alternative lies above or below the diagonal corresponds to whether the type is *uncontested*, as given by the following definition.

**Definition 2** *Type $t$ is uncontested if $v_t(a_t^{\mathrm{E}}) \geq v_{t'}(a_t^{\mathrm{E}})$, where $t' \neq t$. Otherwise, type $t$ is contested.*

An uncontested type is willing to pay weakly more for his efficient alternative than the other type does.[12] Type 1 is uncontested if and only if $a_1^{\mathrm{E}}$ lies weakly below the diagonal; type 2 is uncontested if and only if $a_2^{\mathrm{E}}$ lies weakly above the diagonal.

We observe that at least one type is uncontested. This is because alternative $a_2^{\mathrm{E}}$ lies on a weakly higher 45 degree line than alternative $a_1^{\mathrm{E}}$, that is, $v_2(a_2^{\mathrm{E}}) - v_1(a_2^{\mathrm{E}}) \geq$

[12]The term "uncontested" is motivated by the observation that such a type would outbid the other type in a standard auction for the uncontested type's efficient alternative.

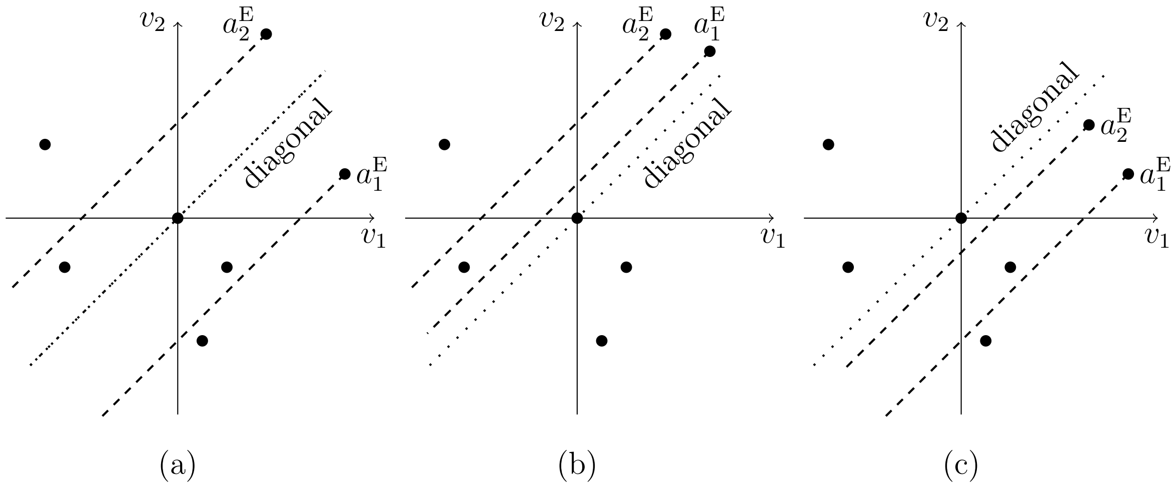



Figure 3: The three possibilities for the positions of the two efficient alternatives. (a) Both types are uncontested. (b) Type 1 is contested and type 2 is uncontested. (c) Type 1 is uncontested and type 2 is contested.

$v_2(a_1^{\mathrm{E}}) - v_1(a_1^{\mathrm{E}})$.[13] Thus, there are three possibilities, depicted in Figure 3: (a) Both types are uncontested ($a_2^{\mathrm{E}}$ lies weakly above the diagonal and $a_1^{\mathrm{E}}$ lies weakly below the diagonal), (b) type 1 is contested and type 2 is uncontested (both alternatives lie strictly above the diagonal), or (c) type 1 is uncontested and type 2 is contested (both alternatives lie strictly below the diagonal). In cases (a) and (c), type 1 is uncontested; in cases (a) and (b), type 2 is uncontested. Without loss of generality, we assume throughout our analysis that type 2 is uncontested. That is, if we are in case (c), switch the labels of the types to obtain case (b). This is formalized by the following assumption, which we maintain for the remainder of the paper.

**Assumption 1** *Without loss of generality, type 2 is uncontested ($v_2(a_2^{\mathrm{E}}) \geq v_1(a_2^{\mathrm{E}})$). That is, either Figure 3, (a) or Figure 3, (b) applies.*

## 3.1 No Tradeoff Lemma

Uncontested types play a crucial role in our characterization of the optimal menus. This is because if a type is uncontested, then the principal does not face a tradeoff between maximizing the surplus that can be extracted from this type and minimizing

[13]This follows because, by definition of $a_1^{\mathrm{E}}$ and $a_2^{\mathrm{E}}$, $v_2(a_2^{\mathrm{E}}) \geq v_2(a_1^{\mathrm{E}})$ and $v_1(a_1^{\mathrm{E}}) \geq v_1(a_2^{\mathrm{E}})$.

the information rent of the other type. This is the content of the following lemma, which shows that in any optimal menu an uncontested type's allocation is efficient and the other type obtains no information rent, regardless of the probabilities of the two types.

**Lemma 1 (No Tradeoff Lemma)** *If a type is uncontested, then in any optimal menu this type's allocation is efficient and the other type's IR constraint holds as an equality.*

**Proof.** Consider an uncontested type, without loss of generality type 2 (by Assumption 1), and an optimal menu $\{(a_1, p_1), (a_2, p_2)\}$. Suppose that the allocation of type 2 is not efficient ($a_2 \neq a_2^{\mathrm{E}}$). We will show that the menu is not optimal by changing type 2's allocation to his efficient alternative $a_2^{\mathrm{E}}$ and increasing his payment by the increase $v_2(a_2^{\mathrm{E}}) - v_2(a_2) > 0$ in his valuation (to leave his utility unchanged). This change increases the principal's revenue (from type 2), provided that none of the constraints are violated.

Clearly, the only constraint that may be violated because of the change is IC1. We will consider three cases based the locations of $a_1$ and $a_2$ relative to the diagonal, show that IC1 is maintained in the first two cases, and show that if IC1 fails in the third case, then a modification of type 1's allocation $a_1$ that overcomes this failure further increases the principal's revenue (from type 1).[14]

Case 1: $a_2$ lies strictly below the diagonal: $v_2(a_2) < v_1(a_2)$.
Because type 2 is uncontested, $v_2(a_2^{\mathrm{E}}) \geq v_1(a_2^{\mathrm{E}})$. Therefore, $v_2(a_2) < v_1(a_2)$ implies that the increase $v_2(a_2^{\mathrm{E}}) - v_2(a_2)$ in type 2's payment exceeds the increase $v_1(a_2^{\mathrm{E}}) - v_1(a_2)$ in type 1's valuation for type 2's allocated alternative. This relaxes IC1, so IC1 is maintained.

Case 2: $a_2$ lies weakly above and $a_1$ lies weakly below the diagonal: $v_2(a_2) \geq v_1(a_2)$ and $v_2(a_1) \leq v_1(a_1)$.
We first observe that IR2 holds as an equality in the menu $\{(a_1, p_1), (a_2, p_2)\}$. Intuitively, because $v_1(a_1) \geq v_2(a_1)$, there is no reason to give type 2 any information

[14] Equivalently, we can think of the principal as offering the modified menu $\{(a_1, p_1), (a_2^{\mathrm{E}}, p_2 + v_2(a_2^{\mathrm{E}}) - v_2(a_2))\}$, from which type 2 chooses $(a_2^{\mathrm{E}}, p_2 + v_2(a_2^{\mathrm{E}}) - v_2(a_2))$, which increases the principal's revenue. Type 1 either continues to choose $(a_1, p_1)$ or chooses $(a_2^{\mathrm{E}}, p_2 + v_2(a_2^{\mathrm{E}}) - v_2(a_2))$, which we show further increases the principal's revenue.

rent. To see this, suppose that IR2 holds as a strict inequality. Then, by optimality of the menu, IC2 and IR1 hold as equalities, so

$$v_2(a_2) - p_2 = v_2(a_1) - p_1 = v_2(a_1) - v_1(a_1) \leq 0,$$

where the first equality follows from IC2 and the second equality follows from IR1 holding as equalities. That $v_2(a_2) - p_2 \leq 0$ contradicts that IR2 holds as a strict inequality, so IR2 holds as an equality.

Now, because IR2 holds as an equality and the change in the menu leaves type 2's utility unchanged, his new payment is $v_2(a_2^{\mathrm{E}})$. Thus, by mimicking type 2, type 1 obtains utility

$$v_1(a_2^{\mathrm{E}}) - p_2 = v_2(a_2^{\mathrm{E}}) - v_2(a_2^{\mathrm{E}}) \leq 0,$$

where the inequality follows because type 2 is uncontested, so IC1 is maintained.

Case 3: $a_2$ lies weakly above and $a_1$ lies strictly above the diagonal: $v_2(a_2) \geq v_1(a_2)$ and $v_2(a_1) > v_1(a_1)$.

That $v_2(a_1) > v_1(a_1)$ implies that IR2 holds as a strict inequality in the menu $\{(a_1, p_1), (a_2, p_2)\}$. This is because

$$v_2(a_2) - p_2 \geq v_2(a_1) - p_1 \geq v_2(a_1) - v_1(a_1) > 0,$$

where the first inequality is IC2 and the second inequality follows from IR1. That IR2 holds as a strict inequality implies (by the assumed optimality of the menu) that IC2 and IR1 hold as equalities, so the first two of the above inequalities are equalities and type 2's utility $v_2(a_2)-p_2$ is equal to $v_2(a_1)-v_1(a_1)$. Thus, to maintain type 2's utility after changing his allocation to $a_2^{\mathrm{E}}$, his payment is increased to $v_2(a_2^{\mathrm{E}}) - (v_2(a_1) - v_1(a_1))$. This payment exceeds $p_1 = v_1(a_1)$, because $v_2(a_2^{\mathrm{E}}) \geq v_2(a_1)$ by definition of $a_2^{\mathrm{E}}$. So if IC1 is violated we can replace type 1's allocation and payment with type 2's modified allocation and payment, that is, replace $(a_1, p_1)$ with $(a_2^{\mathrm{E}}, v_2(a_2^{\mathrm{E}}) - (v_2(a_1) - v_1(a_1)))$ . This change maintains all the constraints and weakly increases the principal's profit.

Having established that $a_2 = a_2^{\mathrm{E}}$, we now show that IR1 holds as an equality. Because type 2 is uncontested, $v_2(a_2^{\mathrm{E}}) \geq v_1(a_2^{\mathrm{E}})$, so there is no reason to give type 1 any information rent. Indeed, suppose that IR1 holds as a strict inequality in the menu $\{(a_1, p_1), (a_2, p_2)\}$. Then, by optimality of the menu, IC1 and IR2 hold as equalities, so

$$v_1(a_1) - p_1 = v_1(a_2^{\mathrm{E}}) - p_2 = v_1(a_2^{\mathrm{E}}) - v_2(a_2^{\mathrm{E}}) \leq 0,$$

where the first equality follows from IC1 and the second equality follows from IR2 holding as equalities. That $v_1(a_1) - p_1 \leq 0$ contradicts that IR1 holds as a strict inequality, so IR1 holds as an equality. This completes the proof of the lemma. ■

## 3.2 Informal Description of the Characterization

Lemma 1 implies that if both types are uncontested, then the optimal menu implements the first-best, because both types' allocations are efficient and both types obtain no information rent. If only one type is uncontested, then the first-best violates the IC constraint of the uncontested type (because the other type is contested). It thus remains to characterize the optimal menus assuming that only type 2 is uncontested (by Assumption 1). Lemma 1 reduces this task to identifying the allocation $a_1$ of type 1. Type 1's allocation optimally resolves the tradeoff between extracting revenue from type 1 and minimizing the information of type 2. The resolution hinges on the relative probabilities of the two types. If the probability of type 1 is high, then revenue extraction is more important, so $a_1$ maximizes (or nearly maximizes) type 1's valuation. In other words, $a_1$ is efficient (or close to efficient) for type 1. This generates information rent for type 2 (because type 1 is contested).[15] As the probability of type 2 increases, reducing his information rent becomes more important, even at the expense of reducing the revenue extracted from type 1. This leads to distorting $a_1$ away from efficiency in order to reduce type 2's information rent. If the probability of type 2 is such that it is optimal to reduce his information rent to zero, so IR2 holds as an equality, it may be possible to do so without excluding type 1, unlike with increasing differences. Indeed, any alternative that type 1 values weakly more than type 2 can be allocated to type 1 at a price equal to type 1's valuation without generating information rent for type 2. These are the alternatives that lie weakly below the diagonal in the space $V$ of valuation pairs. Therefore, among these alternatives it is optimal to allocate to type 1 the alternative he values the most. Thus, type 1 is excluded for a high enough probability of type 2 if and only if the

[15]This means that IR2 is slack so IC2 holds as an equality.

alternative type 1 values the most among those weakly below the diagonal coincides with the outside option. If IR2 holds as an equality, then both types' utilities are zero and IC2 is slack if and only if this alternative lies strictly below the diagonal. This is not the first-best, however, if type 1 is not served efficiently.

For a geometric interpretation, notice that because type 1 obtains no information rent (Lemma 1), his payment is equal to his valuation for his allocated alternative $a_1$. This implies that if $v_2(a_1) > v_1(a_1)$, then by mimicking type 1, type 2 can obtain utility $v_2(a_1) - v_1(a_1)$, so this is type 2's information rent. In the space of valuation pairs, the difference between the types' valuations for an alternative corresponds to the 45 degree line on which the alterative lies. A lower line corresponds to a smaller difference, or less information rent for type 2. This difference is positive if the alternative lies above the diagonal, which is the case for $a_1^{\mathrm{E}}$ if type 1 is contested (by definition). In this case, as the probability of type 2 increases, it becomes optimal to choose an alternative with a lower valuation for type 1 that lies on a lower 45 degree line.[16] This corresponds to finding an alternative that is maximal in the direction of the vector $(1, -q)$. For sufficiently high probability of type 2, this alternative may lie on or below the diagonal, at which point type 2 obtains no information rent by mimicking type 1 (IC2 no longer holds as an equality). It is then optimal to allocate to type 1 the alternative $a_1$ he values the most among those weakly below the diagonal, and to charge type 2 his entire valuation $v_2(a_2^{\mathrm{E}})$, which eliminates type 2's information rent. This case can also be captured by maximizing in the direction of the vector $(1, -q)$ when a "virtual alternative" that corresponds to the projection of $a_1$ on the diagonal is added to the space of valuation pairs.

## 3.3 Formalizing the Characterization

To formalize this discussion and state the characterization precisely, we introduce some additional notation. Let $A^- = \{a : v_2(a) \leq v_1(a)\}$ be the set of alternatives that lie weakly below the diagonal, and let

$$a^{\mathrm{BD1}} = \operatorname*{argmax}_{a \in A^-} \quad v_1(a) \tag{BD1}$$

[16] If an alternative with a higher valuation for type 1 lies on a lower 45 degree line, that alternative would have already been chosen because it leads to both higher revenue extraction from type 1 and lower information rent for type 2.

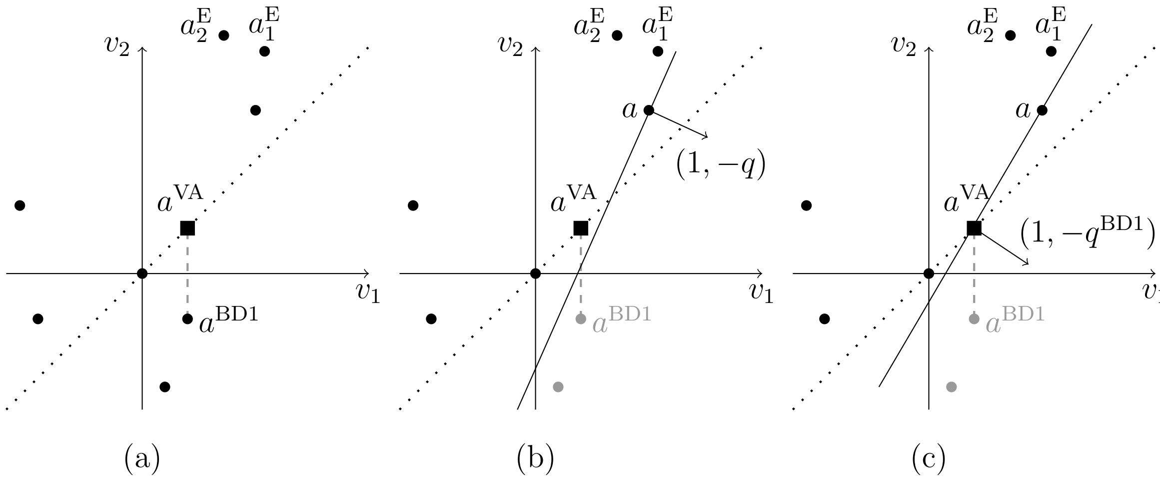



Figure 4: (a) Alternatives $a^{\mathrm{BD1}}$ and $a^{\mathrm{VA}}$. (b) Alternative $a$ maximizes $v_1(\cdot) - qv_2(\cdot)$ among the alternatives in $A^+ \cup \{a^{\mathrm{VA}}\}$, and $A_1'(q) = \{a\}$. (c) $q^{\mathrm{BD1}}$ is the lowest $q$ for which $a^{\mathrm{VA}} \in A_1(q)$, because for $q < q^{\mathrm{BD1}}$ alternative $a$ is farther than $a^{\mathrm{VA}}$ in the direction of the vector $(1, -q)$.

be the alternative that type 1 values the most among those alternatives (BD1 stands for "best alternative below the diagonal for type 1"), and suppose for expositional simplicity only that this alternative is unique.[17] Now, define the "virtual alternative" $a^{\mathrm{VA}}$ to satisfy

$$v_1(a^{\mathrm{VA}}) = v_2(a^{\mathrm{VA}}) = v_1(a^{\mathrm{BD1}}). \tag{V}$$

In particular, if $a^{\mathrm{BD1}}$ lies *on* the diagonal, so $v_1(a^{\mathrm{BD1}}) = v_2(a^{\mathrm{BD1}})$, then $a^{\mathrm{VA}}$ coincides with $a^{\mathrm{BD1}}$. Otherwise, that is, if $v_1(a^{\mathrm{BD1}}) > v_2(a^{\mathrm{BD1}})$, then $a^{\mathrm{VA}}$ is the projection of $a^{\mathrm{BD1}}$ on the diagonal. Alternatives $a^{\mathrm{BD1}}$ and $a^{\mathrm{VA}}$ are shown in Figure 4, (a).

Let $A^+ = \{a : v_1(a) \le v_2(a)\}$ be the set of alternatives that lie weakly above the diagonal. For any $q \in [0, 1]$, let $A_1'(q)$ be the set of alternatives in $A^+ \cup \{a^{\mathrm{VA}}\}$ that are maximal in the direction of the vector $(1, -q)$, as illustrated in Figure 4, (b),

$$A_1'(q) = \operatorname*{argmax}_{a \in A^+ \cup \{a^{\mathrm{VA}}\}} v_1(a) - qv_2(a).$$

Let $A_1(q) \subseteq A^+ \cup \{a^{\mathrm{BD1}}\}$ be the same as $A_1'(q)$ but with $a^{\mathrm{VA}}$ replaced by $a^{\mathrm{BD1}}$ if $a^{\mathrm{VA}}$

[17]Alternative $a^{\mathrm{BD1}}$ always exists because $v_1(0) = v_2(0) = 0$, so $A^-$ is not empty, $V$ is closed by assumption, and type 1's valuation is at most $v_1(a_1^{\mathrm{E}})$.

is in $A'_1(q)$. That is,

$$A_1(q) = \begin{cases} A'_1(q) & \text{if } A'_1(q) \subset A^+ \\ A'_1(q) \setminus \{a^{\mathrm{VA}}\} \cup \{a^{\mathrm{BD1}}\} & \text{otherwise.} \end{cases}$$

Finally, let

$$q^{\mathrm{BD1}} = \min\{q \mid a^{\mathrm{BD1}} \in A_1(q)\} \tag{q-BD1}$$

be the lowest probability $q$ such that $a^{\mathrm{BD1}} \in A_1(q)$, with $q^{\mathrm{BD1}} = 1$ if $a^{\mathrm{BD1}} \notin A_1(q)$ for every $q < 1$.[18] Threshold $q^{\mathrm{BD1}}$ is illustrated in Figure 4, (c). We now state our main result, whose proof is in Appendix A.

**Theorem 1** *Suppose without loss of generality that type 2 is uncontested (Assumption 1). Then, the optimal menu implements the first-best if and only if type 1 is also uncontested. For any probability $q \in (0, 1)$ of type 2, a menu $\{(a_1, p_1), (a_2, p_2)\}$ is optimal if and only if it satisfies the following properties:*

1. *Type 2's allocation is efficient ($a_2 = a_2^{\mathrm{E}}$).*
2. *Type 1's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(a_1).$$

3. *Type 1's alternative $a_1$ is any alternative in $A_1(q)$. Moreover,*
   (a) *If $q < q^{\mathrm{BD1}}$, then $A_1(q) \subset A^+ \setminus \{a^{\mathrm{BD1}}\}$.*
   (b) *If $q > q^{\mathrm{BD1}}$, then $A_1(q) = \{a^{\mathrm{BD1}}\}$.*
4. *If $a_1 \in A^+$, then type 2's incentive constraint IC2 holds as an equality:*

$$v_2(a_2) - p_2 = v_2(a_1) - p_1.$$

5. *If $a_1 = a^{\mathrm{BD1}}$, then type 2's individual rationality constraint IR2 holds as an equality:*

$$p_2 = v_2(a_2^{\mathrm{E}}),$$

[18] Threshold $q^{\mathrm{BD1}}$ is well defined by continuity of $v_1 - qv_2$ in $q$.

*so*

$$a_1 = a^{\mathrm{BD1}}, p_1 = v_1(a^{\mathrm{BD1}}), a_2 = a_2^{\mathrm{E}}, p_2 = v_2(a_2^{\mathrm{E}}).$$

*In particular, type 1's allocation $a_1$ pins down the menu and:*

6. *The resulting profit is $v_1(a_1) - qv_2(a_1) + qv_2(a_2^{\mathrm{E}})$ if $a_1 \in A^+$ and $(1-q)v_1(a^{\mathrm{BD1}}) + qv_2(a_2^{\mathrm{E}})$ if $a_1 = a^{\mathrm{BD1}}$.*

7. *Both $v_1(a_1)$ and $v_2(a_1)$ weakly decrease in $q$.*

8. *The utility of type 2 weakly decreases in $q$.*

9. *As $q$ approaches 1, type 1 is excluded or "almost excluded," that is, $(v_1(a_1), v_2(a_2))$ approaches $(0, 0)$, if and only if $a^{\mathrm{BD1}}$ coincides with the outside option.*

Theorem 1 shows that the first-best is implementable if and only if both types are uncontested, in which case $a^{\mathrm{BD1}} = a_1^{\mathrm{E}}$. Moreover, in any optimal menu type 2 obtains his efficient alternative, IR1 holds as an equality, and type 1's alternative $a_1$ is maximal in the direction $(1, -q)$ among those in $A^+ \cup \{a^{\mathrm{VA}}\}$, with $a^{\mathrm{VA}}$ replaced by $a^{\mathrm{BD1}}$ if $a^{\mathrm{VA}}$ is a maximizer. If $q < q^{\mathrm{BD1}}$, so $a_1 \in A^+$, then IC2 holds as an equality. But if If $q > q^{\mathrm{BD1}}$, so $a_1 = a^{\mathrm{BD1}}$, then IR2 holds as an equality (and IC2 may be slack). If $a_1 = a^{\mathrm{BD1}}$ and $a^{\mathrm{BD1}} \in A^+$ (so $a^{\mathrm{BD1}}$ is on the diagonal), then both IC2 and IR2 hold as equalities. And if IR2 holds as an equality for some probability $q$ of type 2, then the same is true for all higher probabilities of type 2, and the same menu $\{(a^{\mathrm{BD1}}, v_1(a^{\mathrm{BD1}})), (a_2^{\mathrm{E}}, v_2(a_2^{\mathrm{E}}))\}$ remains optimal.

# 4 Convexity and Randomization

In this section we discuss several corollaries of Theorem 1. The first corollary considers the case in which the set $V$ of valuation pairs is convex. The second corollary considers the case in which the principal can offer random alternatives (lotteries over alternatives). We use these corollaries in our applications in Section 5. The third corollary characterizes when random alternatives strictly increase the principal's profit.

## 4.1 A Convex Set of Valuation Pairs

The set $V$ of valuations is convex in some applications, such as the vertical and horizontal product differentiation in Section 5.1, and when random alternatives are allowed (we consider this in the next subsection). Convexity leads to sharper predictions and a simplified formulation of Theorem 1. The reason is that with convexity, if type 1 is contested (so the first-best is not implementable) to identify type 1's alternative $a_1$ in an optimal menu we can maximize $v_1(a) - qv_2(a)$ over the set of *all* alternatives $A$ (and compare this maximum value to $v_1(a^{BD1}) - qv_2(a^{BD1})$), instead of maximizing over the constrained set $A^+ \cup \{a^{\mathrm{VA}}\}$. This is because (i) the alternative $a^{\mathrm{BD1}}$ lies on (and not below) the diagonal, so $a^{\mathrm{BD1}} = a^{\mathrm{VA}}$, and (ii) as $q$ increases, the first alternative $a \in A^-$ that maximizes $v_1(a) - qv_2(a)$ over the set of all alternatives $A$ is $a = a^{\mathrm{BD1}}$.[19] This also implies that for any $q$, in any optimal menu IC2 holds as an equality.

To formalize this, suppose that $V$ is convex, type 1 is contested, so the first-best is not implementable, and, without loss of generality, type 2 is uncontested (Assumption 1). We will relate type 1's alternative in an optimal menu to the set

$$B_1(q) = \operatorname*{argmax}_{a \in A} v_1(a) - qv_2(a)$$

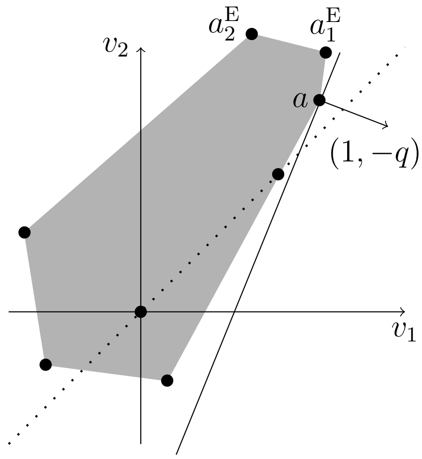
of all the alternatives that are maximal in the direction of the vector $(1, -q)$, as shown in Figure 5, (a). For this, we consider the threshold $q^{\mathrm{OD}} \in [0, 1]$, which is the lowest value $q$ such that some alternative in $B_1(q)$ lies on the diagonal (OD stands for "on the diagonal").[20]

Formally, let $A^= = \{a : v_2(a) = v_1(a)\}$ be the set of alternatives on the diagonal, and let

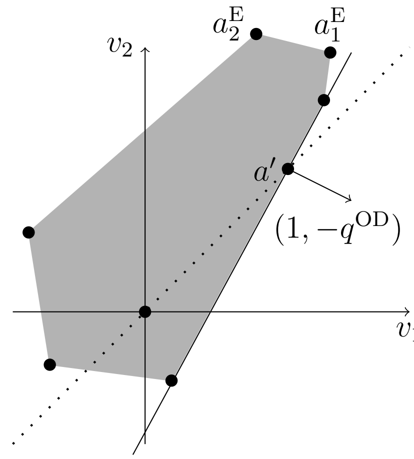
$$q^{\mathrm{OD}} = \min\{q \mid B_1(q) \cap A^= \neq \emptyset\}. \tag{OD}$$

[19] As we show in Section 4.3, this is precisely when allowing for random alternatives does not increase the principal's profit.

[20] This threshold exists because $B_1(0) = \operatorname*{argmax}_{a \in A} v_1(a) = \{a_1^{\mathrm{E}}\}$ lies weakly above the diagonal, the outside option is farther in the direction of the vector $(1, -1)$ than any alternative above the diagonal, so the alternatives in $B_1(1)$ lie weakly below the diagonal, and as $q$ increases the alternatives in $B_1(q)$ move clockwise along the boundary of $V$.




(a)




(b)

Figure 5: (a) Alternative $a$ is maximal in the direction of the vector $(1, -q)$. (b) Alternative $a'$ is an alternative in $B_1(q^{\mathrm{OD}})$ that lies on the diagonal.

Threshold $q^{\mathrm{OD}}$ is shown in Figure 5, (b).

The following corollary, whose proof is in Appendix E, shows that $q^{\mathrm{OD}}$ is well defined and uses it to characterize the optimal menus when the set $V$ is convex.

**Corollary 1** *Suppose that type 2 is uncontested (Assumption 1), type 1 is contested, and $V$ is convex. Then, the threshold $q^{\mathrm{OD}}$ defined in* (OD) *exists and $q^{\mathrm{OD}} = q^{\mathrm{BD1}}$, where $q^{\mathrm{BD1}}$ is defined in* (q-BD1). *For any probability $q \in (0, 1)$ of type 2, a menu $\{(a_1, p_1), (a_2, p_2)\}$ is optimal if and only if it satisfies the following properties:*

1. *Type 2's allocation is efficient ($a_2 = a_2^{\mathrm{E}}$).*

2. *Type 1's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(a_1).$$

3. *Type 1's alternative $a_1$ and type 2's binding constraints are specified given $q$ as follows:*

   *(a) If $q < q^{\mathrm{OD}}$, then $a_1 \in B_1(q)$ and IC2 holds as an equality.*

(b) *If* $q = q^{\text{OD}}$*, then* $a_1 \in B_1(q^{\text{OD}}) \cap A^+$ *and IC2 holds as an equality. If* $a_1 = a^{\text{BD1}}$ *(the unique alternative in* $B_1(q^{\text{OD}}) \cap A^+ \cap A^-$*), then IR2 also holds as an equality.*

*One particular* $a_1 \in B_1(q^{\text{OD}}) \cap A^+$ *is the alternative in* $B_1(q^{\text{OD}})$ *that maximizes* $v_2 - v_1$*.*

(c) *If* $q > q^{\text{OD}}$*, then* $a_1 = a^{\text{BD1}}$ *and both IC2 and IR2 hold as equalities.*

*Properties 7-9 from Theorem 1 hold as stated, and property 6 can be strengthened to the following:*

6. *The resulting profit is* $v_1(a_1) - qv_2(a_1) + qv_2(a_2^{\text{E}})$*,*

*and, in addition, we have the following property:*

10. *As* $q$ *increases from* 0 *to* 1*,* $(v_1(a_1), v_2(a_2))$ *traces the boundary of* $V$ *clockwise from* $(v_1(a_1^{\text{E}}), v_2(a_1^{\text{E}}))$ *to* $(v_1(a^{\text{BD1}}), v_2(a^{\text{BD1}}))$*.*

## 4.2 Allowing for Randomization

So far we did not allow the principal to offer distributions over alternatives, or "random alternatives." Accommodating random alternatives is straightforward, because adding random alternatives convexifies $V$. Thus, we can apply our analysis to the set of random alternatives and the corresponding convex hull of $V$. Doing so also allows us in Section 4.3 to characterize when random alternatives strictly increase the principal's profit.

We use Corollary 1 to characterize, for a given set of alternatives $A$, optimal menus that can include random alternatives. A random alternative is an element of $\Delta(A)$, and type $t$'s valuation for random alternative $x$ is $v_t(x) = E_{a\sim x}[v_t(a)]$.[21] We assume that type 2 is uncontested (Assumption 1), as we did in Theorem 1, and that type 1 is contested (otherwise the first-best is implementable and randomization has no effect on the optimal menus).

[21] The principal's cost of producing $x$ is $c(x) = E_{a\sim x}[c(a)]$, which is equal to 0 because we normalized the production costs of all alternatives to 0.

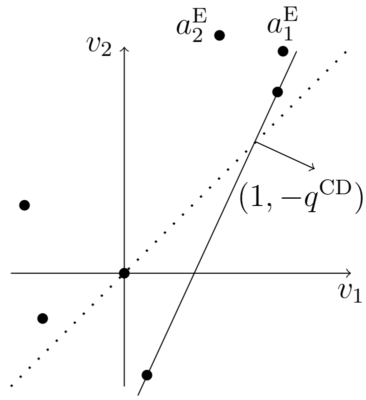



Figure 6: There is no $q$ such that $B_1(q)$ contains an alternative on the diagonal, but $B_1(q^{\mathrm{CD}})$ contains an alternative above and an alternative below the diagonal.

We begin with a generalization of the definition of the threshold $q^{\mathrm{OD}}$ from (OD) that does not require $V$ to be convex. Recall from Section 4.1 that

$$B_1(q) = \underset{a \in A}{\operatorname{argmax}}\ v_1(a) - qv_2(a)$$

is the set of alternatives that are maximal in the direction of the vector $(1, -q)$, and for a convex $V$, $q^{\mathrm{OD}}$ is the lowest threshold at which an alternative in $B_1(q^{\mathrm{OD}})$ lies on the diagonal. With a non-convex $V$, there may not be a maximizer on the diagonal for any $q$, because the alternatives in $B_1(q)$ might "jump" at some point from above the diagonal to below it, as Figure 6 shows. Nevertheless, there exists a threshold at which $B_1(q)$ "crosses" the diagonal, that is, some alternative in $B_1(q)$ lies weakly above the diagonal and some alternative in $B_1(q)$ lies weakly below the diagonal. Formally, let

$$q^{\mathrm{CD}} = \min\{q \mid B_1(q) \cap A^+ \neq \emptyset \text{ and } B_1(q) \cap A^- \neq \emptyset\}. \tag{CD}$$

This threshold $q^{\mathrm{CD}}$ is shown in Figure 6 (CD stands for "crosses the diagonal") and coincides with the threshold $q^{\mathrm{BD1}}$ from Theorem 1 when $q^{\mathrm{BD1}}$ is computed with respect to the set of random alternatives. The following corollary, whose proof is in Appendix F, characterizes the optimal menus when randomization is allowed.

**Corollary 2** *Suppose that type 2 is uncontested (Assumption 1), type 1 is contested, and randomization is allowed. Then, the threshold* $q^{\mathrm{CD}}$ *defined in* (CD) *exists and*

$q^{\mathrm{CD}} = q^{\mathrm{BD1}}$, *where* $q^{\mathrm{BD1}}$, *defined in* (q-BD1), *is computed with respect to the set of random alternatives. For any probability* $q \in (0,1)$ *of type 2, a menu* $\{(x_1, p_1), (x_2, p_2)\}$ *of random alternatives is optimal if and only if it satisfies the following properties:*

1. *Type 2's allocation is efficient* ($x_2 = a_2^{\mathrm{E}}$)*.*

2. *Type 1's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(x_1).$$

3. *Type 1's random alternative* $x_1$ *and type 2's binding constraints are specified given* $q$ *as follows:*

   (a) *If* $q < q^{\mathrm{CD}}$, *then* $x_1$ *is any distribution over the alternatives in* $B_1(q)$*, and IC2 holds as an equality.*

   (b) *If* $q = q^{\mathrm{CD}}$, *then* $x_1$ *is any distribution over the alternatives in* $B_1(q^{\mathrm{CD}})$ *such that*

$$v_2(x_1) \geq v_1(x_1),$$

   *and IC2 holds as an equality. If the above inequality holds as an equality, then IR2 also holds as an equality.*
   *One particular* $x_1$ *that satisfies the above inequality assigns probability 1 to the alternative* $a$ *in* $B_1(q^{\mathrm{CD}})$ *that maximizes* $v_2 - v_1$*.*

   (c) *If* $q > q^{\mathrm{CD}}$, *then* $x_1$ *is any distribution over the alternatives in* $B_1(q^{\mathrm{CD}})$ *such that*

$$v_2(x_1) = v_1(x_1),$$

   *and both IC2 and IR2 hold as equalities.*
   *One particular* $x_1$ *that satisfies the above equality randomizes only over the two alternatives in* $B_1(q^{\mathrm{CD}})$ *that maximize and minimize* $v_2 - v_1$ *(if the two alternatives coincide, then* $x_1$ *assigns probability 1 to this alternative).*

Notice that if $q > q^{\mathrm{CD}}$, then (by property 3(c) of Corollary 2) there may be multiple optimal random alternatives $x_1$, whereas property 3(c) of Corollary 1 states that $a^{\mathrm{BD1}}$

is the unique optimal alternative when $V$ is convex (without randomization). This multiplicity with randomization arises because the principal and the agent do not distinguish between two random alternatives $x \neq x'$ with $v_1(x) = v_1(x')$ and $v_2(x) = v_2(x')$. Except for knife edge cases, however, there will be only two alternatives in $B_1(q^{\text{CD}})$, so for $q > q^{\text{CD}}$ the optimal random alternative $x_1$ is uniquely achieved by appropriately randomizing over these two alternatives.[22]

## 4.3 When does Randomization Help?

We now turn to the question of when offering random alternatives strictly increases the principal's profit.[23] If the set $V$ of valuations pairs is convex, then allowing for random alternatives clearly has no effect. Otherwise, allowing for random alternatives expands the set of menus the principal can offer. This expansion clearly does not affect the principal's profit if the first-best is implementable.

The following corollary characterizes when allowing for random alternatives strictly increases the principal's profit, assuming the first-best is not implementable.

**Corollary 3** *Suppose that type 2 is uncontested (Assumption 1) and type 1 is contested. The following statements are equivalent.*

1. $a^{\text{VA}} = a^{\text{BD1}}$ *and they are both on the boundary of the convex hull of* $V$*.*

2. *The highest point on the diagonal that is also in the convex hull of* $V$ *corresponds to an (non-random) alternative in* $V$*.*

3. *Randomization does not help (does not increase the principal's profit) for any* $q$*.*

*In addition, if either (both) of statements (1)-(2) are violated, then* $q^{\text{CD}} < 1$ *and randomization helps if and only if* $q > q^{\text{CD}}$*.*

Statement (1) of Corollary 3 implies that a sufficient condition for randomization to help (for $q > q^{\text{CD}}$) is that $a^{\text{VA}} \neq a^{\text{BD1}}$, which means that $a^{\text{BD1}}$ lies strictly below

[22] For example, this is the case with probability 1 if the number of alternatives is finite and the agent's valuations for the alternatives are drawn independently across alternatives from a continuous distribution over a compact set.

[23] Thanassoulis (2004), Vincent and Manelli (2007), and Daskalakis et al. (2017) show that this may happen with a multi-product monopolist.

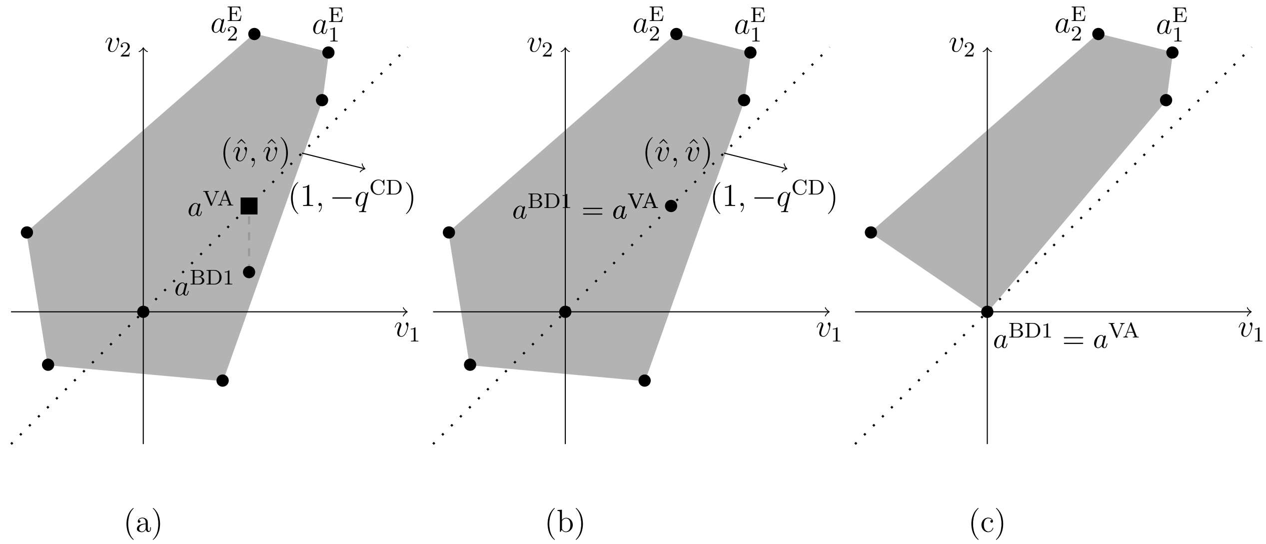



Figure 7: (a) If $a^{\mathrm{BD1}}$ lies strictly below the diagonal, then randomization helps. (b) Randomization helps even though $a^{\mathrm{BD1}}$ lies on the diagonal, as long as $a^{\mathrm{BD1}} \notin B_1(q^{\mathrm{CD}})$. (c) Randomization does not help if no alternative lies strictly below the diagonal.

the diagonal. This is illustrated in Figure 7, (a). But even if $a^{\mathrm{VA}} = a^{\mathrm{BD1}}$, so $a^{\mathrm{BD1}}$ lies on the diagonal, randomization might still help because $a^{\mathrm{VA}} = a^{\mathrm{BD1}}$ might not be on the boundary of the convex hull of $V$. This is illustrated in Figure 7, (b). In both of these cases, for $q > q^{\mathrm{CD}}$, by Corollary 2 the allocation of type 1 in an optimal randomized menu must result in a pair of expected valuations $(\hat{v}, \hat{v})$ for the two types that is the highest point on the diagonal that is also in the convex hull of $V$. But there is no (deterministic) alternative with that pair of valuations because otherwise type 1 would have a higher valuation for such an alternative than for $a^{\mathrm{BD1}}$, contradicting the definition of $a^{\mathrm{BD1}}$. A sufficient condition for randomization *not* to help is that no alternative lies strictly below the diagonal. This is illustrated in Figure 7, (c). This sufficient condition holds under increasing differences because then the outside option lies on the diagonal and every other alternative lies strictly above it (see Appendix D). Thus, Corollary 3 confirms that with increasing differences, randomization does not help.

To gain some intuition for why randomization helps, consider the sufficient condition that $a^{\mathrm{BD1}}$ lies strictly below the diagonal. As Theorem 1 shows, when the probability of type 2 is high enough, IR2 holds as an equality and the principal optimally offers $a_1 = a^{\mathrm{BD1}}$ and charges both types their valuation for their assigned alternative. If $a^{\mathrm{BD1}}$ lies strictly below the diagonal, then IC2 is slack. Since type 1 is not assigned his efficient alternative, the principal would like to assign type 1 a "slightly more efficient" alternative and charge him more, which would not violate IC2 because it is slack. Randomization makes this possible: instead of assigning type 1 alternative $a^{\mathrm{BD1}}$, the principal can increase his revenue by assigning type 1 a random alternative that makes type 2's valuation for this random alternative equal to that of type 1 (so the random alternative lies on the diagonal).

# 5 Applications

## 5.1 Horizontal and Vertical Differentiation

In this application, each alternative corresponds to a product, and each product is characterized by a horizontal value and a vertical value. Both types like higher vertical values but may differ in their ideal horizontal value.

We formalize this as follows. Each alternative $a \in A$ has a "size" $s(a) \in [0, 1]$ and a "location" $\ell(a) \in [0, 1]$, with $s(0) = \ell(0) = 0$ for the outside option. We assume that all size-location pairs in $[0, 1] \times [0, 1]$ are possible, so $\{(s(a), \ell(a)) \mid a \in A\} = [0, 1] \times [0, 1]$. Each type $t$ is also represented by a size $s(t) \in [0, 1]$ and a location $\ell(t) \in [0, 1]$. A type's size represents the intensity of his preferences, and his location represents his ideal horizontal value, so his valuation for an alternative decreases in the distance between his location and that of the alternative. Specifically, type $t$'s valuation for alternative $a$ is $s(t)s(a)cos(2\pi(\ell(t) - \ell(a)))$. If we let $\vec{u}(t)$ be a vector of size $s(t)$ and angle $2\pi\ell(t)$ and $\vec{u}(a)$ be a vector of size $s(a)$ and angle $2\pi\ell(a)$, as shown in Figure 8, (a), then type $t$'s valuation for alternative $a$ is the inner product of the two vectors, $\vec{u}(t) \cdot \vec{u}(a)$. We assume that the cost of producing an alternative is linear in its size, that is, $c(a) = cs(a)$ for some constant $c \geq 0$.

The special case in which both types have the same location corresponds to standard second-degree price discrimination with quality/quantity values $s(a) \in [0, 1]$ and linear costs. This is because in this case, when considering optimal menus we can restrict attention to alternatives whose location coincides with that of the types.[24]

[24] To see this, suppose without loss of generality that both types' location is zero, that is, $\ell(t) =$

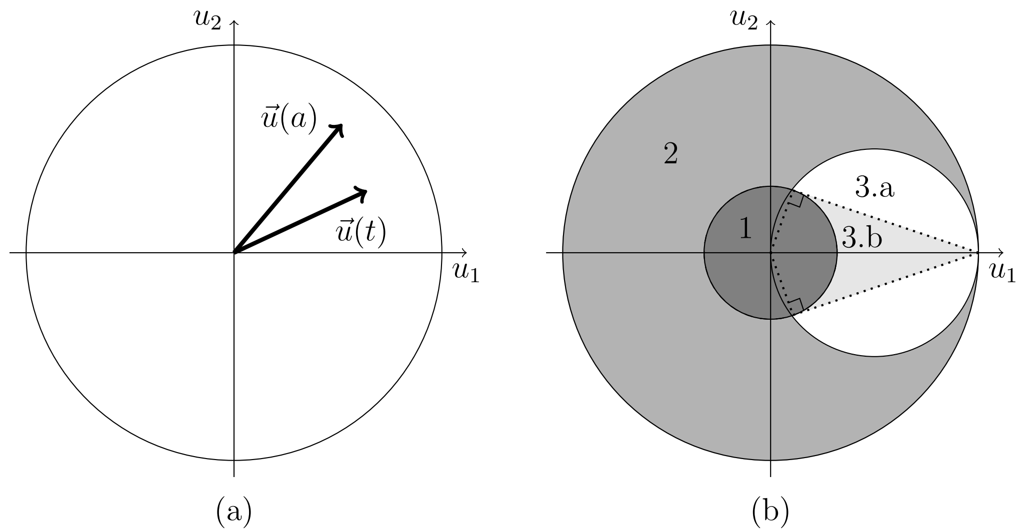



Figure 8: (a) Vector $\vec{u}(t)$ has size $s(t)$ and angle $2\pi\ell(t)$, and vector $\vec{u}(a)$ has size $s(a)$ and angle $2\pi\ell(a)$. (b) The four possible cases for type 1's location in the unit circle.

We now describe the optimal menus for the different possible configurations of the types. A type's efficient alternative is the outside option if the type's size is less than $c$, and has size 1 and the same location as the type's if the type's size exceeds $c$. Thus, if the size of both types is less than $c$, then the first-best is to exclude both types, and this is implementable. Suppose that this is not the case, and assume without loss of generality that type 2's size is weakly higher than type 1's, so type 2 is uncontested. Moreover, normalize type 2's size to 1 and his location to 0.

Figure 8 (b) depicts the various possibilities for type 1's location in the unit circle. The first-best is implementable when type 1 is in the shaded regions 1 and 2, which is when type 1 is uncontested. In region 1, which is the circle with radius $c$ centered at the origin, the size of type 1 is less than $c$, so type 1 is efficiently excluded. When type 1 lies outside of region 1, the efficient alternative of type 1 is the unit vector in his direction. In region 2, which is the area outside of region 1 that is also outside of the unshaded circle with radius $\frac{1}{2}$ centered at point $(\frac{1}{2}, 0)$, type 1 is uncontested so his optimal alternative is the efficient one (the unit vector in his direction). Intuitively, in this case type 1's ideal horizontal value (given his preference intensity) is sufficiently

0. Any alternative $a$ with $\ell(a) \neq 0$ can be replaced with an alternative $a'$ that satisfies $s(a') = s(a)cos(2\pi(\ell(a)))$ and $\ell(a') = 0$. Both types have the same valuation for alternative $a'$ and alternative $a$, but the cost of alternative $a'$ is lower.

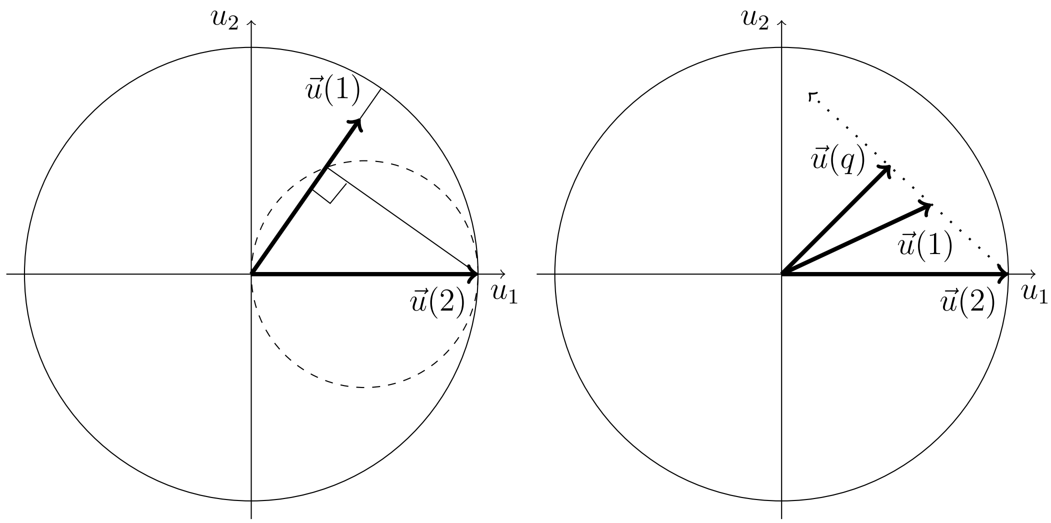



Figure 9: (a) Type 1 has a higher valuation than type 2 for the unit vector in the direction of type 1. (b) Vector $\vec{u}(q)$ is obtained by moving away from $\vec{u}(2)$ and $\vec{u}(1)$.

different from that of type 2, or, equivalently, type 1's preference intensity (given his ideal horizontal value) is sufficiently high that type 1 values his efficient alternative more than type 2. More precisely, type 1's valuation for his efficient alternative is the size of type 1, which is higher than type 2's valuation for type 1's efficient alternative, which is the projection of type 2's vector onto the unit vector in the direction of type 1, as shown in Figure 9, (a).

In regions 3.a and 3.b the first-best is not implementable, so type 1's alternative in any optimal menu is distorted away from efficiency. The type of distortion differs in the two regions and depends on the probabilities of the two types. To describe the distortion, for any probability $q$ of type 2, let

$$\vec{u}(q) = \frac{1}{1-q}\vec{u}(1) - \frac{q}{1-q}\vec{u}(2)$$

be the affine combination of $\vec{u}(1)$ and $\vec{u}(2)$ with weights $\frac{1}{1-q}$ and $\frac{-q}{1-q}$. This $\vec{u}(q)$ is obtained by moving away from $\vec{u}(1)$ on the line that connects $\vec{u}(2)$ and $\vec{u}(1)$, as shown in Figure 9, (b). As $q$ increases, $\vec{u}(q)$ moves further away from $\vec{u}(1)$ and $\vec{u}(2)$.

In region 3.a, for $q < q^{\mathrm{OD}}$, type 1's alternative in an optimal menu is the unit vector in the direction of $\vec{u}(q)$, and type 2 obtains information rent, and for $q > q^{\mathrm{OD}}$ type 1's alternative is the unit vector in the direction of $\vec{u}(q^{\mathrm{OD}})$, and type 2 obtains

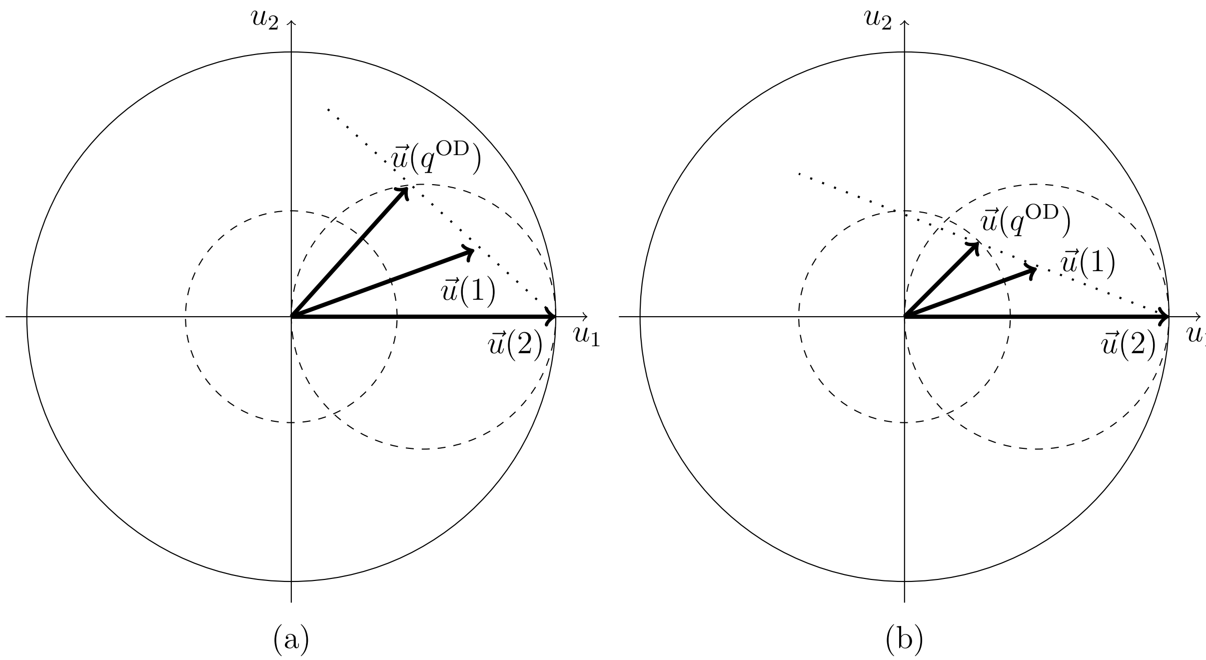



Figure 10: Threshold probability $q^{\mathrm{OD}}$ in regions 3.a and 3.b.

no information rent. Thus, in region 3a, the distortion is in the horizontal dimension only. The threshold probability $q^{\mathrm{OD}}$ is such that vector $\vec{u}(q^{\mathrm{OD}})$ lies on the boundary of the circle with radius $\frac{1}{2}$ centered at point $(\frac{1}{2}, 0)$, as shown in Figure 10, (a). In region 3.b, for $q < q^{\mathrm{OD}}$ type 1's alternative is the unit vector in the direction of $\vec{u}(q)$, and type 2 obtains information rent, similarly to region 3.a, but for $q > q^{\mathrm{OD}}$ type 1's alternative is the outside option and type 2 obtains no information rent. Thus, the distortion is only in the horizontal dimension for low $q$, but leads to inefficient exclusion of type 1 for high $q$. The threshold probability $q^{\mathrm{OD}}$ is such that vector $\vec{u}(q^{\mathrm{OD}})$ lies on the boundary of the circle with radius $c$ centered at the origin, as shown in Figure 10, (b). We formalize this discussion in Appendix G, .

Notice that with no production costs, $c = 0$, type 1 can only be in region 3.b if $\vec{u}(1)$ is on the horizontal axis. Otherwise, type 1 lies above or below the line that connects $\vec{u}(2)$ to the circle with radius zero, which is the origin. So the fact that in the standard price-discrimination setting the low type might be inefficiently excluded (for sufficiently high probability of the high type) when the cost is zero is a non-generic property that only holds when both types' ideal horizontal values coincide.

## 5.2 Bundling with additive valuations

We consider a bundling application with a set $\{1, \ldots, n\}$ of products. Each subset of products corresponds to an alternative $a \subseteq \{1, \ldots, n\}$, and the empty set corresponds to the outside option $a = 0$. We allow for random alternatives, which correspond to the set of all random subsets of products $\Delta(A)$, $A = \{a \mid a \subseteq \{1, \ldots, n\}\}$.[25] Type $t$'s valuation (surplus net of production cost) for product $i$ is $v_t(i)$, which may be negative. We assume that valuations and production costs are additive, so type $t$'s valuation for alternative $a \in A$ is $v_t(a) = \sum_{i \in a} v_t(i)$, with $v_t(0) = 0$. We first discuss a two-product example, and then give the characterization for any number of products.

### 5.2.1 An Example

Suppose that there are two products, 1 and 2, which means that there are four (deterministic) alternatives, $\emptyset$, $\{1\}$, $\{2\}$, and $\{1, 2\}$. Type 1's valuation for each product is positive, so his efficient alternative is the grand bundle $a_1^{\mathrm{E}} = \{1, 2\}$. Type 2's valuation is negative for product 1 and positive for product 2, so his efficient alternative is $a_2^{\mathrm{E}} = \{2\}$. Even though type 2's valuation for product 1 is negative, his valuation for the grand bundle exceeds that of type 1, that is,

$$v_2(\{2\}) > v_2(\{1, 2\}) = v_2(1) + v_2(2) > v_1(1) + v_1(2) = v_1(\{1, 2\}) > v_1(\{2\}),$$

so type 2 is uncontested and type 1 is contested.[26] Figure 11, (a) qualitatively depicts both types' valuations for each of the two products, where type 1's valuations are on the horizontal axis and type 2's valuations are on the vertical axis.

Figure 11, (b) shows both types' valuations for all four *alternatives*. Alternatives $\{1\}$ and $\{2\}$ in panel (b) are in the same location as products 1 and 2 in panel (a). Panel (b) also depicts alternative $\{1, 2\}$ and the outside option, which corresponds to the empty set and is located at the origin. To distinguish the two figures, we say that panel (a) is in the "product space" and panel (b) is in the "alternative space."

Because randomization is allowed, we use Corollary 2 to identify the optimal menus. In any optimal menu, type 2 obtains his efficient alternative $a_2^{\mathrm{E}} = \{2\}$, and

[25] We illustrate the complexities that arise when random alternatives are prohibited in Appendix J.

[26] Specific values that satisfy all the assumptions in this section are $v_1(1) = 1, v_1(2) = 3, v_2(1) = -1$, and $v_2(2) = 6$.

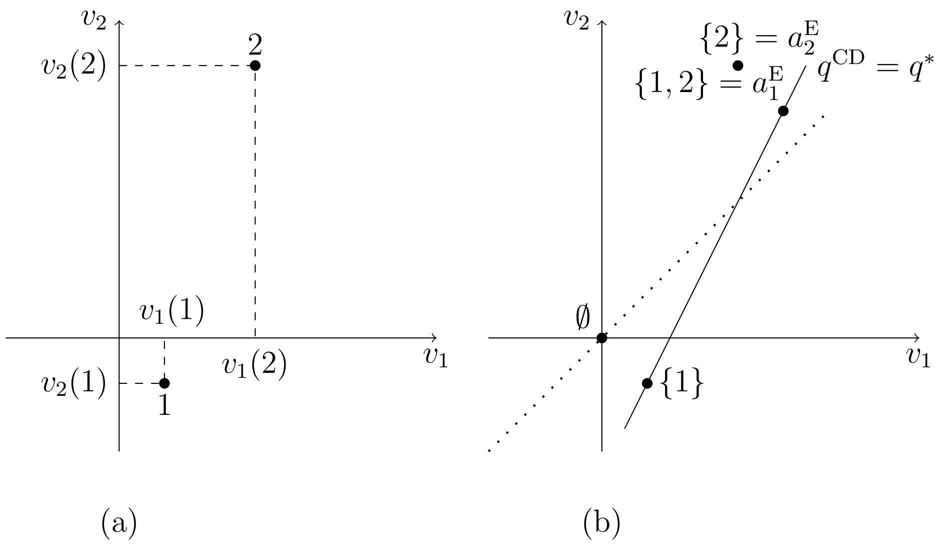



Figure 11: (a) In the product space, each point represents a product. (b) In the alternative space, each point represents an alternative.

IC2 and IR1 hold as equalities. An optimal menu is pinned down by type 1's (possibly random) alternative. To characterize this alternative for each probability $q$ of type 2, we solve for $B_1(q)$, which is the set of alternatives (bundles) that maximize

$$v_1(a) - qv_2(a) = \sum_{i\in a} \Big(v_1(i) - qv_2(i)\Big).$$

Because the expression is additively separable across products, a bundle that maximizes the above expression contains a product $i$ if $v_1(i) - qv_2(i) > 0$, does not contain the product if $v_1(i) - qv_2(i) < 0$, and may or may not contain the product if $v_1(i) - qv_2(i) = 0$ (in which case $B_1(q)$ contains multiple bundles). To visualize this, we can draw a hyperplane in the *product space* that goes through the set of points $(b, c)$ such that $b - qc = 0$, as shown in Figure 12, (a). A product is in a maximizing bundle if it is to the right of the hyperplane, is not in the bundle if it is to the left of the hyperplane, and may or may not be in the bundle if it is on the hyperplane. In Figure 12, (a), the unique alternative in $B_1(q)$ is bundle $\{1\}$.

As $q$ increases, this hyperplane rotates clockwise. We can therefore describe $B_1(q)$ using a threshold $q^*$ at which the hyperplane goes through the point that represents product 2, shown Figure 12, (b). If $q < q^*$, then $\{1, 2\}$ is the unique alternative in

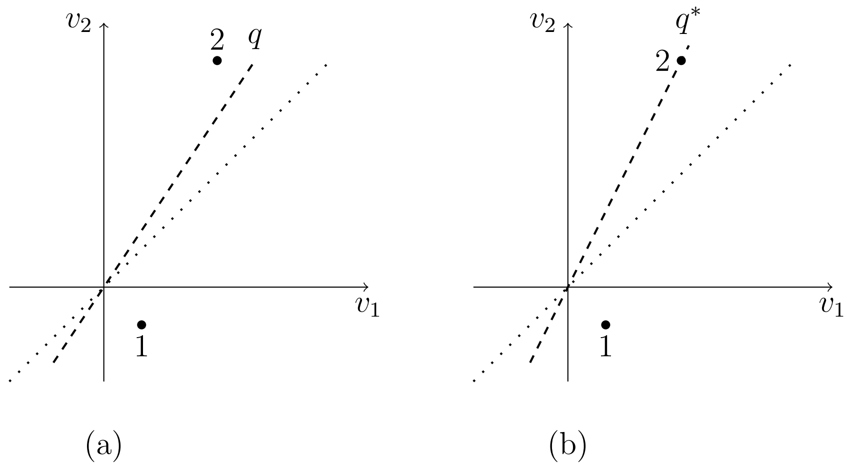



Figure 12: (a) Representation of the example in product space. (b) The hyperplane orthogonal to vector $(1, -q^*)$ goes through product 2.

$B_1(q)$. If $q = q^*$, then both $\{1\}$ and $\{1,2\}$ are in $B_1(q)$. If $q > q^*$, then $\{1\}$ is the unique alternative in $B_1(q)$.

This threshold $q^*$ is in fact threshold $q^{\text{CD}}$ from Corollary 2. This is because in the *alternative space* shown in Figure 11, (b), $\{1,2\}$ lies strictly above the diagonal and $\{1\}$ lies strictly below the diagonal, so $q^*$ is the lowest $q$ such that $B_1(q)$ contains an alternative that lies weakly above the diagonal and an alternative that lies weakly below the diagonal. The locations of $\{1,2\}$ and $\{1\}$ relative to the diagonal correspond to the following inequalities:

$$\sum_{i \in \{1,2\}} \Big( v_2(i) - v_1(i) \Big) > 0 \qquad \text{and} \qquad \sum_{i \in \{1\}} \Big( v_2(i) - v_1(i) \Big) < 0.$$

Thus, there exists some $\tilde{p} \in (0,1)$ such that

$$\tilde{p} \sum_{i \in \{1,2\}} \Big( v_2(i) - v_1(i) \Big) + (1 - \tilde{p}) \sum_{i \in \{1\}} \Big( v_2(i) - v_1(i) \Big) = 0.$$

We can now use Corollary 2 to describe type 1's (possibly random) bundle in any optimal menu. If $q < q^* = q^{\text{CD}}$, then type 1 obtains the grand bundle $\{1,2\}$. If $q = q^* = q^{\text{CD}}$, type 1 obtains (any) random bundle with support over $\{1,2\}$ and $\{1\}$ in which the probability of $\{1,2\}$ is at least $\tilde{p}$. If $q > q^*$, then type 1 obtains the random bundle that assigns probability $\tilde{p}$ to $\{1,2\}$ and probability $1 - \tilde{p}$ to $\{1\}$.

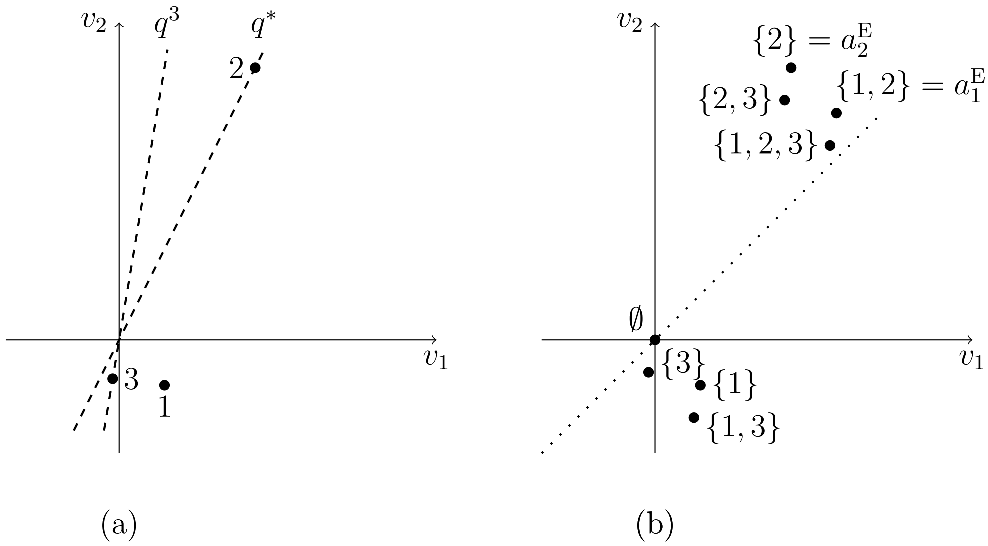



Figure 13: (a) Representation of the example in product space. (b) The hyperplane orthogonal to vector $(1, -q^*)$ goes through product 2.

One feature of the example is that for each product the valuation of at least one type is positive. But, perhaps surprisingly, even products for which *both* types' valuations are negative may be offered as part of an optimal menu. To see this, consider adding product 3, for which both types' valuations, $v_1(\{3\})$ and $v_2(\{3\})$, are negative. Suppose that $v_1(\{3\}) = q^3 v_2(\{3\})$ for some positive $q^3 < q^*$, as shown in Figure 13, (a). Suppose also that in the alternative space, alternative $\{1,2,3\}$ lies above the diagonal, as shown in Figure 13, (b).[27] Then, for any $q \in (q^3, q^*)$, in any optimal menu type 2 obtains his efficient alternative $a_2^{\mathrm{E}} = \{2\}$ and type 1 obtains the grand bundle $\{1,2,3\}$. Intuitively, even though allocating product 3 to type 1 reduces the surplus that can be extracted from type 1, the resulting reduction in type 2's information rent is large enough that, overall, the seller's revenue increases.[28]

[27]Specific values that satisfy all the assumptions are $v_1(1) = 1, v_1(2) = 3, v_1(3) = -1/6, v_2(1) = -1$, $v_2(2) = 6$, and $v_2(3) = -1$.

[28]Recall that valuations in our setting are normalized and represent willingness-to-pay net of costs. Appendix H interprets what "giving a product with a negative valuation" means in the pre-normalized setting.

### 5.2.2 The Characterization with Any Number of Products

Our characterization for any number of products, which we describe informally here and formally in Appendix I, generalizes the above example. Type $t$'s efficient alternative $a_t^{\mathrm{E}} = \{i \mid v_t(i) > 0\}$ is the set of products for which type $t$'s valuation is positive. Suppose without loss of generality that type 2 is uncontested, so $v_2(a_2^{\mathrm{E}}) \geq v_1(a_2^{\mathrm{E}})$, and focus on the case in which type 1 is contested, so the first-best is not implementable, $v_2(a_1^{\mathrm{E}}) > v_1(a_1^{\mathrm{E}})$. In any optimal menu, type 2's allocation is $a_2^{\mathrm{E}}$ and IC2 and IR1 hold as equalities.

The optimal allocation of type 1 is described by the threshold $q^{\mathrm{CD}}$ (which we will shortly identify). As shown in Figure 14, if $q < q^{\mathrm{CD}}$, then type 1's allocation contains every product that is to the right of the hyperplane that goes through the set of points $(v_1, v_2)$ such that $v_1 - qv_2 = 0$, does not contain any product to the left of the hyperplane, and may or may not contain any product that is on the hyperplane. If $q \geq q^{\mathrm{CD}}$, then what matters is the position of a product in comparison to the hyperplane corresponding to $q^{\mathrm{CD}}$, and not $q$: type 1's allocation contains every product that is to the right of the hyperplane that goes through the set of points $(v_1, v_2)$ such that $v_1 - q^{\mathrm{CD}}v_2 = 0$, does not contain any product to the left of the hyperplane, and randomizes among the products on the hyperplane in a way that both types have the same valuation for the bundle. Indeed, the threshold $q^{\mathrm{CD}}$ is exactly where such randomization is possible: $q^{\mathrm{CD}}$ is such that if we include the products on the hyperplane and in the first quadrant, then $v_2 - v_1$ is positive, and if we include the products on the hyperplane and in the third quadrant, then $v_2 - v_1$ is negative. So by randomizing over these products appropriately, we can construct a random bundle for which $v_2 = v_1$.

This characterization clarifies which products may optimally be assigned to type 1. The products in the fourth quadrant and those in the first quadrant and below the hyperplane corresponding to $q^{\mathrm{CD}}$ are always assigned to type 1. Those in the second quadrant and those in the third quadrant and above the hyperplane corresponding to $q^{\mathrm{CD}}$ are never assigned to type 1. Other products might be assigned to type 1 for some values of $q$ and not other values of $q$. In particular, for each product $i$ in the first quadrant that lies above the $q^{\mathrm{CD}}$-hyperplane, define $r_i = v_1(i)/v_2(i)$. Product $i$ is assigned to type 1 whenever $q < r_i$, and is not assigned whenever $q > r_i$. Similarly,

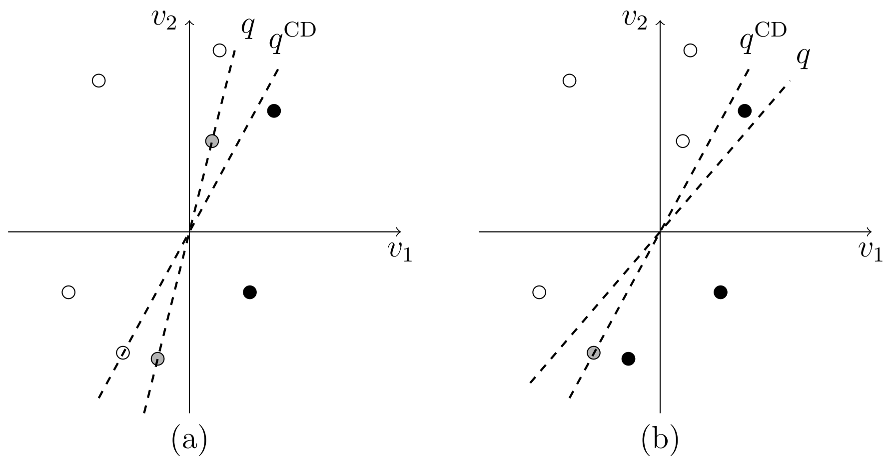



Figure 14: The optimal bundle of type 1 contains all products filled black and may or may not contain products filled gray.

for each product $i$ in the third quadrant that lies below the $q^{\mathrm{CD}}$-hyperplane, define $r_i = v_1(i)/v_2(i)$. Product $i$ is not assigned to type 1 whenever $q < r_i$, and is assigned whenever $q > r_i$.

As $q$ increases, as long as it stays below $q^{\mathrm{CD}}$, some products are added to type 1's bundle and some products are removed from it. The ones that are removed lie in the first quadrant and are above the 45 degree line through the origin. These are products for which both types have a positive valuation, but type 2 values them more, so while removing them reduces the surplus extracted from type 1, it also reduces the information rent of type 2. The products that are added are those in the third quadrant and are below the 45 degree line through the origin. These are products for which both types have a negative valuation, but type 2 values them less, so while adding them reduces the surplus extracted from type 1, it also reduces the information rent of type 2.

# 6 Conclusions

We characterize optimal menus in screening settings with an agent who has quasi-linear preferences and one of two possible types. We introduce the notion of an uncontested type, show that at least one type is uncontested, and prove that the principal faces no tradeoff between maximizing the allocation value for a contested type and minimizing the information rent of the other type. This implies that that the first-best is implementable if and only if both types are uncontested. When only one type is uncontested, an optimal menu is characterized by the allocation of the contested type, which we identify. The characterization enhances our understanding of screening with two types and the role that assumptions like increasing differences play. Some properties of optimal menus under increasing differences, such as the existence of a type that is served efficiently and a type that obtains no information rent, independently of the types' probabilities, generalize and do not require increasing differences (because at least one type is uncontested). Other properties, such as one type being indifferent between his alternative and the other type's alternative (at their respective prices) and exclusion, or "near exclusion," of one type when the probability of the other type is sufficiently high, do not always hold. Appendix D further clarifies the role of increasing differences by ordering the alternatives according to the 45 degree lines on which they lie in the space of valuation pairs. Increasing differences implies that the outside option is the unique lowest alternative in this order and no two alternatives lie on the same 45 degree line. These two properties do not hold in general, which drives the differences between our general characterization of the optimal menus and the one in the standard setting with increasing differences.

Our characterization also helps clarify the role that convexity of the set of valuations plays in the characterization of the optimal menus. This allows us to identify precisely when the ability to offer random menus strictly increases the principal's profit, and explains why randomization does not help the principal in the standard setting with increasing differences.

The characterization makes it possible to study canonical problems for which our current understanding is limited, such as multi-product bundling, as well as new applications, such as combining vertical and horizontal product differentiation. Finally, our analysis relies heavily on the assumption of two types, and an intriguing direction for future research is identifying special cases or certain properties for which some of the analysis applies when there are more than two types.

# Appendix

## A Completing the proof of Theorem 1

That the optimal menu implements the first best if and only if both types are uncontested was shown in Section 3.2. In addition, Lemma 1 proves properties 1 and 2 from Theorem 1. It remains to formalize the proof of properties 3-9 from Theorem 1, which complete the characterization of the optimal menus. As mentioned in Section 3.2, identifying type 1's allocation and payment necessitates a different approach from the one in Section 2.1 because we do not assume increasing differences. Our approach includes four steps, based on the observation that in any optimal menu either IR2 or IC2 (or bod) holds an an equality.[29] In the first step we suppose that IR2 holds as an equality, show that this implies that IC1 holds, and solve the resulting relaxed problem.[30] In the second step, we suppose that IC2 holds as an equality, ignore IC1, and solve the resulting relaxed problem. In the third step, we show that the solution satisfies IC1. In the fourth step, we combine the two cases (IR2 holds as an equality and IC2 holds as an equality) to obtain the characterization and prove the remaining properties from Theorem 1.

Step one: Suppose that IR2 holds as an equality. That $a_2 = a_2^{\mathrm{E}}$ and IR1 and IR2 hold as equalities implies that $p_1 = v_1(a_1)$ and $p_2 = v_2(a_2^{\mathrm{E}})$. Substituting these into the principal's problem and constraints, which allows us to remove IR1 and IR2, the problem becomes:

$$\max_{a_1 \in A} \quad (1-q)v_1(a_1) + qv_2(a_2^{\mathrm{E}})$$
$$\text{subject to} \quad v_1(a_1) - v_1(a_1) \geq v_1(a_2^{\mathrm{E}}) - v_2(a_2^{\mathrm{E}}).$$
$$v_2(a_2^{\mathrm{E}}) - v_2(a_2^{\mathrm{E}}) \geq v_2(a_1) - v_1(a_1).$$

Rearranging, and ignoring $qv_2(a_2^{\mathrm{E}})$ in the target function, because it does not

[29] Otherwise type 2's payment can be increased without violating any of the constraints.

[30] Recall that under increasing differences IR2 can be ignored. This is not the case in general.

depend on $a_1$, we obtain:

$$\begin{aligned} \max_{a_1 \in A} \quad & (1-q)v_1(a_1) \\ \text{subject to} \quad & v_1(a_2^{\mathrm{E}}) - v_2(a_2^{\mathrm{E}}) \leq 0 \\ & v_2(a_1) - v_1(a_1) \leq 0. \end{aligned}$$

Because type 2 is uncontested, the first constraint (IC1) holds and can be ignored. We rewrite the second constraint (IC2) as $a_1 \in A^-$ (the set of alternatives that lie weakly below the diagonal) and obtain:

$$\max_{a_1 \in A^-} \quad (1-q)v_1(a_1). \tag{1}$$

In the solution to this problem, the principal optimally assigns to type 1 the alternative $a^{\mathrm{BD1}}$, which maximizes type 1's utility among the alternatives in $A^-$. The resulting menu is $\{(a^{\mathrm{BD1}}, v_1(a^{\mathrm{BD1}})), (a_2^{\mathrm{E}}, v_2(a_2^{\mathrm{E}}))\}$, which generates a profit of

$$(1-q)v_1(a^{\mathrm{BD1}}) + qv_2(a_2^{\mathrm{E}}). \tag{2}$$

<u>Step two:</u> Suppose that IC2 holds as an equality. That $a_2 = a_2^{\mathrm{E}}$ and IR1 and IC2 hold as equalities implies that $p_1 = v_1(a_1)$ and $p_2 = v_2(a_2^{\mathrm{E}}) - v_2(a_1) + v_1(a_1)$. Substituting these into the principal's problem and constraints, which allows us to remove IC2 and IR1, the problem becomes:

$$\begin{aligned} \max_{a_1 \in A} \quad & (1-q)v_1(a_1) + q(v_2(a_2^{\mathrm{E}}) - v_2(a_1) + v_1(a_1)) \\ \text{subject to} \quad & v_2(a_2^{\mathrm{E}}) - (v_2(a_2^{\mathrm{E}}) - v_2(a_1) + v_1(a_1)) \geq 0, \\ & v_1(a_2^{\mathrm{E}}) - (v_2(a_2^{\mathrm{E}}) - v_2(a_1) + v_1(a_1)) \leq 0. \end{aligned}$$

Rearranging and ignoring $qv_2(a_2^{\mathrm{E}})$ in the target function, because it does not depend on $a_1$, we obtain:

$$\begin{aligned} \max_{a_1 \in A} \quad & v_1(a_1) - q v_2(a_1) \\ \text{subject to} \quad & v_2(a_1) - v_1(a_1) \geq 0, \\ & v_2(a_1) - v_1(a_1) \leq v_2(a_2^{\mathrm{E}}) - v_1(a_2^{\mathrm{E}}). \end{aligned}$$

We rewrite the first constraint (IR2) as $a_1 \in A^+$ (the set of alternatives that lie weakly above the diagonal), ignore the second constraint (IC1), and obtain:

$$\max_{a_1 \in A^+} \quad v_1(a_1) - q v_2(a_1). \tag{3}$$

This problem has the following geometric interpretation. Consider the set of valuation pairs $V^+ = \{(v_1(a), v_1(a)) : a \in A^+\}$ that corresponds to the set of alternatives $A^+$. For any probability $q$ of type 2, the principal finds a pair $(v_1^q, v_2^q)$ in $V^+$ that is maximal in the direction of the vector $(1, -q)$. Any corresponding alternative $a_1 \in A^+$ generates a profit of

$$v_1(a_1) - q v_2(a_1) + q v_2(a_2^{\mathrm{E}}). \tag{4}$$

<u>Step three:</u> We now show that IC1 holds for any solution $(v_1^q, v_2^q)$ to the principal's problem in step 2. We showed in step 2 that IC1 is

$$v_2(a_1) - v_1(a_1) \leq v_2(a_2^{\mathrm{E}}) - v_1(a_2^{\mathrm{E}}),$$

so to show that IC1 holds we have to show that

$$v_2^q - v_1^q \leq v_2(a_2^{\mathrm{E}}) - v_1(a_2^{\mathrm{E}}).$$

For this we first prove the following lemma.

**Lemma 2** *For any $q' \geq q \geq 0$ and $S \subseteq \mathbb{R}^2$, if $(a^{q'}, b^{q'}) \in \underset{(a,b) \in S}{max} (a - q'b)$ and $(a^q, b^q) \in \underset{(a,b) \in S}{max} (a - qb)$, then $b^{q'} - a^{q'} \leq b^q - a^q$.*

**Proof.** Suppose that $q' > q$ but $b^q - a^q < b^{q'} - a^{q'}$, that is, $b^q - b^{q'} < a^q - a^{q'}$. By definition of $(a^{q'}, b^{q'})$, we have that $a^{q'} - q' b^{q'} \geq a^q - q' b^q$, that is, $q'(b^q - b^{q'}) \geq a^q - a^{q'}$.

Thus,

$$q'(b^q - b^{q'}) \geq a^q - a^{q'} > b^q - b^{q'},$$

so $b^q - b^{q'} < 0$. Consequently, because $q' > q$, we have that

$$q(b^q - b^{q'}) > a^q - a^{q'}.$$

But by definition of $(a^q, b^q)$, we have that $a^q - qb^q \geq a^{q'} - qb^{q'}$, that is,

$$q(b^q - b^{q'}) \leq a^q - a^{q'},$$

a contradiction. ■

By Lemma 2 applied so $S = V^+$, it is enough to show that

$$v_2^0 - v_1^0 \leq v_2(a_2^{\mathrm{E}}) - v_1(a_2^{\mathrm{E}}),$$

or equivalently

$$v_1(a_2^{\mathrm{E}}) - v_1^0 \leq v_2(a_2^{\mathrm{E}}) - v_2^0.$$

The right-hand side of the inequality is non-negative by definition of $a_2^{\mathrm{E}}$. The left-hand side is non-positive because $v_1^0$ is, by definition, player 1's maximal valuation across the alternatives in $A^+$, and $a_2^{\mathrm{E}} \in A^+$ because type 2 is uncontested. Thus, the inequality holds.

Step four: We now combine the cases in which IR2 holds as an equality (step one) and IC2 holds as an equality (step two). Because alternative $a^{\mathrm{BD1}}$ maximizes type 1's valuation across the alternatives in $A^-$, and $v_1(a^{\mathrm{VA}}) = v_2(a^{\mathrm{VA}}) = v_1(a^{\mathrm{BD1}})$ by definition of $a^{\mathrm{VA}}$, the value $(1-q)v_1(a^{\mathrm{BD1}})$ of the solution to (1) can be written as $v_1(a^{\mathrm{VA}}) - qv_2(a^{\mathrm{VA}})$. Thus, the following maximization problem combines the cases:

$$\max_{a_1 \in A^+ \cup \{a^{\mathrm{VA}}\}} \quad v_1(a_1) - qv_2(a_1). \tag{5}$$

Denoting by $A_1'(q)$ the set of maximizers that solve (5), we have that for any maximizing alternative $a_1 \in A_1'(q)$ the corresponding profit of the principal is

$$v_1(a_1) - qv_2(a_1) + qv_2(a_2^{\mathrm{E}}).$$

This is because if $a_1 \in A^+$, then in the optimum IC2 holds as an equality and the profit is given by (4), and if $a_1 = a^{\mathrm{VA}}$, then in the optimum IR2 holds as an equality and the profit is

$$(1-q)v_1(a^{\mathrm{BD1}}) + qv_2(a_2^{\mathrm{E}}) = v_1(a^{\mathrm{VA}}) - qv_2(a^{\mathrm{VA}}) + qv_2(a_2^{\mathrm{E}}).$$

This proves property 6 and the first part of property 3 from Theorem 1. Property 3(a) then holds by definition of $a^{\mathrm{BD}}$.

If $a_1 \in A^+$, then $a_1$ solves the principal's problem, and if $a_1 = a^{\mathrm{VA}}$, then $a^{\mathrm{BD1}}$ solves the principal's problem. This proves properties 4 and 5 from the statement of Theorem 1. We now prove the remaining parts of Theorem 1.

Let $\hat{V}^+ = \{(a,b) : a \leq b\} \cup \{(v_1(a^{\mathrm{VA}}), v_2(a^{\mathrm{VA}}))\}$ be the set of valuation pairs that corresponds to the set of alternatives $A^+ \cup \{a^{\mathrm{VA}}\}$. For any $q \in (0,1)$, with a slight abuse of notation denote by $(v_1^q, v_2^q)$ a pair in $\hat{V}^+$ that corresponds to an alternative $a_1 \in A^+ \cup \{a^{\mathrm{VA}}\}$ that solves the combined problem (5). Thus, $(v_1^q, v_2^q)$ maximizes $a - qb$ across all valuation pairs $(a,b)$ in $\hat{V}^+$. By Lemma 2 with $S = \hat{V}^+$, we have that $v_2^q - v_1^q$ weakly decreases in $q$. Now, if $a_1 \in A^+ \cup \{a^{\mathrm{VA}}\}$ solves (5), then in the corresponding menu type 2's utility is $v_2(a_1) - v_1(a_1)$, so type 2's utility weakly decreases in $q$ (property 8 from Theorem 1).[31]

That $v_2^q - v_1^q$ weakly decreases in $q$ also shows that if IR2 holds as an equality, so $a_1 = a^{\mathrm{BD1}}$ is optimal for some $q$, then $a^{\mathrm{BD1}}$ is uniquely optimal for any $q' > q$ (which implies property 3(b) from Theorem 1). To see this, we first observe that since the outside option corresponds to the valuation pair $(0,0) \in \hat{V}^+$, we must have that $v_1^q - qv_2^q \geq 0$. Together with $(v_1^q, v_2^q) \in \hat{V}^+$, we have that $v_2^q \geq v_1^q \geq qv_2^q$. This implies that $v_2^q \geq 0$, and therefore $v_1^q \geq 0$ as well. Now suppose that $q' > q$ and $a_1 = a^{\mathrm{BD1}}$ is optimal for $q$. By definition of $a^{\mathrm{BD1}}$ and $a^{\mathrm{VA}}$, $(v_1(a^{\mathrm{VA}}), v_2(a^{\mathrm{VA}}))$ is the highest pair in $\hat{V}^+$ on the diagonal. And since (i) $a^{\mathrm{BD1}}$ is optimal for $q$ if and only if $(v_1^q, v_2^q) = (v_1(a^{\mathrm{VA}}), v_2(a^{\mathrm{VA}}))$, (ii) $(v_1^{q'}, v_2^{q'})$ lies on a weakly lower 45 degree line than $(v_1^q, v_2^q)$ does (by Lemma 2 with $S = \hat{V}^+$), and (iii) the diagonal, on which the valuation pair $(v_1(a^{\mathrm{VA}}), v_2(a^{\mathrm{VA}})$ lies, is the lowest 45 degree line in the set $\hat{V}^+$, we have that $(v_1^{q'}, v_2^{q'})$ also lies on this line. And because $q' < 1$ implies that $a - q'a < b - q'b$ for

[31]If $a_1 \in A^+$ then IC2 holds as an equality, so $p_2 = v_2(a_2^{\mathrm{E}}) - v_2(a_1) + v_1(a_1)$ and type 2's utility is $v_2(a_1) - v_1(a_1)$, and if $a_1 = a^{\mathrm{VA}}$, then IR2 holds as an equality and type 2's utility is $v_2(a^{\mathrm{VA}}) - v_1(a^{\mathrm{VA}}) = 0$.

any $b > a \geq 0$, we must have $(v_1^{q'}, v_2^{q'}) = (v_1(a^{\mathrm{VA}}), v_2(a^{\mathrm{VA}}))$. Thus, $a^{\mathrm{BD1}}$ is uniquely optimal for $q'$.

We now show that both $v_1(a_1)$ and $v_2(a_1)$ weakly decrease in $q$, where $a_1$ is type 1's alternative in an optimal menu for $q$ (property 7 from Theorem 1). For this we first show that both $v_1^q$ and $v_2^q$ weakly decrease in $q$. Indeed, suppose that $q' > q$ but $v_2^{q'} > v_2^q$. By definition, we have that $v_1^q - qv_2^q \geq v_1^{q'} - qv_2^{q'}$, and since $v_2^{q'} > v_2^q \geq 0$ and $q' > q$, we may conclude that

$$v_1^q - q'v_2^q > v_1^{q'} - q'v_2^{q'},$$

a contradiction. Thus, $v_2^q$ weakly decreases in $q$. Now suppose that $q' > q$ but $v_1^{q'} > v_1^q$. Since $v_2^q \geq v_2^{q'}$, we have that

$$v_1^{q'} - qv_2^{q'} > v_1^q - qv_2^q,$$

a contradiction to the definition of $(v_1^q, v_2^q)$. That both $v_1^q$ and $v_2^q$ weakly decrease in $q$ implies that both $v_1(a_1)$ and $v_2(a_1)$ weakly decrease in $q$ because (i) if $a_1 \in A^+$ then $(v_1^q, v_2^q) = (v_1(a_1), v_2(a_1))$, (ii) if $a_1 = a^{\mathrm{BD1}}$ then $v_1^q = v_1(a^{\mathrm{BD1}})$ and $v_2^q \geq v_2(a^{\mathrm{BD1}})$, and (iii) if $a_1 = a^{\mathrm{BD1}}$ is optimal for some $q$, then $a_1 = a^{\mathrm{BD1}}$ is optimal for any $q' > q$.

It remains to show property 9 from the statement of Theorem 1. Since $(0, 0) \in \hat{V}^+$, by definition of $a^{\mathrm{BD1}}$ we have that $v_1(a^{\mathrm{BD1}}) \geq 0$. If $v_1(a^{\mathrm{BD1}}) > 0$, then, since $v_1(a^{\mathrm{VA}}) = v_1(a^{\mathrm{BD1}})$ and $v_1^q$ decreases in $q$, we have that for any $q$, any optimal $a_1$ satisfies $v_1(a_1) \geq v_1(a^{\mathrm{BD1}}) > 0$, so type 1 is not excluded or "almost excluded." If $v_1(a^{\mathrm{BD1}}) = 0$, then $(0, 0)$ is the pair with the highest value in $\hat{V}^+$ on the diagonal. For $q = 1$, $(0, 0)$ maximizes (5), so for $q$ close to 1, by continuity, the maximizer of (5) must coincide or be close to $(0, 0)$. This completes the proof of Theorem 1.

# References


Adams, W. J. and Yellen, J. L. (1976). Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, pages 475–498.

Armstrong, M. and Vickers, J. (2001). Competitive price discrimination. *rand Journal of economics*, pages 579–605.

Bergemann, D., Bonatti, A., and Smolin, A. (2018). The design and price of information. *American Economic Review*, 108(1):1–48.

Bergemann, D., Bonatti, A., and Smolin, A. (2025). The economics of large language models: Token allocation, fine-tuning, and optimal pricing. *arXiv preprint arXiv:2502.07736*.

Boleslavsky, R. and Kolb, A. (2025). The design and price of influence. *working paper*.

Bolton, P. and Dewatripont, M. (2004). *Contract theory*. MIT press.

Börgers, T. (2015). *An introduction to the theory of mechanism design*. Oxford University Press, USA.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2014). The complexity of optimal mechanism design. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 1302–1318. SIAM.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2017). Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767.

Galperti, S. (2015). Commitment, flexibility, and optimal screening of time inconsistency. *Econometrica*, 83(4):1425–1465.

Ghili, S. (2023). A characterization for optimal bundling of products with nonadditive values. *American Economic Review: Insights*, 5(3):311–326.

Haghpanah, N. and Hartline, J. (2021). When is pure bundling optimal? *The Review of Economic Studies*, 88(3):1127–1156.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, 39(153):41–57.

Maskin, E. and Riley, J. (1984). Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196.

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

Myerson, R. B. (1979). Incentive compatibility and the bargaining proglem. *Econometrica*, 47(1):61–73.

Pycia, M. (2006). Stochastic vs deterministic mechanisms in multidimensional screening. *working paper*.

Riley, J. and Zeckhauser, R. (1983). Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289.

Rochet, J.-C. and Stole, L. A. (2002). Nonlinear pricing with random participation. *The Review of Economic Studies*, 69(1):277–311.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, pages 141–156.

Stigler, G. J. (1963). United States v. Loew's inc.: A note on block-booking. *The Supreme Court Review*, 1963:152–157.

Tadelis, S. and Segal, I. (2005). Lectures in contract theory.

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

Thirumulanathan, D., Sundaresan, R., and Narahari, Y. (2019). Optimal mechanisms for selling two items to a single buyer having uniformly distributed valuations. *Journal of Mathematical Economics*, 82:1–30.

Villas-Boas, J. M. (1999). Dynamic competition with customer recognition. *The Rand Journal of Economics*, pages 604–631.

Vincent, D. and Manelli, A. (2007). Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory*, 137(1):153–185.

Yang, F. (2024). Nested bundling. *working paper*.

# Supplemental Appendix

## B The Example in Section 1.1

The efficient alternative is steak for Applied theorists. It is steak for Theorists if $v \leq 30$ and is fish if $v \geq 30$, as shown in Figure 15. Applied theorists are uncontested, so in any optimal menu their allocation is steak, and Theorists obtain no information rent. Theorists are uncontested if and only if $v \geq 30$. In this case the optimal menu implements the first-best, so steak priced at 60 and fish priced at $v$ are offered. For the rest of the analysis suppose that $v < 30$.

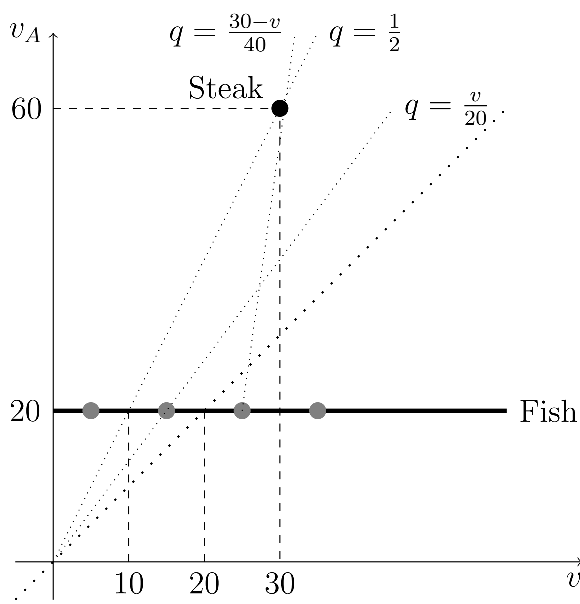



Figure 15: T stands for Theorist and A for Applied theorist. The circle shaded dark shows valuations for steak, and the circles shaded gray show some possible valuations for fish.

We describe the optimal menus when the restaurant cannot offer mixed dishes (randomization is not allowed) and when the restaurant can offer mixed dishes. Corollary 3 shows that randomization makes a difference if and only if $20 < v < 30$ and $q$ is sufficiently high. This is because in this case, alternative $a^{\mathrm{BD1}}$, the best alternative for Theorists among those they like weakly more than Applied theorists, is fish, which lies strictly below the diagonal. When $v \leq 20$, on the other hand, $a^{\mathrm{BD1}}$ is the outside option, which lies on the diagonal and on the convex hull of $V$. Thus, in our analysis of the optimal menus below, we need to consider whether randomization is allowed only for $20 < v < 30$.

First suppose that $v \leq 10$. Then $q^{\text{BD1}} = \frac{1}{2}$ is the threshold $q$ at which both steak and the outside option are maximal in the direction of vector $(1, -q)$. By Theorem 1, the optimal allocation for Theorists is steak if $q \leq q^{\text{BD1}} = \frac{1}{2}$, so only steak priced at 30 is offered, and is the outside option (exclusion) if $q \geq q^{\text{BD1}} = \frac{1}{2}$, so only steak priced at 60 is offered.

Now suppose that $10 \leq v \leq 20$. Then $q^{\text{BD1}} = \frac{v}{20}$ is the threshold $q$ at which both fish and the outside option are maximal in the direction of vector $(1, -q)$. By Theorem 1, the optimal allocation for Theorists is steak if $q \leq \frac{30-v}{40}$ (which is the threshold at which both steak and fish are maximal in the direction of vector $(1, -q)$), so only steak priced at 30 is offered, is fish if $\frac{30-v}{40} \leq q \leq \frac{v}{20}$, so fish priced at $v$ and steak priced at $v+40$ are offered, and is exclusion if $q \geq \frac{v}{20}$, so only steak priced at 60 is offered.

Now suppose that $20 < v < 30$ and randomization is allowed. Then $q^{\text{CD}} = q^{\text{BD1}} = \frac{30-v}{40}$ is the threshold $q$ at which both fish and steak, one below and one above the diagonal, are maximal in the direction of vector $(1, -q)$. By Corollary 2, the optimal allocation for Theorists is steak if $q \leq \frac{30-v}{40}$, so only steak priced at 30 is offered, and is "surf-and-turf" if $q \geq \frac{30-v}{40}$. The surf-and-turf lies on the diagonal, so it consists of a fraction $\frac{30}{v+10}$ of the fish dish and the remaining fraction $\frac{v-20}{v+10}$ of the steak dish. This mixed dish is priced at $\frac{60(v-10)}{v+10}$ and steak is priced 60, which extracts all surplus from both consumer types without excluding either type.

Now suppose that $20 < v < 30$ and randomization is *not* allowed. By Theorem 1, we consider the virtual alternative $a^{\text{VA}}$, which in our example is an alternative for which both types have valuation $v$, as shown in Figure 16. Then $q^{\text{BD1}} = \frac{30-v}{60-v}$ is the threshold $q$ at which both steak and the virtual alternative are maximal in the direction of vector $(1, -q)$. The optimal allocation for Theorists is steak if $q \leq \frac{30-v}{60-v}$, so only steak priced at 30 is offered, and is fish if $q \geq \frac{30-v}{60-v}$, so fish priced at $v$ and steak priced at 60 are offered, which extracts all surplus from both consumer types without excluding either type. Notice that Applied theorists strictly prefer steak priced at 60 to fish priced at $v$. This "slack" is what makes offering "surf-and-turf" more profitable when randomization is allowed.

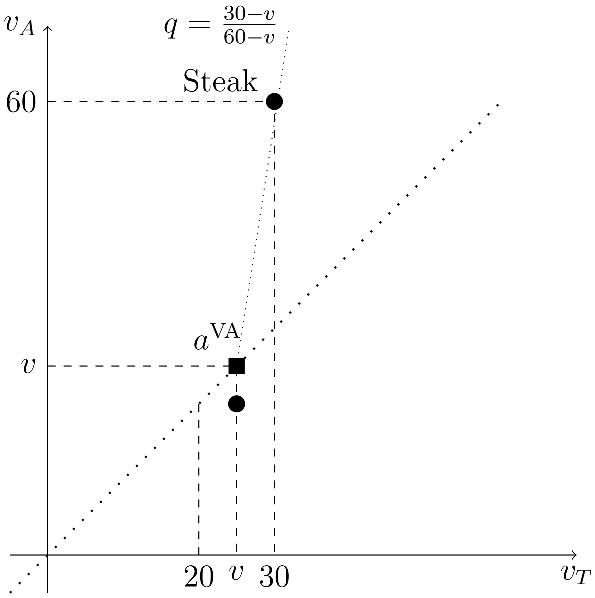



Figure 16: The case without randomization when $20 \leq v \leq 30$.

# C Normalizing Costs and Valuations for the Outside Option to Zero

We now show that we can normalize production costs and the valuations for the outside option to zero by changing the valuations in a way that maintains the same solutions to the principal's problem. For this, let $\hat{v}_t(a) = v_t(a) - c(a)$ for $t = 1, 2$ and every alternative $a \in A$, and consider the following problem.

$$\begin{aligned}
\max_{a_1,a_2 \in A \;\; r_1,r_2 \in \mathbb{R}} \quad & (1-q)r_1 + qr_2 \\
\text{subject to} \quad & \hat{v}_1(a_1) - r_1 \geq \hat{v}_1(0), \\
& \hat{v}_2(a_2) - r_2 \geq \hat{v}_2(0), \\
& \hat{v}_1(a_1) - r_1 \geq \hat{v}_1(a_2) - r_2, \\
& \hat{v}_2(a_2) - r_2 \geq \hat{v}_2(a_1) - r_1.
\end{aligned}$$

For any $a_1, a_2, r_1, r_2$, let $p_1 = r_1 + c(a_1)$ and $p_2 = r_2 + c(a_2)$. Then

$$(1-q)r_1 + qr_2 = (1-q)(p_1 - c(a_1)) + q(p_2 - c(a_2))$$

and each of IR1, IR2, IC1, and IC2 in the first problem holds if and only if the corresponding inequality in the other problem holds. Thus, the solutions to the second problem are the same as the ones to the first problem.

It remains to show that we can normalize the valuations for the outside option to zero. For this, let $\bar{v}_t(a) = \hat{v}_t(a) - \hat{v}_t(0)$ for $t = 1, 2$ and every alternative $a \in A$. Then the second problem is equivalent to the following one.

$$
\begin{aligned}
\max_{a_1,a_2 \in A \;\; r_1,r_2 \in \mathbb{R}} \quad & (1-q)r_1 + qr_2 \\
\text{subject to} \quad & \bar{v}_1(a_1) - r_1 \geq 0, \\
& \bar{v}_2(a_2) - r_2 \geq 0, \\
& \bar{v}_1(a_1) - r_1 \geq \bar{v}_1(a_2) - r_2, \\
& \bar{v}_2(a_2) - r_2 \geq \bar{v}_2(a_1) - r_1.
\end{aligned}
$$

In conclusion, we can normalize production costs and the valuations for the outside option to zero by changing type $t$'s valuation from $v_t(a)$ to $v_t(a) - c(a) - (v_t(0) - c(0))$ for every alternative $a \in A$. The cost $c(0)$ of the outside option represents the principal's disutility from not contracting with the agent, and is typically set to 0.

# D Increasing Differences

In this appendix we formally define the increasing differences assumption that is used in the screening literature to identify optimal menus and review the standard analysis under this assumption. We then discuss an alternative, weaker notion of increasing differences that our setting satisfies. This weaker notion is used in the literature to study what is *implementable* (instead of optimal).

## D.1 The Standard Analysis

We consider the standard setting often studied in the screening literature and review the corresponding standard arguments used to characterize optimal menus. This setting involves alternatives that are linearly ordered, with the outside option being the lowest alternative, and assumes that one of the types, say type 2, is willing to pay more than the other type, type 1, for any increase in the alternative. Type 2 is referred to as the "high type" and type 1 is referred to as the "low type." We formalize this standard setting as follows.

**Definition 3** *Increasing differences (ID) holds if:*

1. $A \subset \mathbb{R}_+$, *and*

2. *for any two alternatives* $a, a' \in A$ *with* $a > a'$,

$$v_2(a) - v_2(a') > v_1(a) - v_1(a').$$

Increasing differences implies that $v_2(a) \geq v_1(a)$ for every alternative $a$. This property implies that IR2 can be ignored when solving for the optimal menus, which is the first step in the analysis below. Without increasing differences, a different proof is required.

Increasing differences is satisfied, for example, if $v_t(a) = t \cdot a - c(a)$ for some cost function $c$ (recall that we normalized the valuations to be net of production costs). Increasing differences facilitates the analysis because it implies that the problem of finding an optimal menu $\{(a_1, p_1), (a_2, p_2)\}$ can be simplified as follows:[32]

1. IR2 can be ignored because the first inequality in

$$v_2(a_2) - p_2 \geq v_2(a_1) - p_1 \geq v_1(a_1) - p_1 \geq 0,$$

   follows from IC2, the second inequality follows from ID, the outside option 0 being the lowest alternative, and $v_1(0) = v_2(0) = 0$, and the last inequality follows from IR1.

2. IR1 holds as an equality (otherwise, because IR2 can be ignored, $p_1$ and $p_2$ can be increased by the same small amount).

3. IC2 holds as an equality (otherwise, because IR2 can be ignored, $p_2$ can be increased by a small amount).

[32]This standard analysis mostly follows Tadelis and Segal (2005), but also highlights some features of the optimal menus. Section 2.1.3 of Bolton and Dewatripont (2004) provides a similar analysis that leads to the same conclusions under some additional assumptions.

4. IC1 can be ignored. Indeed, if it cannot be ignored, then the solution to the problem subject only to IR1 and IC2 violates IC1. And then the principal could do better by offering both types the same (alternative, price) pair: $(a_2, p_2 + \epsilon)$ for a small $\epsilon > 0$ if $p_2 \geq p_1$, and $(a_1, p_1)$ otherwise. This trivially maintains IC2, and maintains IR1 because IC1 was violated.

Because IR1 and IC2 hold as equalities, $p_1$ and $p_2$ are pinned down by $a_1$ and $a_2$. Thus, because IR2 and IC1 can be ignored, the principal's problem becomes:

$$\max_{a_1, a_2 \in A} (1-q)v_1(a_1) + q(v_2(a_2) - v_2(a_1) + v_1(a_1)).$$

Grouping the elements that contain $a_1$ separately from the one that contains $a_2$ we obtain:

$$\max_{a_1, a_2 \in A} v_1(a_1) - qv_2(a_1) + qv_2(a_2).$$

Thus, we can solve for $a_1$ and $a_2$ separately. The optimal $a_2$ is clearly $a_2 = a_2^{\mathrm{E}}$, and the optimal $a_1$ maximizes $v_1(\cdot) - qv_2(\cdot)$. That is, the high type is optimally served efficiently, and the low type's allocation may be distorted in order to reduce the utility $v_2(a_1) - v_1(a_1)$ (information rent) of the high type. As the probability $q$ of the high type approaches 1, the low type's allocation $a_1$ becomes similar or identical to the outside option in that $(v_1(a_1), v_2(a_2))$ approaches $(0, 0)$.[33] These observations are summarized by the following characterization of the optimal menus.

**Theorem 2** *[Tadelis and Segal (2005)] If increasing differences holds, then for any probability $q \in (0, 1)$ of the high type, a menu $\{(a_1, p_1), (a_2, p_2)\}$ is optimal if and only if it satisfies the following properties:*

1. *The high type's alternative is efficient ($a_2 = a_2^{\mathrm{E}}$).*

2. *The high type's incentive constraint IC2 holds as an equality:*

$$v_2(a_2) - p_2 = v_2(a_1) - p_1.$$

[33] For $q = 1$ this is immediate because maximizing $v_1(a_1) - qv_2(a_1)$ is identical to minimizing the information rent $v_2(a_1) - v_1(a_1)$ of the high type, which is uniquely achieved in $A \subset \mathbb{R}_+$ by $a_1 = 0$. For $q$ close to 1 this follows from a continuity argument.

*3. The low type's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(a_1).$$

*4. The low type's alternative $a_1$ maximizes $v_1(\cdot) - qv_2(\cdot)$.*

*In particular, the low type's alternative pins down the menu and:*

*5. The low type's alternative may be inefficient ($a_1 \neq a_1^{\mathrm{E}}$) and he may be excluded ($a_1 = 0$).*

*6. As $q$ approaches 1, the low type is excluded or "almost excluded," that is, $(v_1(a_1), v_2(a_2))$ approaches $(0,0)$.*

In addition, under increasing differences the optimal menu coincides with the first-best if and only if the low type's efficient allocation is the outside option. This is because if $a_1^{\mathrm{E}} = 0$, then the menu $\{(0,0),(a_2^{\mathrm{E}}, v_2(a_2^{\mathrm{E}}))\}$ satisfies all the constraints; and if $a_1^{\mathrm{E}} \neq 0$, then $p_2 < v_2(a_2^{\mathrm{E}})$ in any menu $\{(a_1^{\mathrm{E}}, p_1),(a_2^{\mathrm{E}}, p_2)\}$ that satisfies IR1 and IC2, because $v_2(a_1^{\mathrm{E}}) - v_1(a_1^{\mathrm{E}}) > 0$ by increasing differences and $v_1(0) = v_2(0) = 0$. This is summarized by the following observation.

**Observation 0** *If increasing differences holds, then the optimal menu implements the first-best if and only if $a_1^{\mathrm{E}} = 0$.*

## D.2 A Weaker Notion

Our general setting satisfies a weaker notion of increasing differences. This weaker notion is used in the literature to study which allocations are *implementable*, but not which ones are optimal. We first formalize this notion and then discuss how it differs from the standard notion of increasing differences in Definition 3 that facilitates the standard characterization in Theorem 2.

**Definition 4** *Weak Increasing differences (WID) holds if there exists complete and transitive orders $\succeq_T$ and $\succeq_A$ on types and alternatives such that:*

$$t \succeq_T t', a \succeq_A a' \Rightarrow v_t(a) - v_{t'}(a) \geq v_t(a') - v_{t'}(a').$$

*Write $a \succ_A a'$ if $a \succeq_A a'$ but not $a' \succeq_A a$ and $a \backsim_A a'$ if $a \succeq_A a'$ and $a' \succeq_A a$.*

If a setting satisfies WID, then an allocation of alternatives to types can be implemented by some IC and IR menu if and only if the allocation assigns a higher ranked alternative to a higher ranked type (see, for example, Proposition 5.5 in Börgers, 2015).

Our setting satisfies this weaker notion without imposing any additional structure on values. For this, consider ranking type 2 above type 1, $2 \succeq_T 1$, and the following order $\succeq_2$ on the set $A$ of alternatives:

$$a \succeq_2 a' \iff v_2(a) - v_1(a) \geq v_2(a') - v_1(a').$$

That is, we assign to each alternative $a$ the value $v_2(a) - v_1(a)$, and order the alternatives by their value. Notice, however, that WID is also satisfied with the *reverse* orders: $1 \succeq_T 2$ and the following order $\succeq_1$ on the set $A$ of alternatives:

$$a \succeq_1 a' \iff v_1(a) - v_2(a) \geq v_1(a') - v_2(a').$$

Thus, according to WID either type can be considered the "high" or the "low" type, so WID alone cannot meaningfully single out one of the types and characterize the optimal menus. There must be some substantial difference between WID and increasing differences, which affects the optimal menus.

There are, in fact, two differences. The first is that under increasing differences, there are no ties in the rankings of alternatives, but WID allows for ties. In fact, in many applications, including the ones we consider in Section 5, it is natural for many alternatives to have the same value $v_2(\cdot) - v_1(\cdot)$. And different alternatives with the same value can play different roles in the optimal menus. For example, $a^{\text{VA}} \backsim_2 0$ (and $a^{\text{BD1}} \backsim_2 0$ if $V$ is convex) but, as Theorem 1 shows, $a^{\text{VA}}$ and $a^{\text{BD1}}$ play a different and significant role from the outside option 0 in the optimal menus. The second difference is that increasing differences implicitly assumes that the outside option is the unique lowest alternative. This is because the outside option is 0 and property 1 in Definition 3 states that $A \subset \mathbb{R}_+$. With $\succ_2$, in contrast, the outside option may be weakly higher than any number of alternatives.

This distinction is a key difference between the case of increasing differences and the general case. Indeed, if $a_2^{\text{E}} \succeq_2 0 \succeq_2 a_1^{\text{E}}$, then both types are uncontested and the first-best is implementable. Otherwise, either $a_2^{\mathrm{E}} \succeq_2 a_1^{\mathrm{E}} \succ_2 0$ or $0 \succ_2 a_2^{\mathrm{E}} \succeq_2 a_1^{\mathrm{E}}$. In the former case, type 1 is contested (and type 2 is uncontested)..[34] In the latter case, type 2 is contested (and type 1 is uncontested). But in both cases 0 may not be the unique $\succ_2$-lowest alternative in $A$.

If $a_2^{\mathrm{E}} \succeq_2 a_1^{\mathrm{E}} \succ_2 0$, which is the case we have considered, but 0 is not the unique $\succ_2$-lowest alternative in $A$, it might be tempting to restrict attention to the alternatives that $\succeq_2$-exceed 0 and remove the other alternatives. This would be a mistake for two reasons. First, if $0 \succ_2 a^{\mathrm{BD1}}$, then, as Theorem 1 shows, removing $a^{\mathrm{BD1}}$ affects the optimal menus when IR2 holds as an equality. This is also true if $a^{\mathrm{BD1}} \backsim_2 0$ but $v_1(a^{\mathrm{BD1}}) > 0$. Only if $a^{\mathrm{BD1}}$ coincides with the outside option is it without loss to remove the alternatives that are $\succ_2$-lower than 0. In this case we are essentially in an increasing differences setting, IC2 always holds as an equality in an optimal menu, and type 1 is excluded or "almost excluded" as $q$ approaches 1. Finally, if randomization is allowed and 0 is the unique $\succ_2$-lowest alternative in $A$, which increasing differences satisfies, then randomization does not help.

In conclusion, what makes increasing differences special is that it identifies $\succ_2$ as the "correct" order on the alternatives and it implies that 0 is the unique $\succ_2$-lowest alternative and that no two alternatives are $\succ_2$-equivalent. Figure 17 illustrates the order $\succ_2$, a case in which 0 is the unique $\succ_2$-lowest element in $A$ in panel (a), and two cases in which 0 is not the unique $\succ_2$-lowest element in $A$ in panels (b) and (c).

# E Proof of Corollary 1

The following lemma, which will be used below to prove Corollary 1, shows that $q^{\mathrm{OD}}$ is well defined. Moreover, when $q^{\mathrm{OD}} < 1$, we have that $B_1(q^{\mathrm{OD}}) \cap A^{=}$ is the singleton $a^{\mathrm{BD1}} = a^{\mathrm{VA}}$, the best alternative for type 1 that lies weakly below the diagonal. In addition, $q^{\mathrm{OD}}$ coincides with the threshold $q^{\mathrm{BD1}}$ from Theorem 1.

**Lemma 3** *Suppose that type 2 is uncontested (Assumption 1), type 1 is contested, and $V$ is convex. Then*

1. *The threshold $q^{\mathrm{OD}} \leq 1$ defined in* (OD) *exists.*

[34] This is because $a_1^{\mathrm{E}} \succ_2 0$ means that $v_2(a_1^{\mathrm{E}}) - v_1(a_1^{\mathrm{E}}) > v_2(0) - v_1(0) = 0$, where the equality follows from our normalization $v_1(0) = v_2(0) = 0$.

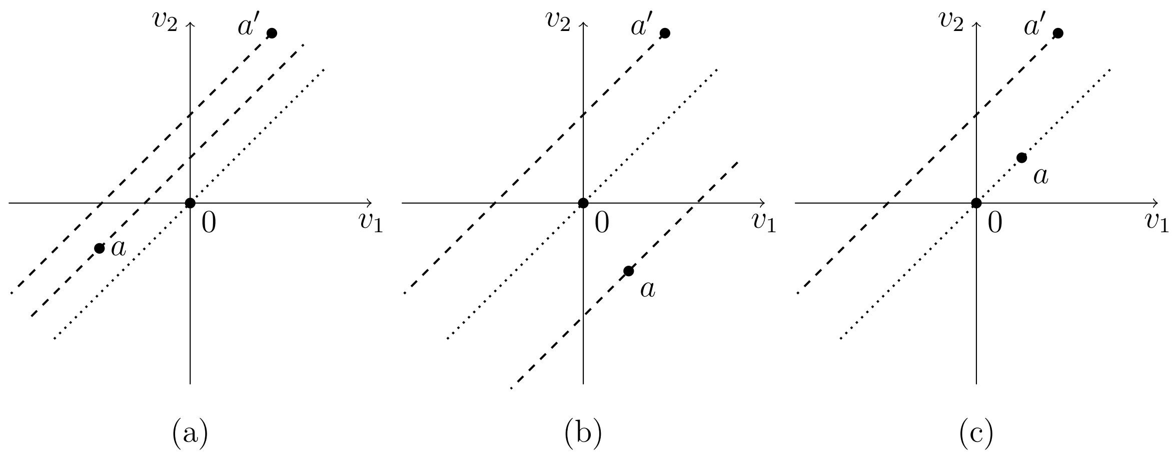



Figure 17: (a) $a' \succ_2 a \succ_2 0$. (b) $a' \succ_2 0 \succ_2 a$. (c) $a' \succ_2 0 \backsim_2 a$.

2. *If $q < q^{\mathrm{OD}}$, then $B_1(q) \subseteq A^+$.*

3. *Alternative $a^{\mathrm{BD1}}$ lies on the diagonal, that is, $a^{\mathrm{VA}} = a^{\mathrm{BD1}}$.*

4. *If $q^{\mathrm{OD}} < 1$, then $B_1(q^{\mathrm{OD}}) \cap A^= = \{a^{\mathrm{BD1}}\} = \{a^{\mathrm{VA}}\}$.*

5. *Threshold $q^{\mathrm{OD}}$ coincides with threshold $q^{\mathrm{BD1}}$ defined in* (q-BD1).

**Proof.** We first show that a $q^{\mathrm{OD}}$ that satisfies (OD) exists. For this, let

$$\Delta v(q) := \min_{a \in B_1(q)} v_2(a) - v_1(a),$$

be the lowest difference between $v_2 - v_1$ among all the alternatives in $B_1(q)$, which exists because the set of valuation pairs in $B_1(q)$ is closed and convex.

We first observe that, by Lemma 2, $\Delta v(q)$ is non-increasing in $q$. We also observe that $\Delta v(0) > 0$ and $\Delta v(1) \leq 0$. Indeed, $B_1(0) = \{a_1^{\mathrm{E}}\}$ and $v_2(a_1^{\mathrm{E}}) > v_1(a_1^{\mathrm{E}})$ because type 1 is contested, so $\Delta v(0) > 0$. And if $a \in B_1(1)$, then, by definition, for alternative $a' = 0$ we have

$$v_1(a) - v_2(a) \geq v_1(a') - v_2(a') = 0,$$

so $\Delta v(1) \leq 0$.

Now consider

$$q^* = \sup\{q \mid \Delta v(q) > 0\}.$$

Because $\Delta v(q)$ is non-increasing, we have $\Delta v(q) > 0$ for all $q < q^*$ and $\Delta v(q) \leq 0$ for all $q > q^*$.

We argue that $B_1(q^*)$ contains an alternative on the diagonal by showing that there exist alternatives $a, a' \in B_1(q^*)$ such that

$$v_2(a) - v_1(a) \geq 0 \text{ and } v_2(a') - v_1(a') \leq 0.$$

This suffices because $B_1(q^*)$ is convex. To see that such alternatives $a$ and $a'$ exist, consider a sequence $a(q)$ of alternatives as $q$ approaches $q^*$ from below, where $a(q)$ is an alternative in $B_1(q)$ that minimizes $v_2 - v_1$. Because $V$ is closed, there exists some $v^* \in V$ such that

$$(v_1^*, v_2^*) = \lim_{q \uparrow q^*} (v_1(a(q)), v_2(a(q))).$$

Let $a^*$ be an alternative with $v^* = (v_1(a^*), v_2(a^*))$. Because $v_1(a) - qv_2(a)$ is continuous in $q$, we have that $a^* \in B_1(q^*)$. And because $v_2(a(q)) - v_1(a(q)) > 0$ for all $q$ in the sequence,

$$v_2(a^*) - v_1(a^*) \geq 0.$$

A similar argument, but approaching $q^*$ from above, shows that there exists an alternative $a' \in B_1(q^*)$ such that

$$v_2(a') - v_1(a') \leq 0.$$

We now claim that $q^{\mathrm{OD}} = q^*$. We have already shown that $B_1(q^*)$ contains an alternative on the diagonal. That there is no $q < q^*$ with this property follows because $\Delta v(q) > 0$ for all $q < q^*$ means that $v_2(a) - v_1(a) > 0$ for all $a \in B_1(q)$ and $q < q^*$.

We now show the second statement. As argued above, for any $q < q^{\mathrm{OD}}$, $a \in B_1(q)$, and $a' \in B_1(q^{\mathrm{OD}})$, we have that $v_2(a) - v_1(a) \geq v_2(a') - v_1(a')$. So, because

$v_2(a') - v_1(a') = 0$ for some $a' \in B_1(q^{\text{OD}})$, we have that $v_2(a) - v_1(a) \geq 0$ for all $a \in B_1(q)$, so $B_1(q) \subseteq A^+$.

To see that $a^{\text{BD1}}$ lies on the diagonal, so $a^{\text{VA}} = a^{\text{BD1}}$, suppose that $a^{\text{BD1}}$ lies strictly below the diagonal, so $v_1(a^{\text{BD1}}) > v_2(a^{\text{BD1}})$. Since type 1 is contested, $a_1^{\text{E}}$ lies strictly above the diagonal. Thus, the line connecting $a^{\text{BD1}}$ and $a_1^{\text{E}}$ intersects the diagonal at a single point, which corresponds to some alternative $a$ by convexity of $V$. And $v_1(a) > v_1(a^{\text{BD1}})$ because $v_1(a_1^{\text{E}}) > v_1(a^{\text{BD1}})$ by definition of $v_1(a_1^{\text{E}})$. This contradicts the definition of $a^{\text{BD1}}$, so we conclude that $a^{\text{BD1}}$ lies on the diagonal.

Now suppose that $q^{\text{OD}} < 1$. The set $B_1(q^{\text{OD}})$ is the set of alternatives in $V$ that lie on some line $L$ whose slope strictly exceeds 1 because $L$ is orthogonal to vector $(1, -q^{\text{OD}})$. Because $V$ is convex and closed, $B_1(q^{\text{OD}})$ is a convex and closed segment of $L$, which by definition of $q^{\text{OD}}$ has non-empty intersection with the diagonal. Because the slope of this segment strictly exceeds 1, the point of intersection is unique. Let $\tilde{a}_1$ be the unique element in $B_1(q^{\text{OD}}) \cap A^=$.

To see that $\tilde{a}_1 = a^{\text{BD1}}$, consider any alternative $a \in A^-$, so $v_1(a) \geq v_2(a)$, and write

$$(1 - q^{\text{OD}})v_1(a) \leq v_1(a) - q^{\text{OD}}v_2(a) \leq v_1(\tilde{a}_1) - q^{\text{OD}}v_2(\tilde{a}_1) = (1 - q^{\text{OD}})v_1(\tilde{a}_1),$$

where the second inequality follows because $\tilde{a}_1 \in B_1(q^{\text{OD}})$, and the equality follows because $\tilde{a}_1$ is on the diagonal. That $(1 - q^{\text{OD}})v_1(a) \leq (1 - q^{\text{OD}})v_1(\tilde{a}_1)$ implies that $v_1(a) \leq v_1(\tilde{a}_1)$, so $\tilde{a}_1$ coincides with $a^{\text{BD1}}$ by definition of $a^{\text{BD1}}$ (which is assumed to be unique).

We conclude the proof by showing that $q^{\text{OD}} = q^{\text{BD1}}$, where $q^{\text{BD1}}$ is defined in (q-BD1). We showed in the first part of the proof that $\Delta v(q) > 0$ for every $q < q^{\text{OD}}$. This implies, by definition of $q^{\text{BD1}}$, that $q^{\text{OD}} \leq q^{\text{BD1}}$. In particular, if $q^{\text{OD}} = 1$, then $q^{\text{OD}} = q^{\text{BD1}}$. And we saw that if $q^{\text{OD}} < 1$, then $a^{\text{BD1}} \in B_1(q^{\text{OD}})$. This implies, again by definition of $q^{\text{BD1}}$, that $q^{\text{BD1}} \leq q^{\text{OD}}$. ■

Given Lemma 3, we now complete the proof of Corollary 1.

**Proof of Corollary 1.** Property 6 follows from property 6 of Theorem 1 because IC2 holds as an equality for all $q \in (0, 1)$. Property 10 follows from properties 7 and 8 of Theorem 1 along with the fact that, by linearity, the alternatives in $A_1(q)$ lie on the boundary of $V$.

We now prove properties 3(a), (b), and (c). Because $a^{\mathrm{VA}} = a^{\mathrm{BD1}} \in A^{=}$ (by statement (3) of Lemma 3),

$$\operatorname*{argmax}_{a \in A^{+} \cup \{a^{\mathrm{VA}}\}} v_1(a) - qv_2(a) = \operatorname*{argmax}_{a \in A^{+}} v_1(a) - qv_2(a),$$

so $a_1$ is any alternative in

$$\operatorname*{argmax}_{a \in A^{+}} v_1(a) - qv_2(a).$$

Consider $q < q^{\mathrm{OD}}$. By statement (2) of Lemma 3, we have that $B_1(q) \subseteq A^{+}$, that is,

$$\operatorname*{argmax}_{a \in A} v_1(a) - qv_2(a) = \operatorname*{argmax}_{a \in A^{+}} v_1(a) - qv_2(a),$$

so $a_1$ is any alternative in $B_1(q)$, and IC2 holds as an equality by property 4 of Theorem 1.

Consider $q = q^{\mathrm{OD}}$. Because $B_1(q^{\mathrm{OD}})$ has an alternative in $A^{+}$,

$$\max_{a \in A} v_1(a) - qv_2(a) = \max_{a \in A^{+}} v_1(a) - qv_2(a),$$

so $a_1$ is any alternative in $B_1(q^{\mathrm{OD}}) \cap A^{+}$. The binding constraints follow from properties 4 and 5 of Theorem 1.

Consider $q > q^{\mathrm{OD}}$. Since $q^{\mathrm{OD}} = q^{\mathrm{BD1}}$ and $a^{\mathrm{BD1}} \in A^{=} \subset A^{+}$ (by statement (3) of Lemma 3), property 3(b) of Theorem 1 implies that $a_1 = a^{\mathrm{BD1}}$, and properties 4 and 5 of Theorem 1 imply that both IC2 and IR2 hold as equalities. ■

# F Proof of Corollary 2

**Proof.** Denote by

$$V^{\mathrm{R}} = \{(v_1(x), v_2(x)) : x \text{ is a } \textit{random} \text{ alternative}\}$$

the set of valuation pairs associated with random alternatives. Because $V^{\mathrm{R}}$ is convex, Corollary 1 characterizes the optimal menus. Thus, properties 1 and 2 hold, and it remains only to prove property 3.

For this consider the set

$$B^{\mathrm{R}}(q) = \operatorname*{argmax}_{x \in \Delta(A)} v_1(x) - qv_2(x),$$

of maximizers of $v_1 - qv_2$ over all random alternatives. By linearity of $v_1 - qv_2$, $B^{\mathrm{R}}(q)$ is the convex hull of the set $B_1(q)$ of all (non-random) alternatives that maximize $v_1 - qv_2$, that is, $\Delta(B_1(q)) = B^{\mathrm{R}}(q)$, so $B_1(q) \subseteq B^{\mathrm{R}}(q)$. Because $V^{\mathrm{R}}$ is convex, Lemma 3 shows that there is a lowest threshold, call it $q^{OD-R}$, such that some (random) alternative $x \in B^{\mathrm{R}}(q^{OD-R})$ is on the diagonal, that is,

$$v_1(x) = v_2(x).$$

Notice that for any $q$, because $\Delta(B_1(q)) = B^{\mathrm{R}}(q)$, there is a random alternative $x \in B^{\mathrm{R}}(q)$ with $v_1(x) - v_2(x) = 0$ if and only if there are two (deterministic) alternatives $a, a' \in B_1(q)$ such that $v_1(a) - v_2(a) \geq 0$ and $v_1(a') - v_2(a') \leq 0$, that is, $B_1(q) \cap A^+$ and $B_1(q) \cap A^-$ are non-empty. Thus, $q^{\mathrm{CD}} = q^{OD-R} = q^{\mathrm{BD1}}$, where the second equality follows from part 5 of Lemma 3.

Now consider the three cases in property 3 and apply Corollary 1. If $q < q^{\mathrm{CD}} = q^{OD-R}$, then $x_1$ is any alternative in $B^{\mathrm{R}}(q) = \Delta(B_1(q))$, that is, any distribution over the alternatives in $B_1(q)$, and IC2 holds as an equality.

Next consider $q = q^{\mathrm{CD}} = q^{OD-R}$. Then $x_1$ is any random alternative in $B^{\mathrm{R}}(q^{OD-R})$ that is on or above the diagonal. These are distributions $x$ over alternatives in $B_1(q^{\mathrm{CD}})$ such that

$$v_2(x) \geq v_1(x),$$

and IC2 holds as an equality. If the inequality holds as an equality, then IR2 also holds as an equality.

Finally consider $q > q^{\mathrm{CD}} = q^{OD-R}$. Then, $x_1$ is the "unique alternative" in $B_1(q^{OD-R}) \cap A^+ \cap A^-$. Thus, $x_1$ is any random alternative in $B^{\mathrm{R}}(q^{OD-R})$ that is on the diagonal, that is, any distribution $x$ over the alternatives in $B_1(q^{\mathrm{CD}})$ such that $v_1(x) = v_2(x)$, and both IC2 and IR2 hold as equalities. ■

# G Characterizing the Optimal Menus in the Horizontal and Vertical Differentiation Application

Consistent with our model, we normalize costs to zero by reflecting the cost in the valuations and writing $v_t(a) = s(t)s(a)cos(2\pi(\ell(t) - \ell(a))) - cs(a)$. To simplify notation we write $cos(t, a)$ instead of $cos(2\pi(\ell(t) - \ell(a)))$, so we have $v_t(a) = s(t)s(a)cos(t, a) - cs(a)$. In inner product form, this is

$$v_t(a) = \vec{u}(t) \cdot \vec{u}(a) - cs(a).$$

Without loss of generality we assume that the size of type 2 weakly exceeds that of type 1, and we normalize the size of type 2 to 1, so $s(2) = 1$. We also assume without loss of generality that the location of type 2 is 0.

The following result formalizes the description of the optimal menus from Section 5.1.

**Proposition 1** *Suppose without loss of generality that* $s(2) = 1 \geq s(1)$*,* $l(2) = 0$*, and* $c < 1$*. For any probability* $q \in (0, 1)$ *of type 2, a menu* $\{(a_1, p_1), (a_2, p_2)\}$ *is optimal if and only if it satisfies the following properties:*

1. *Type 2's allocation is efficient (*$a_2 = a_2^{\mathrm{E}}$ *is the unit vector* $(0, 1)$*).*

2. *Type 1's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(a_1).$$

3. *If type 1 is in region 1 in Figure 8, (b), then the first-best is implementable and* $a_1$ *is the outside option.*

4. *If type 1 is in region 2 in Figure 8, (b), then the first-best is implementable and* $a_1$ *is the unit vector in the direction of* $\vec{u}(1)$*.*

5. *If type 1 is in region 3.a or 3.b in Figure 8, (b), then* $a_1$ *and type 2's binding constraints are specified given* $q$ *as follows.*

   (a) *If type 1 is in region 3.a, then* $q^{\mathrm{OD}}$ *is such that* $\vec{u}(q^{\mathrm{OD}})$ *lies on the boundary of the circle with radius* $\frac{1}{2}$ *centered at* $(\frac{1}{2}, 0)$ *and:*

i. *If* $q \le q^{\mathrm{OD}}$, *then* $a_1$ *is the unit vector in the direction of* $\vec{u}(q)$ *and IC2 holds as an equality.*

ii. *If* $q > q^{\mathrm{OD}}$, *then* $a_1$ *is the unit vector in the direction of* $\vec{u}(q^{\mathrm{OD}})$ *and both IC2 and IR2 hold as equalities.*

(b) *If type 1 is in region 3.b, then* $q^{\mathrm{OD}}$ *is such that* $\vec{u}(q^{\mathrm{OD}})$ *lies on the boundary of the circle with radius* $c$ *centered at* $(0,0)$ *and:*

i. *If* $q < q^{\mathrm{OD}}$, *then* $a_1$ *is the unit vector in the direction of* $\vec{u}(q)$ *and IC2 holds as an equality.*

ii. *If* $q = q^{\mathrm{OD}}$, *then* $a_1$ *is either the outside option or any vector in the direction of* $\vec{u}(q)$, *and both IC2 and IR2 hold as equalities.*

iii. *If* $q > q^{\mathrm{OD}}$, *then* $a_1$ *is the outside option and both IC2 and IR2 hold as equalities.*

As outlined in Section 5.1, the efficient alternative of each type is the unit vector in the direction of that type if the size of the type is more than $c$, and is the outside option if the size is less than $c$. Moreover, the first-best is clearly implementable if type 1 is either in region 1 or region 2 of Figure 8 (b). To complete the proof of Proposition 1, it thus remains to prove property 5, which corresponds to type 1 being in region 3.a or 3.b.

To characterize the optimal menus in this case, we observe that the set of valuation pairs of the two types across all alternatives is convex. Thus, we can apply Corollary 1, which shows that type 2's allocation is efficient and IR1 and IC2 hold as equalities. Any optimal menu is therefore pinned down by type 1's alternative $a_1$. To identify $a_1$, we need to identify the set $B_1(q)$ of alternatives that maximize $v_1(a) - qv_2(a)$ across all alternatives, where $q$ is the probability of type 2.

Among all the alternatives that have some fixed size $\bar{s}$, the alternative $a$ that maximizes the objective is in the direction of $\vec{u}(q)$, in which case the value of the objective is $\vec{u}(q) \cdot \vec{u}(a) - c\bar{s} = \Big(s(\vec{u}(q)) - c\Big)\bar{s}$. Thus, we can focus on maximizing $\Big(s(\vec{u}(q)) - c\Big)s(a)$. If $s(\vec{u}(q)) > c$, then the unique maximizer is the unit vector in the direction of $\vec{u}(q)$. If $s(\vec{u}(q)) < c$, then the unique maximizer is the outside option. If $s(\vec{u}(q)) = c$, then the outside option and every vector in the direction of $\vec{u}(q)$ are maximizers. This characterizes the set $B_1(q)$.

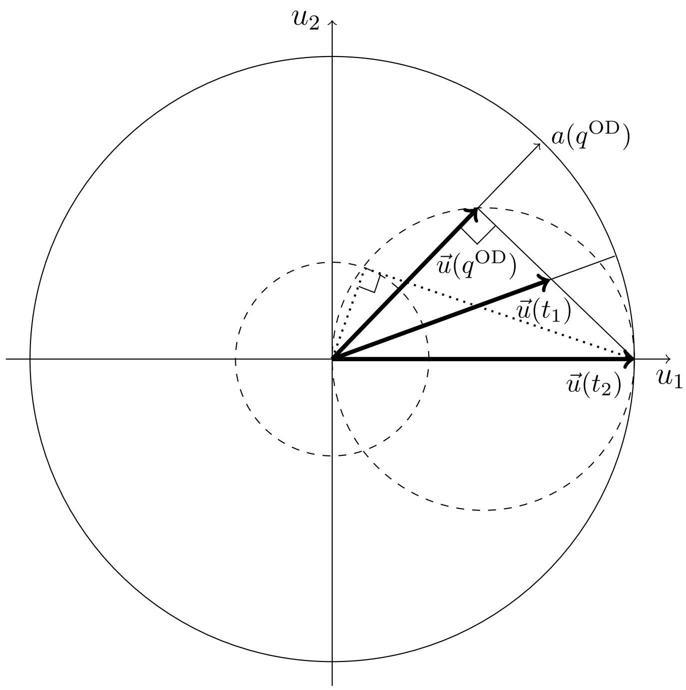



Figure 18: Vector $\vec{u}(t_1)$ is above the tangent from $\vec{u}(t_1)$ to the circle with radius $c$ centered at the origin. Threshold $q^{\mathrm{OD}}$ is such that $\vec{u}(q^{\mathrm{OD}})$ is on the boundary of the circle with radius 0.5 centered at $(0, 0.5)$.

Now consider two cases. The first case is when $\vec{u}(1)$ lies in region 3.a. In this case, the line that connects $\vec{u}(2)$ and $\vec{u}(1)$ never enters the circle with radius $c$, as shown in Figure 18. So $\vec{u}(q)$ is outside the circle with radius $c$, that is, $s(\vec{u}(q)) > c$, for all $q$ and therefore $B_1(q)$ consists of the unit vector in the direction of $\vec{u}(q)$, which we denote by $a(q)$. The threshold $q^{\mathrm{OD}}$ defined in (OD) is such that

$$v_1(a(q^{\mathrm{OD}})) = v_2(a(q^{\mathrm{OD}})).$$

For both types to have the same valuation for $a(q^{\mathrm{OD}})$, the projections of $\vec{u}(1)$ and $\vec{u}(2)$ onto $a(q^{\mathrm{OD}})$ must have the same size, which means that $\vec{u}(q^{\mathrm{OD}})$ is on the boundary of the circle with radius 0.5 centered at $(0, 0.5)$, as shown in Figure 18. By Corollary 1, type 1's alternative $a_1$ is $a(q)$ for all $q \leq q^{\mathrm{OD}}$, and $a_1 = a(q^{\mathrm{OD}})$ for all $q > q^{\mathrm{OD}}$ (since $a^{\mathrm{BD1}} = a(q^{\mathrm{OD}})$ is the unique element in $B_1(q^{\mathrm{OD}}) \cap A^{=}$).

The second case is when $\vec{u}(1)$ lies in region 3.b. In this case, as shown in Figure 19, the line that connects $\vec{u}(2)$ to $\vec{u}(1)$ touches the circle at some point. This point

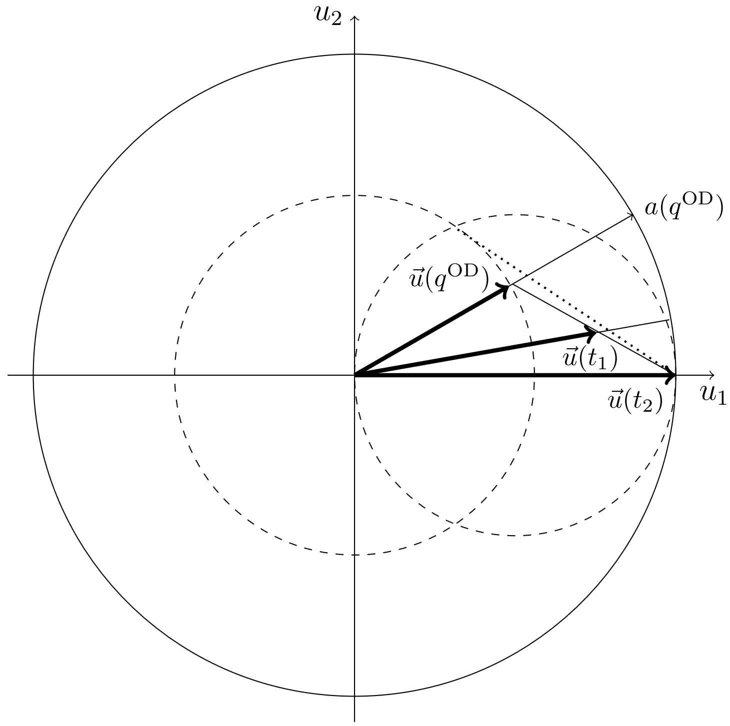



Figure 19: Vector $\vec{u}(t_1)$ is below the tangent from $\vec{u}(t_1)$ to the circle with radius $c$. The threshold $q^{\mathrm{OD}}$ is such $\vec{u}(q^{\mathrm{OD}})$ has size $c$.

corresponds to $q^{\mathrm{OD}}$. Indeed, for $q < q^{\mathrm{OD}}$, we have that $s(\vec{u}(q)) > c$, so $B_1(q^{\mathrm{OD}})$ is the single alternative $a(q)$ corresponding to the unit vector in the direction of $\vec{u}(q)$. Since the projection of $\vec{u}(2)$ onto the unit vector in the direction of this vector $\vec{u}(q)$ is larger than the projection of $\vec{u}(1)$ onto this vector, we have that $v_2(a(q)) > v_1(a(q))$. For $q > q^{\mathrm{OD}}$, we have that $s(\vec{u}(q)) < c$, so the single alternative in $B_1(q)$ is the outside option, for which both types have valuation 0. Thus, $q^{\mathrm{OD}}$ given by (OD) is such that $\vec{u}(q)$ touches the circle with radius $c$, so $s(\vec{u}(q^{\mathrm{OD}})) = c$ and $B_1(q^{\mathrm{OD}})$ consists of the outside option and all the vectors in the direction of $\vec{u}(q^{\mathrm{OD}})$, for which type 2's valuation is weakly higher than type 1's. By Corollary 1, type 1's alternative $a_1$ is $a(q)$ (the unit vector in the direction of $\vec{u}(q^{\mathrm{OD}})$) for all $q < q^{\mathrm{OD}}$, $a_1$ is the outside option for all $q > q^{\mathrm{OD}}$, and can be the outside option or any vector in the direction of $\vec{u}(q^{\mathrm{OD}})$ for $q = q^{\mathrm{OD}}$.

# H Interpretation of Giving a Product with Negative Valuation

Recall that valuations in our setting are normalized and represent willingness-to-pay net of costs. To interpret what “giving a product with negative valuation” means in the pre-normalized setting, consider two cases. First, suppose every type has a positive willingness-to-pay (so products are “goods”) but costs might be higher than willingness-to-pay so normalized valuations are negative. Then the seller might sell a product for which both normalized valuations are negative, even though he is “losing money” on that product because it costs more to produce them than either type is willing to pay. Second, suppose some type might have a negative willingness-to-pay for a product (so some products are “bads”), in which case we have to distinguish the cases where free disposal is possible and when it is not (free disposal might be natural with pre-mixed products such as nutritional supplements or assembled devices, but not natural in some settings such as a cable TV subscription). Without free disposal, the interpretation is the same: products might be assigned even if it costs more to produce them than either type is willing to pay. With free disposal, if a type’s willingness to pay for a product is negative, then we can replace it with zero and assume no-free-disposal without loss. But then, if both types have negative willingness-to-pay for a product, those valuations are replaced with zero. Then with positive costs, our result says the seller would never find it optimal to assign such a product.

# I Bundling Characterization with Any Number of Products

If both types are uncontested, then the first-best is implementable. So, for the rest of the discussion, assume that type 2 is uncontested (Assumption 1), as we did in Theorem 1, and that type 1 is contested, so $v_2(a_1^{\mathrm{E}}) > v_1(a_1^{\mathrm{E}})$. This assumption means that

$$\sum_{i:v_1(i)>0} \Big(v_2(i) - v_1(i)\Big) > 0.$$

We characterize the optimal random menus by using Corollary 2. An optimal menu is pinned down by type 1's random alternative $x_1$. To characterize this random alternative, we identify the set

$$B_1(q) = \operatorname*{argmax}_{a \in A} v_1(a) - qv_2(a) = \operatorname*{argmax}_{a \in A} \sum_{i \in a} \Big(v_1(i) - qv_2(i)\Big).$$

By definition, a bundle $a$ is in $B_1(q)$ if and only if it contains every product $i$ for which $v_1(i) - qv_2(i) > 0$ and contains no product $i$ for which $v_1(i) - qv_2(i) < 0$, that is,

$$\{i \mid v_1(i) - qv_2(i) > 0\} \subseteq a \subseteq \{i \mid v_1(i) - qv_2(i) \geq 0\}.$$

Each set of products in $B_1(q)$ can be visualized as a bundle that contains every product to the right of the hyperplane that goes through the set of points $(v_1, v_2)$ such that $v_1 - qv_2 = 0$, and may or may not contain the products on the hyperplane, as shown in Figure 20, (a).

Corollary 2 says that there exists a threshold $q^{\text{CD}}$ such that $B_1(q^{\text{CD}}) \cap A^+$ and $B_1(q^{\text{CD}}) \cap A^-$ are non-empty. That is, there are two alternatives $a, a' \in B_1(q^{\text{CD}})$ such that

$$\sum_{i \in a} v_2(i) - v_1(i) \geq 0, \text{ and,} \sum_{i \in a'} v_2(i) - v_1(i) \leq 0.$$

Corollary 2 singles out the alternatives in $B_1(q^{\text{CD}})$ that maximize or minimize $v_2 - v_1$. These alternatives are easy to characterize in our application. The alternative that maximizes $v_2 - v_1$ among those in $B_1(q)$ is the bundle that, among the products $i$ that are on the hyperplane $v_1 - qv_2 = 0$, contains exactly those for which $v_2(i) - v_1(i)$ is positive. Because $0 < q < 1$, those are the products that are in the first quadrant,

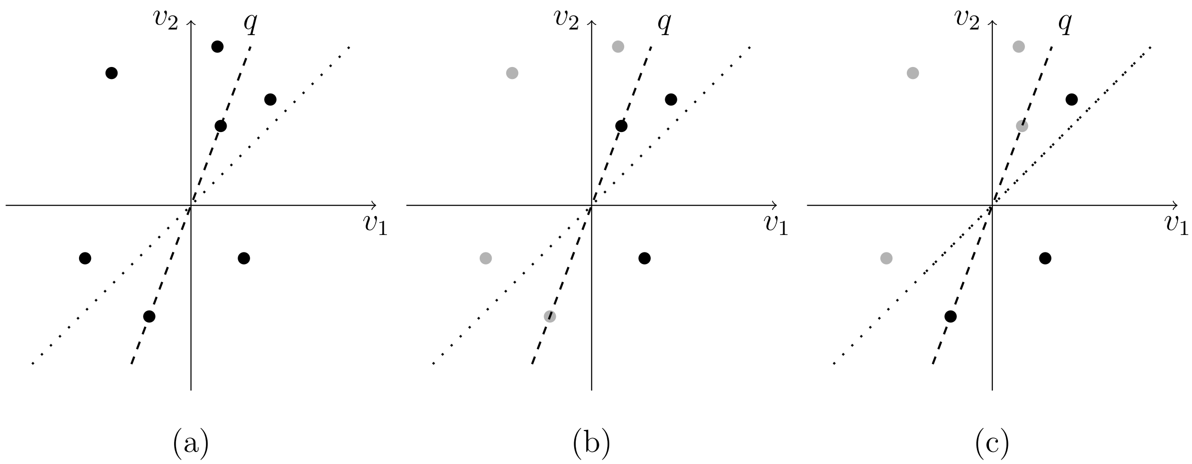



Figure 20: (a) A bundle in $B_1(q)$ contains the products to the right of the hyperplane $q$ and may or may not contains the products on the hyperplane. (b) $a^{1Q}(q)$ contains the products shaded dark. (c) $a^{3Q}(q)$ contains the products shaded dark.

as shown in Figure 20, (b). Formally, the maximizer is the bundle

$$a^{1Q}(q) = \left\{ i \;\middle|\; \Big(v_1(i) - qv_2(i) > 0\Big) \text{ or } \Big(v_1(i) - qv_2(i) \geq 0, v_1(i), v_2(i) \geq 0\Big) \right\}. \quad (6)$$

Similarly, the minimizer of $v_2 - v_1$ among those in $B_1(q)$ is the bundle that contains the products on the hyperplane that are in the third quadrant, as shown in Figure 20, (c), that is,

$$a^{3Q}(q) = \left\{ i \;\middle|\; \Big(v_1(i) - qv_2(i) > 0\Big) \text{ or } \Big(v_1(i) - qv_2(i) \geq 0, v_1(i), v_2(i) \leq 0\Big) \right\}. \quad (7)$$

Given the above definitions, we state our characterization result as follows.

**Proposition 2** *Suppose that randomization is allowed and* $\sum_{i:v_1(i)>0}(v_2(i) - v_1(i)) > 0$. *For any probability* $q \in (0,1)$ *of type 2, a menu* $\{(x_1, p_1), (x_2, p_2)\}$ *is optimal if and only if it satisfies the following properties:*

1. *Type 2's allocation is efficient* ($x_2 = a_2^{\mathrm{E}} = \{i \mid v_2(i) > 0\}$).

2. *Type 1's individual rationality constraint IR1 holds as an equality:*

$$p_1 = v_1(x_1).$$

3. *Type 1's allocation and type 2's binding constraints are specified given $q$ as follows:*

   (a) *If $q < q^{\mathrm{CD}}$, then $x_1$ is any randomized bundle that contains every product $i$ with $v_1(i) - qv_2(i) > 0$ with probability 1 and every product $i$ with $v_1(i) - qv_2(i) < 0$ with probability 0, and IC2 holds as an equality.*

   (b) *If $q = q^{\mathrm{CD}}$, then $x_1$ is any randomized bundle that contains every product $i$ with $v_1(i) - qv_2(i) > 0$ with probability 1 and every product $i$ with $v_1(i) - qv_2(i) < 0$ with probability 0 and such that*

   $$E_{a \sim x_1}[v_2(a) - v_1(a)] \geq 0$$

   *and IC2 holds as an equality. If the above inequality holds as an equality, then additionally IR2 holds as an equality.*
   *One particular $x_1$ that satisfies the above inequality assigns probability 1 to bundle $a^{1Q}(q^{\mathrm{CD}})$ defined in* (6).

   (c) *If $q > q^{\mathrm{CD}}$, then $x_1$ is any randomized bundle that contains every product $i$ with $v_1(i) - qv_2(i) > 0$ with probability 1 and every product $i$ with $v_1(i) - qv_2(i) < 0$ with probability 0 and such that*

   $$E_{a \sim x_1}[v_2(a) - v_1(a)] = 0$$

   *and IR2 holds as an equality.*
   *One particular $x_1$ that satisfies the above equality randomizes only over the two bundles $a^{1Q}(q^{\mathrm{CD}})$ and $a^{3Q}(q^{\mathrm{CD}})$, defined in* (6) *and* (7).

One possibility for generating the alternative $x_1$ in property 3(c) of Proposition 2 is to randomize only over the two alternatives $a^{3Q}(q^{\mathrm{CD}})$ and $a^{1Q}(q^{\mathrm{CD}})$. In the resulting random alternative, products to the right of the hyperplane $v_1 - qv_2 = 0$ are included with probability 1, each product on the hyperplane in the first quadrant is included with some probability $p$, and each product on the hyperplane in the third quadrant is included with the complementary probability $1-p$. Except for knife edge cases, however, there is a single product $i$ on the hyperplane. The resulting random alternative $x_1$ then consists of all the products to the right of the hyperplane along

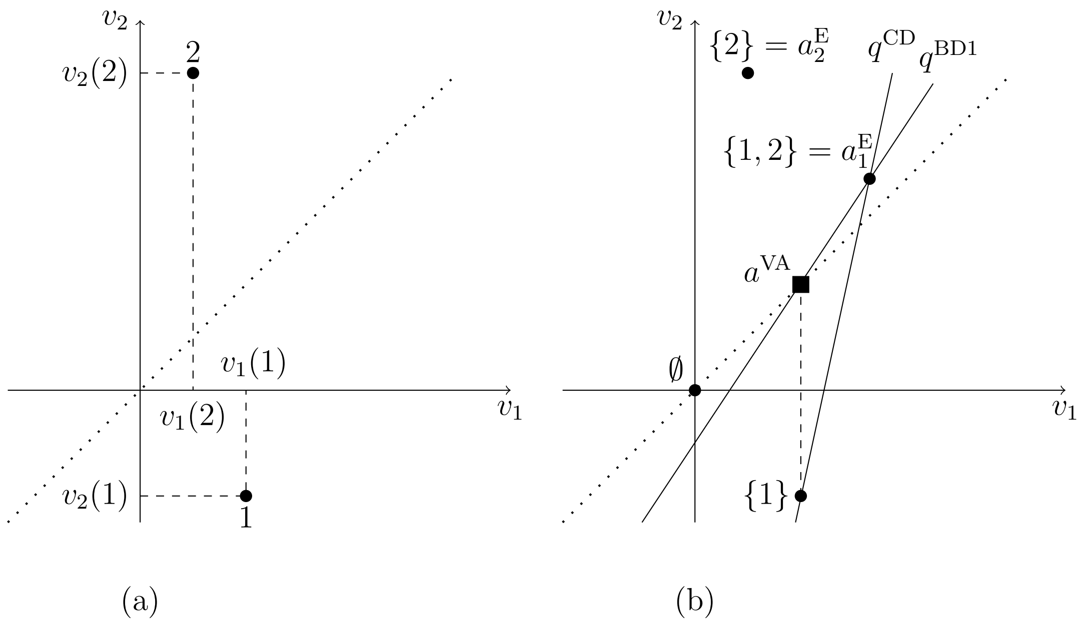



Figure 21: (a) The product space. (b) The alternative space with thresholds $q^{\mathrm{BD1}}$ and $q^{\mathrm{CD}}$.

with product $i$ with some probability.

# J A Bundling Example when Randomization is not Allowed

We here return to the example from Section 5.2 and describe the optimal menus when randomization is not allowed. For convenience, Figure 21 recreates the figures in the product space and the alternative space from Section 5.2.

We identify the optimal menus using Theorem 1. For this, we need to identify the threshold $q^{\mathrm{BD1}}$, which is the lowest $q$ at which the virtual alternative $a^{\mathrm{VA}}$ is maximal in the direction of the vector $(1, -q)$ among the alternatives $A^+ \cup \{a^{\mathrm{VA}}\}$. In our example, $A^+$ consists of alternatives $\{2\}$ and $\{1,2\}$, and the threshold $q^{\mathrm{BD1}}$ is such that alternative $\{1,2\}$ and the virtual alternative $a^{\mathrm{VA}}$ lie on a hyperplane orthogonal to the vector $(1, -q^{\mathrm{BD1}})$, so

$$v_1(\{1,2\}) - q^{\mathrm{BD1}} v_2(\{1,2\}) = v_1(a^{\mathrm{VA}}) - q^{\mathrm{BD1}} v_2(a^{\mathrm{VA}}).$$

This is shown in Figure 21, (b). For any $q < q^{\mathrm{BD1}}$, alternative $\{1,2\}$ is maximal in the direction of the vector $(1,-q)$ among the alternatives $A^+ \cup \{a^{\mathrm{VA}}\}$, and for any $q > q^{\mathrm{BD1}}$, $a^{\mathrm{VA}}$ is maximal in the direction of that vector. Recall that $q^{\mathrm{CD}}$ is the threshold at which $\{1,2\}$ and $a^{\mathrm{BD1}}$ are maximal (among all alternatives) in the direction of the vector $(1,-q^{\mathrm{CD}})$, and notice that $q^{\mathrm{BD1}} > q^{\mathrm{CD}}$, as shown in Figure 21, (b).

We now describe type 1's alternative in any optimal (non-random) menu. If $q < q^{\mathrm{BD1}}$, then type 1's alternative is the grand bundle $\{1,2\}$. If $q = q^{\mathrm{BD1}}$, then type 1's alternative can be either $\{1,2\}$ or $a^{\mathrm{BD1}} = \{1\}$. If $q > q^{\mathrm{BD1}}$, then type 1's alternative is $a^{\mathrm{BD1}} = \{1\}$.

There is a connection between the optimal menus with and without randomization. For $q > q^{\mathrm{BD1}} > q^{\mathrm{CD}}$, with randomization type 1's alternative is a distribution over $\{1,2\}$ and $\{1\}$; without randomization type 1's alternative is $a^{\mathrm{BD1}} = \{1\}$. One might wonder if this is a general property: is it always the case that $a^{\mathrm{BD1}}$, which is type 1's alternative in the optimal deterministic menu for any $q > q^{\mathrm{BD1}}$, is also in the support of type 1's random alternative in an optimal random menu for $q > q^{\mathrm{CD}}$? This property is not in fact general. In the example, $a^{\mathrm{BD1}} \in B_1(q^{\mathrm{CD}})$ and, moreover, it is the only alternative in $B_1(q^{\mathrm{CD}})$ that lies below the diagonal in the alternative space, so it has to be used in type 1's random alternative for any $q > q^{\mathrm{CD}}$. We now give an example in which $a^{\mathrm{BD1}} \notin B_1(q^{\mathrm{CD}})$, so $a^{\mathrm{BD1}}$ is not in the support of type 1's random alternative in the optimal random menu for any $q > q^{\mathrm{CD}}$, but $a^{\mathrm{BD1}}$ is type 1's alternative in the optimal deterministic menu for any $q > q^{\mathrm{BD1}}$.

For the example, we add a third product, product 3, without changing the other two products. The result in the product space is shown in Figure 22, (a). To depict the result in the alternative space, we take each of the four existing alternatives and move it in the direction of the vector $(v_1(3), v_2(3))$ to obtain four new alternatives, each containing product 3, for a total of eight alternatives. Both types have a negative valuation for product 3, so the efficient alternatives remain unchanged. For certain parameters, the threshold $q^{\mathrm{CD}}$ and the set $B_1(q^{\mathrm{CD}}) = \{\{1\},\{1,2\}\}$ remain unchanged.[35] Importantly, $B_1(q^{\mathrm{CD}})$ does not contain the bundle $a^{\mathrm{BD1}} = \{1,2,3\}$. Thus, for $q > q^{\mathrm{BD1}}$ type 1's alternative in the optimal deterministic menu is the bun-

[35] Specific valuations that work are $v_1(1) = 2$, $v_1(2) = 1$, $v_1(3) = -0.5$, $v_2(1) = -2$, $v_2(2) = 6$, and $v_2(3) = -2$.

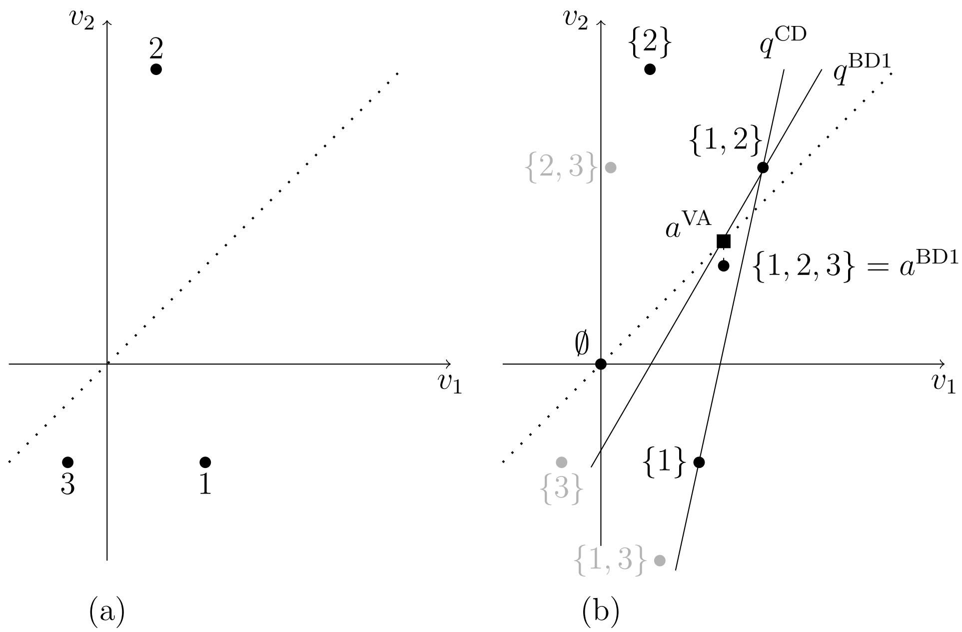



Figure 22: (a) The product space with three products. (b) The corresponding alternative space with eight alternatives. Alternative that are not important in the example are lightly shaded to ease visualization.

dle $\{1, 2, 3\}$, but in the optimal random menu $\{1, 2, 3\}$ is not in the support of type 1's random alternative, which randomizes over bundles $\{1\}$ and $\{1, 2\}$.