Department of Economics
Johannes Kepler University of Linz

Value Design in Optimal Mechanisms

by

Anja PRUMMER
Francesco NAVA

Working Paper No. 2305
Juli 2023

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

anja.prummer@jku.at

Value Design in Optimal Mechanisms*

Anja Prummer[†] Francesco Nava[‡]

July 6, 2023

Abstract

A principal allocates a single good to one of several agents whose values are privately and independently distributed, employing an optimal mechanism. The principal shapes the distribution of the agents' values within general classes of constraints. Divisive product designs, which are either highly favored or met with indifference, can simultaneously enhance surplus and diminish information rents by making agents' values more readily discernible. However, such designs also reduce competition among agents. Divisive designs are optimal under various design constraints, as the main drivers of revenue lie in increasing surplus and minimizing information rents, while competition plays a secondary role.

Keywords: Value Design, Mechanism Design, Differentiation

JEL Codes: D82, D46, L15

1 Introduction

Mechanisms are utilized across numerous markets for the allocation of goods, with collectibles, like art, antiques, or vintage cars frequently auctioned. Service contracts are commonly procured through government auctions. In these contexts, the

*We thank Chris Bidner, Daniele Condorelli, Dan Kovenock, David McAdams, Balazs Szentes for excellent comments. We are grateful to participants at various seminars and conferences for their insightful questions and comments. All remaining errors are ours. First draft: February 2020.

[†]Johannes-Kepler-University Linz, anja.prummer@jku.at

[‡]London School of Economics, f.nava@lse.ac.uk

principal not only determines the allocation mechanism but also has the ability to select or design the goods being auctioned, thereby influencing the agents' values for those goods. Traditional contributions in canonical mechanism design have primarily focused on identifying optimal methods for allocating goods (Myerson (1981), Bulow and Roberts (1989)), while neglecting the question of which goods should be offered for sale.

We grant the principal the power to both implement the optimal mechanism and influence valuations through product design. The principal necessarily faces some limitations in his product design and thus is restricted in the value distributions he can induce. We consider numerous different classes of constraints and provide a comprehensive characterization of jointly optimal value and mechanism design. Our study reveals that the principal consistently finds it advantageous to create dispersion in values within and across agents. Value dispersion is profitable as it can simultaneously increase surplus (the value of the agent with the highest willingness to pay) and enhance value predictability, which reduces the information rent. Consequently, divisive goods emerge as the optimal choice under various general constraints on the principal's ability to design values.

The design constraints faced by a principal depend on the specific context at hand. For instance, when a painter sells their artwork through a mechanism, they have the ability to shape the product's features. They can either create a piece of art that appeals to a broad range of clients or focus on a more niche customer base, crafting a divisive product that perfectly caters to their specific desires. Similarly, in the realm of collectibles, an auctioneer can choose to showcase peculiar pieces or those that have a wider appeal. Another feature of the good to be designed may be its location. If the designer selects where to sell a product, the distance from the venue selected will affect customer or bidders values for the product. In the context of service contracts, the location of where the work is carried out influences a potential contractor's value of the project. However, choosing a location closer to one agent necessarily implies greater distance to the other.

In all these scenarios, there exists an inherent limit to how desirable a good can be made by the principal. Moreover, a natural trade-off arises between creating a good that strongly appeals to a few individuals but is less appealing to others, as opposed to a good that moderately satisfies a larger number of agents. Taking these factors into consideration, we delve into the discussion of optimal value design under a variety of broad and abstract constraints that capture the limitations and trade-offs in valuations.

Similar to the approach taken in Myerson (1981), the principal uses an optimal mechanism to allocate the designed product, encouraging agents to truthfully disclose their private values for the good. The mechanism determines both the transfers made from the agents to the principal and the probability of allocating the good to each agent based on the induced profile of valuations. Through product design, the principal shapes the distribution of agents' values, which has three effects: (i) it directly affects agents' valuations and contributes to the overall surplus, (ii) it influences the allocation probability, thereby impacting the transfers, and (iii) it alters the information rents provided by the principal, incentivizing agents to truthfully reveal their induced valuations and ensuring incentive compatibility of the mechanism. These factors collectively determine the optimal value design for the principal, within the bounds of constraints.

These constraints can be categorized as separable and joint constraints. Separable design constraints pertain to limitations on the distribution of a single agent, while joint design constraints revolve around restrictions on the joint distribution of agents' valuations. Separable constraints capture scenarios in which the principal can affect an agent's willingness to pay for the product without affecting the value of others. On the other hand, joint constraints allow the principal to redistribute valuations among agents (increasing the value for some at the expense of reducing the value for others), and they are useful in modeling situations where the principal cannot affect an agent's desire for the product without affecting the value of others.[1]

Separable Design Constraints As a benchmark, we examine a scenario in which the principal possesses the capability to generate any distribution for the agents. In this

[1] It is worth noting that if the principal can increase valuations for all agents, they will always do so. Conversely, they will never reduce values for all agents. Therefore, we focus on the more nuanced case where an increase in valuation for one agent is accompanied by a reduction in value for another agent. This aligns with our motivating examples.

setting, minimizing the information rents would involve rendering the agents' values deterministic, thereby making them fully predictable. This would enable complete surplus extraction and eliminate information rents. Since higher values correspond to higher transfers, the principal also aims to maximize the value of the highest-ranked agent as much as possible, thereby increasing the surplus. However, certain values may be impractical or unattainable, which necessitates imposing an upper bound on the product's value, denoted as $\overline{v}$. In such cases, it is optimal to design a product that is valued at $\overline{v}$ by at least one agent with certainty. The principal can then set $\overline{v}$ as the transfer required to obtain the good, thereby extracting the entire surplus. Given that the principal cannot achieve a better outcome than obtaining the full surplus, having a single agent with a high value suffices.

Next, we consider scenarios in which the principal faces more stringent design constraints, allowing the generation of any value distribution with an expected value below a constant $k < \overline{v}$. This constraint aligns with the concept of "Bayes' plausibility" proposed by Kamenica and Gentzkow (2011). Under such constraints, the optimal value design entails binary value distributions for all agents, where mass is concentrated solely at the highest possible value, $\overline{v}$, and is zero elsewhere, with a mean equal to k. The principal benefits from increasing the expected value, and therefore designs products that, on average, all agents value at k, the upper bound. To extract the entire surplus, the principal creates a binary distribution with positive mass assigned to exactly one positive value, reducing the information rent to zero. Furthermore, the designed value distributions influence the probability of allocating the good to competing agents. Since an agent wins the object only by paying the exact high value, the principal aims to decrease the likelihood of this event to engage in more frequent trading with competing agents. By doing so, the competitors' expected transfers to the principal increase, despite the reduction in competition when the designed agent has a low value. Consequently, the principal optimally sets the high value at the maximum possible value, $\overline{v}$, thereby minimizing the probability of an agent having a high value.

Contrary to the widespread belief that higher variance increases agents' information rents, we demonstrate that this is not necessarily the case. Variance among non-excluded values diminishes revenue, as it limits the ability to extract values due to agents' private information. However, variance in the overall distribution of values, across both excluded and non-excluded values can actually benefit the principal. This occurs when the overall increase in variance enhances the predictability of values with positive allocation probability, allowing the principal to learn and extract the agent's value during trade.

The optimal value design when the constraint imposes a cap on the mean, given by a binary spread with mass at $\overline{v}$ and zero, that averages to k, is second-order stochastically dominated by any distribution with a mean of k. Therefore, the optimal design under the bounded mean remains optimal even under design constraints that restrict the principal to generating distributions that are second-order stochastically dominated by an arbitrary distribution with a mean of k. Opting for a riskier distribution consistently benefits the principal. However, if the principal is limited to creating distributions that are first-order stochastically dominated by a specified distribution, they refrain from making any adjustments. In such cases, reducing the information rent of the agent by increasing the predictability of their value comes at the cost of diminishing values and surplus, with the surplus motive prevailing. Thus, the principal selects the bounding distribution as the optimal value design for all agents, since he cannot benefit from damaging the product.

Joint Design Constraints When design constraints are not separable, the principal is tasked with jointly generating value distributions for all agents and may have the ability to reallocate value across them. Such constraints are natural and can arise, for example, when the principal decides where to sell the good, benefiting agents who are closer to the venue while potentially disadvantaging others. We begin by assuming that the sum of the expected values of all agents cannot exceed a certain constant. This constraint resembles the first constraint analyzed in the separable design case, making it a sensible starting point. In this setting, the principal assigns all the value to a single agent, effectively reducing the values of the other agents to zero. The intuition behind this result aligns with the benchmark case without any constraints: having one agent with a known valuation is sufficient to extract the entire surplus. This once again leads to the creation of a divisive product that is highly favored by a few and largely disliked by the majority. However, in contrast to the earlier case, the principal now distributes values across agents rather than within agents.

The outcome of this result is striking and prompts further inquiry into its general applicability. To investigate this, we consider a scenario with two agents (without loss of generality) and define an arbitrary measure with a mass of two, represented by the cumulative function $H(v)$, where v denotes the induced value for the good. The principal can design any two value distributions for the two agents, provided that their cumulative distribution adds up to $H(v)$ pointwise. Once again, the optimal strategy involves creating divisive goods with maximally distinct value distributions. Specifically, one agent's values are confined solely below the median of $H(v)$, while the other agent's values lie exclusively above the median of $H(v)$. By designing such maximally distinct distributions, the principal not only increases the total surplus but also extracts higher transfers by reducing the agents' information rents. Maximal differentiation renders the values of both agents more predictable compared to any other design, resulting in lower information rents for the agents. While this designed inequality diminishes the transfer received from the disadvantaged agent, the gain in transfer from the advantaged agent more than compensates for it, making maximally divisive designs the optimal choice.

Our findings provide valuable insights into why principals should introduce dispersion either within or across agents whenever feasible. Specifically, principals should strategically design niche products that not only maximize their ability to learn about the underlying values of agents but also maintain a high level of total surplus.

As a result, it is optimal for principals to create goods that are divisive in nature, appealing to a select few while being met with indifference by a larger audience. This phenomenon can be observed in various contexts, such as the art market, where individual pieces often attract significant attention from a small number of collectors but fail to resonate with the wider public.

Related Literature Our contribution adds to the existing literature on value, or product, design in markets and mechanisms. The initial key insights in this area were developed by Johnson and Myatt (2006), focusing on environments where a price-setting monopolist can influence the market demand for the good it sells through the choice of product features or advertising. Their analysis centers on reshaping demand through rotations and shifts. They demonstrate that both value distributions with minimal dispersion (mass products) and distributions with maximal demand dispersion (niche products) can be optimal. Similarly, our approach allows the seller to determine both the features of the product sold and the mechanism to allocate this. However, we consider settings with multiple buyers, thereby changing the mechanism used for selling (from a posted price to an optimal independent private value mechanism), and we consider several different classes of design constraints which can encompass demand rotations. Our findings diverge as we establish that dispersion in values within and across buyers generally benefits the designer.

Our results, demonstrating the optimality of maximally divisive value designs, rely on the assumption that the seller has the ability to simultaneously design the mechanism and the value distributions. Moreover, we allow for surplus to vary. Contrary to our finding, Cantillon (2008) shows that when surplus is fixed, revenue is reduced by asymmetric value designs in first and second price auctions without exclusion.[2]

We further complement Condorelli and Szentes (2020, 2022). Their earlier work focuses on buyer-optimal value design with a single buyer, while the subsequent paper can be leveraged to characterise the optimal profit arising from differ demand functions in Cournot competition.[3] Adding to this strand, we consider seller-optimal designs with multiple buyers while focusing on optimal designs.

Methodologically, some of the insights derived for separable design problems with mean bounds depend on the convexity of surplus in n-buyer allocation problems, and

[2]In contrast to our focus on designing value distributions, Deb and Pai (2017) fix value designs and establish that symmetric mechanism can heavily discriminate between agents when their value designs differ.

[3]The analysis in Condorelli and Szentes (2022) is closely related to the pioneering work of Bergemann, Brooks, and Morris (2015) who first identified all combinations of consumer surplus and producer surplus that can arise in a monopoly setting, with posted prices, if the seller designs the information that buyers receive about their value for the product.

are conceptually linked to majorization results in Kleiner, Moldovanu, and Strack (2021). Results concerning separable design problems with pointwise bounds closely resemble a result in Hart and Reny (2015) establishing that a seller never benefits from selecting first order stochastically dominated designs in single buyer settings. We generalize this result to settings with multiple buyers while having to devise a new proof strategy, as optimal allocation rules become more intricate when multiple buyers compete for the object.

Finally, our work can be viewed through the lens of information design.[4] The key difference lies in the broader classes of design constraints that we can entertain, which can encompass Bayes plausibility but are more general. Instead of merely providing information about the features of the good, as in Bergemann and Pesendorfer (2007), Eso and Szentes (2007), Sorokin and Winter (2018) , and Ganuza and Penalva (2019), the principal can design product features in our setting. Despite these differences, Ganuza and Penalva (2019) obtain an insight that complements some of ours. They show that the principal discloses more information as the number of bidders grows, when information disclosure is public. Our results on separable designs instead imply that when the product sold is optimally designed, the principal discloses the private value of the product fully to all but possibly one bidder (for any possible number of bidders), while the remaining bidder may or may not be informed of the private value (learning only their mean value for the product). Therefore, their insights align with our finding that maximally spread values are optimal, connecting the two approaches of value versus information design.

Section 2 presents our model and the key design constraints. Section 3 solves the optimal value design problems. It first considers a benchmark design problem where the principal does not face constraints in Section 3.1. Then, we analyse separable design problems in which the principal faces a design constraint for each agent in Section 3.2. We turn to joint design constraints where value can be redistributed across agents in Section 3.3. Section 4 concludes.

[4]For instance, Roesler and Szentes (2017) and Bobkova (2019) examine trade environments with a single buyer and a single seller, and characterize the buyer-optimal information designs and the costs associated with information acquisition.

2 A Model of Constrained Value Design

We begin by revisiting classical independent private value mechanism design insights that will be used in our analysis. Then, we present the value design problems at the heart of the paper. Results and the model are presented for two agents for sake of clarity, but extensions to n-agents settings are discussed throughout the paper.

2.1 The Environment and the Optimal Mechanism

Two agents, A and B, compete for a good. The value of the good for agent $i \in \{A, B\}$ is denoted by v_i. Agents are privately informed of their value, and values are independently distributed. The cumulative distribution of value v_i is denoted F_i and its support is denoted by $V_i \subset \mathbb{R}_0^+$. Let $\underline{v}_i$ and $\overline{v}_i$ respectively identify the smallest and the largest value in the support V_i. The distributions of values F_A and F_B are commonly known by the two agents and by the principal. The preferences of each agent are separable in value and transfers, as standard in optimal mechanisms (Myerson (1981)). They simply amount to the value for the good minus the transfers made to the principal if the agent wins the good and to minus the transfer made to the principal otherwise.

The principal maximizes revenue by allocating the good with an optimal mechanism.[5] By the revelation principle, it is without loss of generality to focus on direct, incentive compatible (IC) and individually rational (IR) mechanisms to allocate the object.[6] A direct mechanism specifies an allocation rule x as well as a transfer t,

$$(x, t) : V_A \times V_B \to \Delta_-^2 \times \mathbb{R}^2. \tag{1}$$

The allocation rule $x(\mathbf{v}) = (x_A(\mathbf{v}), x_B(\mathbf{v}))$ identifies the probability with which agents A and B win the good for any profile of reported values $\mathbf{v} \in V_A \times V_B$. These probabilities are contained in Δ_-^2, the set of allocation probabilities. The transfer

[5]We set the cost of sourcing the good to zero, which implies that revenue equals profit. Our analysis would be unchanged if the principal faced a fixed cost of sourcing the good, as this would merely lower profits.

[6]The revelation principle states that for any mechanism and any Bayes Nash equilibrium of that mechanism, there exists a Bayes Nash equilibrium in a corresponding direct mechanism with the same outcomes and in which all agents participate and reveal their value truthfully.

rule $t(\mathbf{v}) = (t_A(\mathbf{v}), t_B(\mathbf{v}))$ identifies the transfer of each agent to the principal for any profile of reported values $\mathbf{v} \in V_A \times V_B$. The allocation probabilities and transfers are such that the mechanism is IR and IC, meaning that there exists a Bayes Nash equilibrium in which all agents participate and report their type truthfully.[7]

The characterization of the optimal mechanism further relies on the revenue equivalence theorem which states that revenue in an incentive compatible mechanism, in which the lowest type of each agent is indifferent between participating and not, amounts to expected virtual surplus,

$$\mathbb{E}_F\left[\sum_i t_i(\mathbf{v})\right] = \mathbb{E}_F\left[\sum_i \psi_i(v_i) x_i(\mathbf{v})\right], \tag{2}$$

where $\psi_i(v_i)$ denotes the virtual value of agent i which depends on the distribution of the value, F_i, and where the expectation $\mathbb{E}_F$ is taken over the joint distribution of values. A direct mechanism is therefore optimal if it maximizes expected virtual surplus subject to (i) the interim allocation rule $x_i(v_i)$ being non-decreasing (which ensures incentive compatibility) and (ii) the lowest type of each agent being indifferent between participating or not (which combined with incentive compatibility implies individual rationality).[8]

Value distributions shape expected virtual surplus by affecting both the joint distribution of values F, and the virtual values, $\psi_i(v_i)$. For any player i and any value v_i at which the cumulative F_i is differentiable, the virtual value is defined as

$$\psi_i(v_i) = v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}, \tag{3}$$

where $f_i(v_i) = F_i'(v_i)$ denotes the density. We refer to the difference between the value and the virtual value as the *information rent.*

[7] For completeness, we formalise an agent's decision. With a slight abuse of notation (as we use the same operator to denote the ex-post rules and the interim rules), denote the interim (or expected) allocation probability for agent i with value $v_i \in V_i$ by $x_i(v_i) = \mathbb{E}_{F_{-i}}[x_i(\mathbf{v})] = \int_{\underline{v}_{-i}}^{\overline{v}_{-i}} x_i(\mathbf{v}) dF_{-i}(v_{-i})$. Similarly for the transfer rule denote the interim (or expected) transfer by $t_i(v_i) = \mathbb{E}_{F_{-i}}[t_i(\mathbf{v})] = \int_{\underline{v}_{-i}}^{\overline{v}_{-i}} t_i(\mathbf{v}) dF_{-i}(v_{-i})$. When cumulative distributions are discontinuous on the support, Riemann-Stijltes integrals will be used to calculate the expectations implicitly. IR and IC for agent i with value v_i then together require that $v_i x_i(v_i) - t_i(v_i) \geq \max\{v_i x_i(z_i) - t_i(z_i), 0\}$ for any $z_i \in V_i$.

[8] The agent with the lowest type is indifferent for a transfer $t(\underline{v}_i) = \underline{v}_i x_i(\underline{v}_i)$.

To establish our results when cumulative distributions are not continuously differentiable on their support, we translate the problem to the *quantile space* as needed, following an approach pioneered in Bulow and Roberts (1989).[9] For any distribution F_i, define the quantile associated with value $v_i \in V_i$ as $q_i(v_i) = 1 - F_i(v_i)$. Similarly, define the value $v_i(q)$ associated with any quantile $q \in [0, 1]$ as $v_i(q) = \inf \{v_i \in V_i | q \geq 1 - F_i(v_i)\}$. This definition encompasses cases in which the cumulative distribution is discontinuous or contains atoms.[10] At quantiles q at which $v_i(q)$ is differentiable, the virtual value can be defined in the quantile space as

$$\phi_i(q) = \psi_i(v_i(q)) = v_i(q) + v_i'(q)q = \frac{\partial (v_i(q)q)}{\partial q}. \tag{4}$$

Similarly, the interim allocation probability for agent i in the quantile space can be defined as $y_i(q) = x_i(v_i(q))$ and must be non-increasing by incentive compatibility.

2.2 The Value Design Problem

Having discussed key properties of the optimal mechanism used in our results, we now turn to the value design problem at the heart of the analysis. To keep the problem compact, we posit throughout that the principal can only design distribution of values with support contained in $[\underline{v}, \overline{v}]$ for some $\underline{v} \geq 0$ and some $\overline{v} < \infty$ – where $\underline{v}$ and $\overline{v}$ represent respectively the lowest and the highest possible values for the good. We refer to any pair of distributions (F_A, F_B) with supports contained in $[\underline{v}, \overline{v}]$ as a *value design*. The principal's value design problem in the quantile space amounts to

$$\max_{F_A, F_B} \quad R(F_A, F_B) = \mathbb{E}[\phi_A(q)y_A(q) + \phi_B(q)y_B(q)] \tag{5}$$

$$\text{s.t. Distributional Constraints,}$$

where $\mathbb{E}$ denotes the expectation of the quantile relative to the uniform distribution. We consider several abstract and general design restrictions. These constraints capture limitations that the principal may face when sourcing and producing products,

[9] See Hartline (2013) for a recent comprehensive survey.

[10] When a cumulative distribution is continuous and strictly increasing, $v_i(q) = F_i^{-1}(1 - q)$.

and can reflect costs.

In all the design problems considered, we posit that the principal cannot increase the value for the good infinitely and that all agents' values remain bounded– a natural assumption in the context of product design. However, the limitations in product design depend on the specific good considered. To capture a wide range of potential limitations, we consider both design constraints disciplining only the expectation of agents' value distributions, and design constraints disciplining the value designs pointwise. The former gives the principal substantial freedom in choosing value designs, the latter will be more demanding on the set of feasible designs restricting the likelihood of any given value. Design restrictions can further be parsed in two classes: separable design constraints and joint design constraints. The former class considers constraints affecting the distribution of values for each agent in isolation. The latter class considers restrictions that jointly constrain the design of the value distributions for all agents, thereby allowing for value reallocation across agents. Separable constraints may be more appropriate in settings where product design affects consumers in the same manner, while joint designs imply trade offs across consumers.[11]

Our analysis of separable designs focuses on two key constraints. The first of these constraints allows the principal to choose any value design (F_A, F_B) such that the expected value of the distribution lies below some constant $k \in [\underline{v}, \overline{v}]$. Formally, the constraint allows the principal to select any value design satisfying for all $i \in A, B$

$$\mathbb{E}_{F_i}[v] \leq k, \tag{6}$$

where $\mathbb{E}_{F_i}[v]$ denotes the expected value of the distribution F_i. In an information design interpretation, this constraint allows for all all value designs satisfying Bayes Plausibility given some prior with mean k. But the designer can do more in our setting, as they may destroy value or alter the support of the true value distribution. Results will also establish why constraint (6) encompasses the set of distributions

[11]Of course, value designs increasing values for all agents would be desirable for the principal. Our analysis will focus on scenarios in which all such designs have been implemented and the remaining options lower value for one agent whenever the value for the other agent increases.

that are second order stochastically dominated by some bounding distribution G with $\mathbb{E}_G[v] = k$.

The second natural constraint considered in this class is first order stochastic dominance which allows the principal to design any distribution of values F_i for agent i satisfying

$$F_i(v) \geq G(v) \text{ for all } v \in [\underline{v}, \overline{v}], \tag{7}$$

where the cumulative distribution G provides a pointwise upper-bound on the distributions that the principal can choose to design.

Our analysis of joint designs focuses on two additional constraints. These capture settings in which values can be reallocated across agents, meaning that enhancing the distribution of values of one agent comes at the cost of deteriorating the distribution of values for the other agent.

Mirroring constraint (6) in the separable design problem, we first allow the principal to generate any two distributions for which the sum of expected values across distributions is below some constant $k \in [\underline{v}, \overline{v}]$. Formally, any two distributions of values must satisfy

$$\mathbb{E}_{F_A}[v] + \mathbb{E}_{F_B}[v] \leq k. \tag{8}$$

To account for distributional constraints, we then take an arbitrary measure with mass two and an associated cumulative distribution $H(v)$, and ask how the principal would split such a measure to create the two value distributions for the two agents.[12] In essence, we allow the principal to design any pair of value distributions satisfying

$$F_A(v) + F_B(v) = H(v) \text{ for all } v \in [\underline{v}, \overline{v}]. \tag{9}$$

This design constraint would endogenously arise in a costly design setting with design costs fulfilling a weak form of linearity, as explained in Appendix B.

[12]Formally, let $\mathcal{H}$ denote the measure and define $H(v) = \mathcal{H}([\underline{v}, v])$.

2.3 Discussion of Modelling Assumptions

IPV Design In our modelling interpretation, the principal can directly affect agents' value distributions while realised values remain independent and private. Such assumptions appear palatable for several applications in which the principal can design goods while carrying out research to measure the appetite for the good conditional on observable characteristic of the agents, but cannot fully learning the value for the good of agents. Maintaining the independence of value distributions across design problems is a natural assumption if one believes that the underlying preferences of agents are an independent trait that cannot be affected. One natural interpretation is that the designer faces agents whose types are uniformly distributed between zero and one and he can lay over this innate type a value distribution.[13] Independence further leads to a more challenging problem, as it requires accounting for agents' information rents rather than focusing purely on surplus design. In Appendix A, we consider extensions to setting in which the agents' values are not independent, and show that our results carry over to these environments.

Costs of Value Design To provide a clear characterisation of the forces at play in the value design problem, the analysis abstracts from cost considerations. Instead, we characterize the optimal distributions that maximize revenue in the optimal mechanism subject to design constraints that capture the principal's ability to disperse, destroy, or reallocate value across agents. The principal will freely choose among all possible distributions of values, F_A and F_B, subject to natural restrictions. These constraints are related to assumptions on design costs as will be discussed in Appendix B.

3 Optimal Value Design

To provide a benchmark, we begin by characterising the optimal distribution when no constraints are in place, and then analyse the separable and joint design problems introduced in the previous section. All proofs of propositions and corollaries can be

[13] In particular, the quantile space approach lends itself to such an interpretation.

found in Appendix D. The proofs of remarks are omitted as they follow immediately from other results.

3.1 Designing Value without Constraints

Suppose the principal can design any value distributions with support in $[\underline{v}, \overline{v}]$. Then the principal's optimal value designs are the ones in which at least one agent values the good at $\overline{v}$ with certainty. Such value designs are optimal as the surplus cannot exceed the maximal value for the good $\overline{v}$. If an agent has this value with probability one, the principal can simply allocate the good to that agent with certainty while asking them to transfer $\overline{v}$, thereby extracting the maximal surplus.

Remark 1. *In any optimal value design, there is at least one agent with a value distribution satisfying $v(q) = \overline{v}$ for all quantiles $q \in [0, 1]$.*

The result highlights two forces at play which re-emerge throughout our further analysis. First, the principal would like to increase total surplus (or equivalently the maximal value for the good) as much as possible, because a higher value induces a higher transfer. Second, the principal would like for the value of the agents to be as predictable as possible, because knowing precisely the agent's value reduces the information rent that the principal pays to the agent to ensure incentive compatibility. To see this, note that in the quantile space, the information rent satisfies $v_i(q) - \phi_i(q) = v_i'(q)q$, and thus equals zero when the value does not change in the quantile space. Therefore, creating a more narrow distribution and making the value easily recognisable, increases the revenue of the principal.

It suffices to increase the value for one agent to $\overline{v}$, as the principal, knowing the value of that agent with certainty, is able extract the full surplus from them without ever selling to the other agent. If the principal adjusts the distribution of both agents, then one agent obtains the good with probability $p \in [0, 1]$, while the other receives it with probability $1 - p$. This reduces the transfer of the agent who previously obtained the good with certainty, while simultaneously increasing the transfer of the agent who never received the good, by the same amount. Thus, the revenue is the same, independently of whether the principal increases the value for one or more agents.

3.2 Separable Design Constraints

We turn our attention to optimal value distributions for the separable design problems set out earlier. We begin by considering constraint (6), which bounds the mean of the possible distributions by a constant $k \in [\underline{v}, \overline{v}]$.

To state results it is useful to define a class of two-atom distributions with atoms at values $h \in [k, \overline{v}]$ and $\underline{v}$, and mean k. Formally, for any $h \in [k, \overline{v}]$, let the probability distribution P^h satisfy

$$P^h(v) = \begin{cases} \frac{h-k}{h-\underline{v}} & \text{if} \quad v = \underline{v} \\ \frac{k-\underline{v}}{h-\underline{v}} & \text{if} \quad v = h, \end{cases} \tag{10}$$

where $P^h(v)$ denotes the probability that the value equals v. Let F^h denote the cumulative distribution associated to P^h. With this definition note that $F^{\overline{v}}$ coincides with the maximally spread two-atom distribution, while F^k amounts to a distribution with a single atom at the mean k.

The first key result for distributions fulfilling constraint (6) establishes that in any optimal design, the principal designs a good that is divisive for at least one agent. For that agent, the expected value of the good amounts to the mean bound k, but the realized value of the good is either maximal or minimal. For the other agent, several two-atom designs can be optimal when the minimal value equals 0, but the unique optimal design for positive minimal values is one in which the value equals the mean bound k with certainty.

Proposition 1. *A value design, (F_A^*, F_B^*), is optimal among all designs satisfying $\mathbb{E}_{F_i}[v] \leq k$ for all $i \in \{A, B\}$, if and only if*
(1) $F_i^* = F^{\overline{v}}$ for some agent $i \in \{A, B\}$,
(2) $F_j^* = F^h$ for agent $j \neq i$ with $h = k$ if $\underline{v} > 0$ and $h \in [k, \overline{v}]$ if $\underline{v} = 0$.

Before discussing the applied relevance and content of the result, we provide some intuition for the proof of the result. The proof begins by characterizing value designs maximizing total surplus (or gains from trade).[14] It does so by fixing the distribution of values for one agent j and a realized value for the other agent i, and by establishing that the expected total surplus is convex in the realized value of agent i. The latter observation coupled with the linearity of constraint (6) is then exploited to show that a surplus maximizing design for player i must be the maximally spread two-atom distribution $F^{\overline{v}}$. Further when $F_i = F^{\overline{v}}$, total surplus is only affected by the mean of j's value distribution, because when $v_i = \overline{v}$, surplus amounts to the value of player i as $v_i \geq v_j$ for any realized value for player j, and when $v_i = \underline{v}$, surplus amounts to the expected value of player j distribution as $v_j \geq v_i$ for any realized value for player j. The proof then concludes by showing that all the designs in the statement maximize surplus and leave no information rent to agents (argued below), and therefore maximize revenue, while all other surplus maximizing designs would have to pay an information rent to agents to fulfill incentive compatibility. As a result, in all optimal designs, there is no conflict between reducing information rents and increasing surplus.

The optimal design is unique up to agents' permutations as long as minimal values are strictly positive, $\underline{v} > 0$. But when minimal values equal zero, $\underline{v} = 0$, and competitor i has a maximally spread design, $F_i^* = F^{\overline{v}}$, multiple designs can be optimal for the remaining agent j. All these designs assigns mass to exactly one strictly positive value h with $h \in [k, \overline{v}]$ and have expected value k. In these designs if i's valuation is $\overline{v}$, the principal awards the good to agent i at transfer $\overline{v}$ and no information rent is paid. While if i's valuation equals $\underline{v}$, the principal awards the good to agent j only if they agree to pay a transfer h. So, the principal extracts k in expectation from j conditional on not trading with i, and no information rent is paid – irrespective of the exact value of h.

The optimal mechanism used for any of such optimal designs is simple and requires first offering the good to agent i at to price $\overline{v}$ and if i rejects offering the good to agent j at price h. In these value designs, the principal designs a product that is divisive for one of the two agents because volatility in values increases surplus by adding variance. To highlight that volatility in valuations is a key feature of

[14] *Total surplus* can formally be defined as $E_F[\max\{v_A, v_B\}]$.

optimal designs when design constraints affect only the mean valuation, consider the generalization of the previous statement to settings with $n+1$ agents. Let $N \subset \mathbb{N}$ denote a set of $n+1$ agents. In any optimal design, n agents exhibit a maximally dispersed two atom distribution, while the remaining agent's distribution is as described in Part (2) of Proposition 1.

Remark 2. *A value design is optimal among designs satisfying $\mathbb{E}_{F_i}[v] \leq k$ for all $i \in N$, only if: $F_i^* = F^{\overline{v}}$ for all agents $i \neq j$ for some $j \in N$.*
The revenue loss from a sub-optimal value design in which all agents draw value from distribution $F^{\overline{v}}$ amounts to at most $P^{\overline{v}}(\underline{v})^{n+1}\underline{v}$, and converges to 0 as $n \to \infty$.

The previous insight establishes that in large markets a designer seeks to design or source goods that are divisive for all but possibly one agent. Further, it shows that the loss from making the good divisive for all agents is small, and non-existent when minimal values equal zero. Recognizing that values in this setting represent the willingness or ability to pay, divisive goods can be interpreted as goods that are highly valued by some types and that are undesirable or unaffordable to the rest. For instance, in luxury good markets, mark-ups are often extremely high meaning that goods have been designed for a handful of potential buyers with extremely high valuations while the remaining buyers are priced out of the market, and have thus a low willingness to pay. These insights are consistent with our value design analysis when $\overline{v}$ is large and consequently $P^{\overline{v}}(\overline{v})$ is small. Optimal designs naturally lead to a high variance in observed transfers to the principal and are exploitative to agents, leaving all of them without any surplus. However, these designs deliver the highest possible revenue, which is worth approximately $(n+1)k$ when $\overline{v}$ is large and $\underline{v} = 0$.

Before progressing to the analysis of pointwise design constraint, it is instructive to fix a distribution for agent A and focus on the optimal design for agent B given F_A. This can be relevant if the principal's product only influences the distribution of one of his customers, while the other customer remains unaffected. We argue that this is a useful benchmark as product design may not impact all customers equally. Therefore, we ask what happens if the other customer is not affected, before analysing the joint design problem where catering to one consumer alienates the other. Throughout this benchmark analysis, we assume that $\underline{v} = 0$. Yet having minimal values equal zero seems most natural in the context of a value design problem, because a principal should always be able to make the good less desirable at no cost.

Proposition 2. *Fixing F_A, value design $F_B^* = F^{\overline{v}}$ is optimal among all value designs satisfying $\mathbb{E}_{F_B}[v] \leq k$. Further, such a design is uniquely optimal when $F_A \neq F^{\overline{v}}$. But when $F_A = F^{\overline{v}}$, any design $F_B^* = F^h$ for some $h \in [k, \overline{v}]$ is also optimal.*

The result is closely related to the earlier multi-agent results, yet the proof strategy differs considerably, because information rents may have to be paid to the agent whose distribution is not being designed. It relies on the minimal value $\underline{v}$ being equal to zero, because when $\underline{v} > 0$ and F_A is dispersed, the optimal design may not be maximally spread, but rather maximally concentrated, as shown in Proposition 1 for $F_A = F^{\overline{v}}$. From a mechanism design perspective, the result is perhaps surprising since it demonstrates an optimal design can maximize variance (the next corollary makes this point explicit). But by maximizing the volatility of valuations, the principal minimizes the dispersion in valuations (and consequently the information rent) for non-excluded types of the agent who win the good with positive probability. Further, since the principal does not care about agents who never win the good, he awards them with the smallest possible value so as to not lose surplus.

To gain intuition for the result, note that the principal has once again two, possibly conflicting, motives: maximizing surplus and reducing information rents. In the proposed design problem, these motives are again aligned. The first motive is driven by the principal's desire to design a good which is highly valued in order to extract more surplus from the agents through their transfers. This aim is met by selecting the maximally spread distribution of values $F^{\overline{v}}$ for which constraint (6) binds, as shown in Proposition 1. The second motive is driven by the principal's desire to make the agent's valuation easily identifiable in order to minimize the information rent paid to agents for truthfully revealing their type. As the value of player B is deterministic whenever gains from trade are strictly positive, no information rent is left to the agent. As the principal does not benefit by leaving any value to types that never win the good, the lower atom must be at zero. This establishes that a bimodal distribution is optimal, with one atom at a strictly positive value and another zero. The strictly positive value is generically equal to $\overline{v}$, because a change in B's distribution affects the allocation probability for *both* agents. As we keep the expected value of B's distribution at k, having the positive atom at a higher value decreases the probability that the value is positive and not zero which in turn increases probability that A obtains the good and consequently A's transfer.

It is worth noting that Proposition 2 also establishes that $F^{\overline{v}}$ is the optimal design relative to any other distribution that is second order stochastically dominated by a distribution G with mean $\mathbb{E}_G[v] = k$. This follows by two insights. First, any distribution F_B that is second order stochastically dominated by G satisfies

$$\int_0^v G(t)dt \leq \int_0^v F_B(t)dt \quad \text{for all } v \in [\underline{v}, \overline{v}], \tag{11}$$

and therefore has a lower mean than G, fulfilling constraint (6). Second, $F^{\overline{v}}$ is second order stochastically dominated by G. Thus, $F^{\overline{v}}$ is also the optimal distribution among all F_B satisfying constraint in (11).

Corollary 2.1. *Fixing F_A, the value design $F_B^* = F^{\overline{v}}$ is optimal among all value designs satisfying (11). It is the unique optimal design if $F_A \neq F^{\overline{v}}$.*

Thus, the principal always prefers a riskier value distribution for agent B given that it is always possible to use the other agent to hedge such risk.

Proposition 2 also implies that in any optimal design the principal is unwilling to reduce the expected valuation for the good. This raises the question of whether the principal would ever be willing to adjust the value distribution for a agent if this necessarily came at the expense of a reduction in the expected value. The next remark exploits the proof of Proposition 2 to show that this is indeed the case, as long as the reduction in expected value is sufficiently small. Define the two-atom probability distribution with atoms at values h and $\underline{v}$ and mean $m \in [\underline{v}, h]$,

$$P_m^h(v) = \begin{cases} \frac{h-m}{h-\underline{v}} & \text{if } \; v = \underline{v} \\ \frac{m-\underline{v}}{h-\underline{v}} & \text{if } \; v = h \end{cases} \tag{12}$$

where $P_m^h(v)$ denotes the probability that the value equals v. Let F_m^h denote the cumulative distribution associated to P_m^h.

Remark 3. *Consider any value design* (F_A, F_B). *For any* $m \geq \mathbb{E}_{F_B}[\max\{\psi_B(v), 0\}]$, *value design* $(F_A, F_m^{\overline{v}})$ *increases revenue compared to the initial value design* (F_A, F_B).

Condition $m \geq \mathbb{E}_{F_B}[\max\{\psi_B(v), 0\}]$ guarantees that any reduction in expected value, $\mathbb{E}_{F_B}[v] - m$, is more than compensated by the reduction in information rent, $\mathbb{E}_{F_B}[v] - \mathbb{E}_{F_B}[\max\{\psi_B(v), 0\}]$. Thus, even if the principal loses some surplus on average by reducing the mean, doing so still increases the revenue due to the reduction in information rents.

Next, we turn to the other separable value design problem, namely the one requiring designed value distributions to be first order stochastically dominated by an arbitrary distribution G. As in the previous design problem, we begin by fixing the value distribution of agent A and characterising the optimal design for agent B subject to $G(v) \leq F_B(v)$ for all v. First order stochastic dominance implies a ranking of quantiles $q_G(v)$ and $q_B(v)$ – so that for all $v \in [\underline{v}, \overline{v}]$:

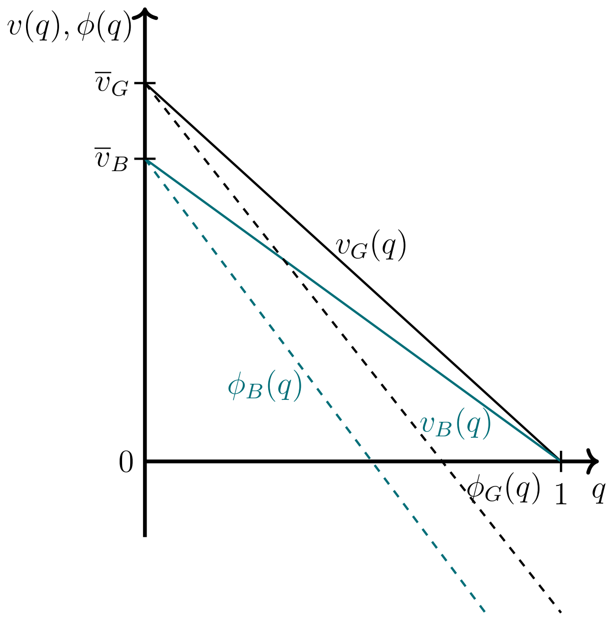
$$q_G(v) = 1 - G(v) \geq 1 - F_B(v) = q_B(v). \tag{13}$$

It follows that for any given quantile, the value under distribution G is higher than under any other distribution that is first order stochastically dominated by G, formally $v_G(q) \geq v_B(q)$ for any $q \in [0, 1]$. This observation is summarised in Figure 1, and is instrumental in proving the next optimality result.

Proposition 3. *The optimal value design,* (F_A^*, F_B^*), *among all designs satisfying* $F_i(v) \geq G(v)$ *for all* $v \in [\underline{v}, \overline{v}]$ *and all* $i \in \{A, B\}$ *is* $(F_A^*, F_B^*) = (G, G)$.

We begin by establishing that $F_B^* = G$ is optimal for all F_A, before turning to the design of both distributions. If the principal can only design distributions that are first order stochastically dominated by G, agent B is assigned the bounding distribution G in the unique optimal design. In this design problem, designs differing from G may be able reduce information rents by making the agent's value more predictable, but at the expense of reducing values and thus surplus. The result

Figure 1: G FOSD F_B



establishes that any gain in information rent cannot outweigh the loss in value. Proposition 2 in Hart and Reny (2015) establishes the same result for one-bidder settings. Our result generalizes theirs by showing that first order stochastically dominated distributions yield lower revenues even with competition and no matter what the competitors' value distributions are. Thus, damaging a product for all types of an agent cannot be optimal even in competitive setting.

To gain more intuition for the result, consider any design F_B that differs from G and is first order stochastically dominated by G. The proof establishes that if so, it is always possible to construct an alternative design, $\hat{F}_B$, which is first order stochastically dominated by G, but first order stochastically dominates F_B, and that is associated to a higher revenue. The change in distribution, from F_B to $\hat{F}_B$, has two effects, it affects the virtual value for the agent whose value is being designed and it affects the probability of winning the good for *both* agents. Our proof strategy fixes the allocation rule that was optimal for value design (F_A, F_B) and shows that revenue increases when B's distribution is transformed to $\hat{F}_B$ due to the adjustment of virtual values. Formally, the key step of the proof establishes the second inequality in the following expression

$$\underbrace{\mathbb{E}[\hat{\phi}_B(q)\hat{y}_B(q)]}_{\text{Distribution:}\hat{F}_B,\ \text{Allocation:}\hat{F}_B} \geq \underbrace{\mathbb{E}[\hat{\phi}_B(q)y_B(q)]}_{\text{Distribution:}\hat{F}_B,\ \text{Allocation:}F_B} > \underbrace{\mathbb{E}[\phi_B(q)y_B(q)]}_{\text{Distribution:}F_B,\ \text{Allocation:}F_B} \tag{14}$$

The first inequality holds since revenue under the old allocation rule, associated with value design (F_A, F_B), and new virtual value, derived from $\hat{F}_B$, is a lower bound on revenue obtained from the optimal allocation rule for the design $(F_A, \hat{F}_B)$. This follows as an adjustment of the allocation rule cannot decrease revenue, or else the principal would select the old allocation rule. We then turn to a comparison of virtual values, to show that $\mathbb{E}\left[\left(\hat{\phi}_B(q) - \phi_B(q)\right) y_B(q)\right] > 0$. Integrating by parts leads to a comparison of values, as $\phi(q) = \frac{\partial(v(q)q)}{\partial q}$. First order stochastic dominance implies that the value associated to each quantile under distribution $\hat{F}_B$ exceeds the value under distribution F_B, see also Figure 1. As such a distribution $\hat{F}_B$ can be constructed for any distribution F_B which is first order stochastically dominated by G, it follows that the optimal value design for agent B is in fact G.

This result immediately implies that when both agents' distributions can be designed the principal will find the design (G, G) to be optimal and generalizes to n-agent settings. It establishes that consistently destroying value by a first order stochastic dominance shift in the distribution relative to upper-bound G can never be optimal, since direct losses in surplus are never compensated by gains in information rents.

Our contributions on separable value design problems establish that whereas destroying value is never optimal for the principal, spreading value always is. In particular, the principal will seek to create divisive bimodal value designs in which agents either value the good very highly or not at all. For applications, this insight implies that different types display distinct valuations for a good and trade with vastly different likelihoods, as designed goods will be loved by few types and met with indifference by most.

3.3 Joint Design Constraints

We now turn to settings where the principal designs the values of their agents jointly. As demonstrated earlier, if the principal could design a good that was deemed more appealing by all agents, then he would do so. We therefore focus on scenarios in which various designs reallocate value across agents – meaning that, increasing the valuation for one agent will have to be compensated by decreasing the valuations for other agents.

The first design problem posits that any design (F_A, F_B) is feasible as long as the sum of expected values does not exceed a given bound, $\mathbb{E}_{F_A}[v]+\mathbb{E}_{F_B}[v] \leq k$ for some $k \in [2\underline{v}, \underline{v}+\overline{v}]$.[15] Under this constraint in any optimal design, the principal allocates all the disposable value to only one agent, leaving the disadvantaged agent with the smallest possible value $\underline{v}$ with certainty. The optimal design for the favoured agent allows the principal to extract all the value from the agent eliminating the information rent by selecting designs that either assign value $k-\underline{v}$ with certainty to the agent, or take on only two values, namely $\underline{v}$ and some value $h \in (k-\underline{v}, \overline{v}]$, while having $k-\underline{v}$ as a mean. Recall that we defined the class of two atom distributions with a given mean in expression (12).

Proposition 4. *A value design, (F_A^*, F_B^*), is optimal among all designs satisfying $\mathbb{E}_{F_A}[v] + \mathbb{E}_{F_B}[v] \leq k$:*
(1) for $k \in [2\underline{v}, \underline{v} + \overline{v}]$ if and only if $F_i^* = F_{k-\underline{v}}^h$ with $h \in [k - \underline{v}, \overline{v}]$ for some agent $i \in \{A, B\}$, while $v_j^*(q) = \underline{v}$ for all quantiles $q \in [0,1]$ for agent $j \neq i$;
(2) for $k \in (\underline{v} + \overline{v}, 2\overline{v}]$ if and only if $v_i^*(q) = \overline{v}$ for all quantiles $q \in [0,1]$ for some agent $i \in \{A, B\}$.

The intuition follows from insights developed in the benchmark case and in the separable design case with a bound on the mean of the designed distributions. As in the benchmark case, the principal only needs one agent to have a high valuation and therefore selects agent i to receive all the disposable expected value $k-\underline{v}$, while the other agent's expected value is reduced to the lower-bound $\underline{v}$ with certainty. As in the separable design case, the principal can opt for designs with different

[15]Note that by restrictions on possible values, we must have that $k \in [2\underline{v}, 2\overline{v}]$.

volatility for the favoured agent provided he can extract the full surplus in any of these. Consequently, a principal facing two potentially identical agents will induce sharp differences in how they value the good – for instance, by designing divisive products that cater to only one of the agents. Similar insights would go through in a multi-agent extension with all the disposable value being allocated to a single agent. When the mean bound $k \in [\underline{v} + \overline{v}, 2\overline{v}]$, some disposable value must be given to both players by feasibility. In any optimal design one agent is given the maximal value $\overline{v}$ with certainty, while the other agent is left with any other distribution, as the principal can extract the maximal surplus from the favoured agent.

The result demonstrates that it is profitable for the principal to allocate as much value as possible to one agent, while keeping the other agent at minimum value. It raises questions as to whether divisive goods would also be optimal when values can be reallocated, but more stringent constraints are imposed on the joint design. To address this, we allow for any design (F_A, F_B) to be feasible as long as $F_A(v) + F_B(v) = H(v)$ for all v, where $H(v)$ is the cumulative associated to an arbitrary measure with mass *two*, with $v \in [\underline{v}, \overline{v}]$. It turns out that in this setting divisive designs remain optimal as long as the measure H assigns mass to more than one value – a symmetric design cannot be optimal. Rather, the principal maximizes revenue by creating two maximally differentiated value distributions. To state the result, it is useful to define the median for the measure H, as the value v^M such that $\int_{\underline{v}}^{v^M} dH(v) = 1$, when such value exists. As such value may not exist when H is atomic, in these instances the median will instead be defined as the smallest value v^M such that $\int_{\underline{v}}^{v^M} dH(v) \geq 1$.

Proposition 5. *The value design, (F_A^*, F_B^*), is optimal among all designs satisfying $F_A(v) + F_B(v) = H(v)$ for all $v \in [\underline{v}, \overline{v}]$ if for some $j \in \{A, B\}$ and $i \neq j$*

$$\begin{aligned} F_i^*(v) &= H(v) - 1 \qquad &&\text{if } v \in [v^M, \overline{v}], \\ F_j^*(v) &= H(v) \qquad &&\text{if } v \in [\underline{v}, v^M]. \end{aligned}$$

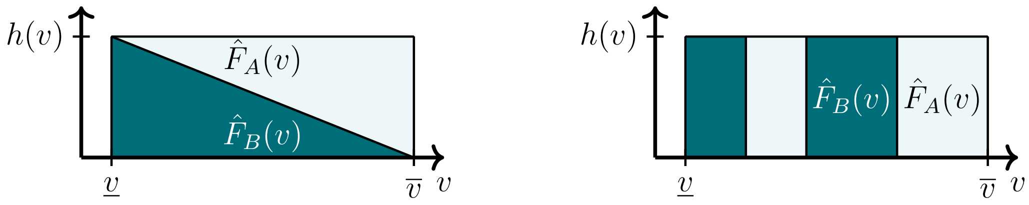
Maximally divisive goods, with one agent's values below the median of H and the other agent's values above the median of H, yield a higher revenue than any other design fulfilling the constraint. Such a design has two benefits. First, it maximizes surplus, $\mathbb{E}[\max\{v_A, v_B\}]$, since it generates the highest possible expected value for one of the agents. This is better than any other design, which necessarily yields more similar expected values, due to the convexity of the maximum operator. Second, a maximally divisive design narrows the support of both agents' distributions, and thereby reduces information rents.

Such a design however increases the probability that the good is allocated to the agent with high values thereby increasing their willingness to pay, and decreases the probability that the good is allocated to the agent with low values thereby reducing their willingness to pay. In spite of the decreased competition, the gains in transfers from the high value agent more than compensate the losses from the low value agent. The key takeaway of this design problem is that when value redistribution is possible, the principal benefits from divisive designs that treat agents differently and discriminates between them. This provides a novel rationale for designing products catering only to a subset of the market. The result generalizes to n-agent settings, but in those settings only the distribution of the two agent with highest values would be pinned down by optimality – with one agent receiving all the highest values with mass 1, another agent receiving all the remaining highest values with mass 1, while the distributions of the remaining agents would be arbitrary but for the need to fulfill the constraint.

To establish the optimality of maximally divisive designs, the proof first posits that H is differentiable and compares optimal revenue in such a design to revenue in any other design (F_A, F_B) in which two distributions F_A and F_B have the same support. An example of an alternative design is provided in Figure 2a. The proof then carries out the same comparison but for designs (F_A, F_B) in which the supports of the of the two distributions are disjoint supports, while allowing the measure H to be atomic. One instance of such a comparison is given in Figure 2b. The proof then concludes by establishing that the maximally divisive design necessarily leads to higher revenue than any design belonging to one of these two classes, and by arguing that the revenue in any other design simply amounts to linear combination of revenue in two designs belonging respectively to each of these two classes.

Next, we provide some intuition for the proof of the result for the case in which H

Figure 2: Alternative Distributions



(a) Same Support

(b) Disjoint Support

is differentiable and the supports of the designed distributions coincide. To capture an arbitrary split of the density $h(v) = H'(v)$, consider any design $(\hat{F}_A, \hat{F}_B)$ assigning a share $a_A(v) \in (0, 1)$ of the density $h(v)$ to $\hat{F}_A$, and the remaining share $a_B(v) = 1 - a_A(v)$ to $\hat{F}_B$. Derive allocation probabilities, $\hat{x}_A(v)$ and $\hat{x}_B(v)$, for such a setting. We compare the revenue generated by this arbitrary design to the revenue generated by a maximal divisive design. This is challenging as we need to keep track of the differences in virtual values as well as allocation probabilities.

To do so, it is useful to rank the two allocation probabilities $\hat{x}_A(v)$ and $\hat{x}_B(v)$ for each possible value v according to magnitude. If the value is above the median, then the maximum over $\hat{x}_A(v)$ and $\hat{x}_B(v)$ is assigned to $\overline{x}_A(v)$ and the minimum becomes $\overline{x}_B(v)$. If the value is below the median, then the maximal allocation probability is absorbed by $\overline{x}_B(v)$, while the minimum is now $\overline{x}_A(v)$. After this change, for any value v, it is possible to identify the ranking of the allocation probabilities.

With this relabelling, we can the show that the revenue for maximally divisive design is higher than revenue at any alternative design by establishing that

$$\begin{aligned}
\underbrace{\mathbb{E}[\psi_A^*(v)x_A^*(v)] + \mathbb{E}[\psi_B^*(v)x_B^*(v)]}_{\text{Distribution:}F_A^*, F_B^*,\ \text{Allocation:}x_A^*, x_B^*} &\geq \underbrace{\mathbb{E}[\psi_A^*(v)\overline{x}_A(v)] + \mathbb{E}[\psi_B^*(v)\overline{x}_B(v)]}_{\text{Distribution:}F_A^*, F_B^*,\ \text{Allocation:}\overline{x}_A, \overline{x}_B} \\
&\geq \underbrace{\mathbb{E}[\overline{\psi}_A(v)\overline{x}_A(v)] + \mathbb{E}[\overline{\psi}_B(v)\overline{x}_B(v)]}_{\text{Distribution:}\overline{F}_A, \overline{F}_B,\ \text{Allocation:}\overline{x}_A, \overline{x}_B} \\
&= \underbrace{\mathbb{E}[\hat{\psi}_A(v)\hat{x}_A(v)] + \mathbb{E}[\hat{\psi}_B(v)\hat{x}_B(v)]}_{\text{Distribution:}\hat{F}_A, \hat{F}_B,\ \text{Allocation:}\hat{x}_A, \hat{x}_B},
\end{aligned} \tag{15}$$

where $\overline{\psi}$ is the virtual value associated with $\bar{x}$. The first inequality follows again by replacing the optimal allocation probabilities x^* for the maximally divisive design with the constructed allocation probabilities, $\overline{x}$. The allocation probabilities x^* must be associated with a weakly higher revenue – otherwise they would not be optimal. However, it is not necessarily the case that $\overline{x}_A(v)$ and $\overline{x}_B(v)$ are feasible allocation probabilities under the new distributions. In Appendix D, we verify that these expected allocation probabilities are indeed interim feasible (Border (1991)) and can be be derived by taking expectations over a well-defined allocation rule, thus validating our approach.

The key to establishing the result then remains the second inequality in expression (15) since the last equality holds by definition as we only relabelled allocation probabilities and associated virtual values. Rewriting the relevant inequality from (15), yields

$$\begin{aligned}&\int_{v^M}^{\overline{v}} \psi_A^*(v)\overline{x}_A(v)h(v)dv + \int_{\underline{v}}^{v^M} \psi_B^*(v)\overline{x}_B(v)h(v)dv \\ &\quad > \int_{\underline{v}}^{\overline{v}} \left(a_A(v)\overline{\psi}_A(v)\overline{x}_A(v) + a_B(v)\overline{\psi}_B(v)\overline{x}_B(v)\right) h(v)dv. \end{aligned} \tag{16}$$

Observe that the density for each v in each integral is identical. This allows us to compare integrands pointwise for each v. Further, we assume for now that virtual values are positive, which yields

$$\begin{aligned}&\psi_i^*(v)\overline{x}_i(v) - a_A(v)\overline{\psi}_A(v)\overline{x}_A(v) - a_B(v)\overline{\psi}_B(v)\overline{x}_B(v) \\ &\geq \left(\psi_i^*(v) - a_A(v)\overline{\psi}_A(v) - a_B(v))\overline{\psi}_B(v)\right)\overline{x}_i(v), \end{aligned} \tag{17}$$

where the inequality follows from the ranking of the allocation probabilities. For $v > v^M$, $\overline{x}_A(v) > \overline{x}_B(v)$ and so we can replace $\overline{x}_B(v)$ by $\overline{x}_A(v)$. Similarly, for $v < v^M$, $\overline{x}_A(v) < \overline{x}_B(v)$, allowing us to replace $\overline{x}_A(v)$ with $\overline{x}_B(v)$. This simplifies the problem to comparing virtual values under the maximally divisive design to the average auxiliary virtual value. If the difference in virtual values is positive for both agents and all values v, revenue is higher in the divisive design compared to the chosen design. This is straightforward, and we obtain that the difference in virtual values is zero for values above the median v^M, while it is equal to one for values below the median. This implies that gains in marginal transfer contributions occur for values below the median. Above the median, the sum of information rents paid to agents coincide for any design chosen by the principal. But for values below the median, the sum of information rents paid to agents is minimized by assigning all such value to a single agent as this maximizes at once both $F(v)$ and $f(v)$ for the agent to whom such values are assigned.

If the virtual values are not positive, then more involved arguments, along the lines made above are necessary. These are provided in Appendix D.

We follow a similar approach to establish that designs with disjoint supports that differ from the maximally divisive design are sub-optimal. This follows because the maximally divisive design minimizes the information rent paid to agents with values below the median at the cost of increasing the information rent for agents above the median. For any other value design, the gain in information rent for higher values cannot make up for the loss at lower values – because the information rent is highest when values are low.

With disjoint supports, we also allow for atoms, and because of this, the proof for this case is developed in the quantile space. This approach is appealing since proofs can be stated in terms of values rather than virtual values, as the mapping between the two is straightforward. We therefore specify revenue in the quantile space, and using the same strategy as with split densities, we replace the allocation probabilities for the optimal design with those of some arbitrary designs. This creates a lower bound on revenue for the maximally divisive design. We then compare this lower bound to different designs, proving for each of them that the maximally divisive design is associated with a lower information rent, verifying its optimality.

Our proof establishes that the maximally divisive design increases virtual values by minimizing information rents. The increase in virtual values does not necessarily translate into a higher transfer from both agents. It is straightforward to show that the transfer for the low value agents will decrease as the value distributions become more unequal.[16] Transfers depend on both virtual values and allocation probabilities. The latter deteriorate for low values, ultimately triggering the reduction in

[16] A worked out example that includes transfers was included in a previous version and is available upon request.

transfer for the low value agent. Recall that the expected virtual surplus is given by

$$\mathbb{E}[\psi_A(v)x_A(\mathbf{v})] + \mathbb{E}[\psi_B(v)x_B(\mathbf{v})]. \tag{18}$$

Unequal distributions which favor A at the expense of B imply that x_A increases, while x_B decreases. Crucially, this is beneficial, as A's virtual values are higher than B's, again due to the distribution favouring A. Therefore, a divisive design leads to a higher allocation probabilities to higher values, while matching low virtual values with low allocation probabilities. This is another force as to why maximally divisive goods are optimal, and the reason why we can just fix the allocation probabilities: allowing them to adjust would only strengthen the result. In practice, this implies that the principal is better off by facing two agents, with distinct value distributions, instead of facing two agents with some average value distribution. Once again, inequality between agents is profitable for the principal, indicating that it is better to focus on goods that are divisive.

4 Conclusion

For separable design problems, bimodal divisive product designs are optimal when the constraint restricts the induced expected valuation or when the induced distribution is second order stochastically dominated by a given distribution. Increasing revenue is never achieved by destroying value in a first order stochastically dominated sense. Surplus growth plays a pivotal role in all value design problems, and first order stochastically dominated distributions result in reduced surplus, rendering them sub-optimal. In joint design problems, our focus lies in reallocating value, and we find that the principal once again favors divisive designs that make the good most appealing to only one agent at the expense of others.

In all such optimal divisive designs, the principal introduces inequalities within agents' types and across agents, as these inequalities reduce the information rents of the agents by making the valuation of the good more predictable during trade. Given that such divisive product designs prove to be profitable under a wide range of constraints, we anticipate their emergence in real-world economic settings. For example, auctioneers are likely to source goods that are highly valuable to a small group of agents but undesirable or unaffordable for the majority, especially when the cost of sourcing products depends solely on their mean valuation. Such designs enable the principal to focus on types for which the good is most appealing by setting high reserve prices, thereby increasing revenue, as observed in our separable design constraint with mean bounds. When the principal possesses knowledge of the agents' preferences and can favor one agent over others through the design or sourcing of goods, they do so by selecting products that cater exclusively to that agent, thereby enhancing predictability of their value. This pattern is evident in both of our joint design problems.

Our approach abstracts from costs, which we explore further in Appendix B. We focus on independent values, arguing that every agent values each good independently. Correlated values simplify our problem as we demonstrate in Appendix A, reducing our analysis to surplus-maximizing value design.

Last, there is the question of implementation. Following the principles outlined in Myerson (1981), a winner-pay implementation of the optimal mechanism may be a good fit for trade settings, but all-pay or other implementations may be more desirable in other applications.

A Correlated Private Values

Maintaining the independence of value distributions across design problems is a natural assumption if one believes that a bidder should not be able to infer anything for his value about values of his competitors, irrespective of the value design. The independence assumption further leads to a more challenging problem, as it requires accounting for agents' information rents rather than focusing purely on surplus design. However, our results carry over to a setting in which signals are not independent and bidders have some information about the value of their competitors.

Consider such a correlated information setting (Crémer and McLean (1985)) in which revenue coincides (generically) with surplus, $\mathbb{E}[\max\{v_A, v_B\}]$. To provide some insight while remaining close to the core of the analysis, we allow the principal to design marginal value distributions for both bidders, but not the correlation structure between their values which is fixed. Further, we assume that bidders know their value and have some information about their competitors' values too. The approach coincides with our baseline analysis where the principal designed marginal value distributions for both agents, but was unable to affect the independence of agents' values.

To fix the correlation structure, take any joint distribution in the quantile space $Q : [0,1]^2 \to [0,1]$. The joint distribution $Q(q_A, q_B)$ has uniform marginal distributions by construction, but can display arbitrary correlation structures – such as independence, perfect positive correlation (*concordance*), and perfect negative correlation (*discordance*). Posit that Q determines the underlying correlation of tastes across agents and cannot be affected by the principal, as was the case in the original setup where we always had that $Q(q_A, q_B) = (1 - q_A)(1 - q_B)$. As before, what the principal designs are the values associated to each quantile for both agents, $v_A(q_A)$ and $v_B(q_B)$. This is equivalent to designing the two marginal value distributions, $F_A(v_A)$ and $F_B(v_B)$. We consider this to be a suitable approach when quantiles reflect the underlying preferences of agents, since the correlation in tastes across agents cannot be designed.

Separable Design Insights: When designing marginal distributions subject to the mean-bound constraint $\mathbb{E}_{F_i}[v_i] \leq k$ for all i, extreme bimodal designs remain optimal. This follows because variance increases the expected value of the maximum of two random variables, $\mathbb{E}[\max\{v_A, v_B\}]$ by the convexity of the maximum operator. As an example, let quantiles be discordant, so that $q_A = 1 - q_B$ for any pair of quantiles (q_A, q_B) in the support of Q. Additionally assume that $\overline{v} = 2$, and that $k = 1$. If so, surplus, or equivalently revenue, is maximized by a value design (F_A, F_B) in which both agents value the good at $v_i(q_i) = 2$ if $q_i \leq 1/2$ and at $v_i(q_i) = 0$ otherwise. Such design corresponds to maximal dispersion as described in Proposition 2. Under such a maximally spread two-atom distribution, surplus is exactly equal to 2, as at least one agent must value the good at 2. Thus, revenue also equals 2, and the principal secures such a surplus by awarding the good with certainty to the agents with realized value equal 2. No other value design can lead to higher surplus, as values never exceed 2. Therefore, designs in which both distributions are maximally spread are optimal when values are negatively correlated.

With concordance, the design in which an agent always values the good at the mean, $k = 1$, while the other is maximally spread, as by Proposition 1, yields a surplus of 1.5. This exceeds the surplus from the design in which both agents are maximally spread which instead yields a surplus of 1. Similarly, if both agents have mass at the mean, the surplus equals 1. With positively correlated values the principal is better off when one agent has all mass at the mean, whereas the other agent's values are maximally spread. In this case, the designed distributions are maximally distinct subject to the constraint to ensure that surplus is extracted even when one agent has low value. Therefore, allowing for such positive value correlation strengthens our insight that the principal benefits from inequalities.

Joint Design Insights: When designing marginal distributions subject to the reallocation constraint $F_A(v) + F_B(v) = H(v)$ for all v, maximally divisive designs remain optimal. This follows because such designs increases surplus, $\mathbb{E}[\max\{v_A, v_B\}]$, by minimizing the chance of having two agents with high values. As an example, let quantiles be concordant, so that $q_A = q_B$ for any pair of quantiles (q_A, q_B) in the support of Q. Additionally assume that $\overline{v} = 2$, and that $H(v) = v$ for all $v \in [0, 2]$. If so, surplus with a maximally divisive design simply amounts to the expected value, $\mathbb{E}_{F_i^*}[v_i]$, of the agent i with the high marginal distribution $F_i^*(v_i) = H(v_i) - 1$, which is equal to 1.5. Thus, revenue also amounts to 1.5. Such a transfer can be obtained by never awarding the good to bidder $j \neq i$ and awarding the good to bidder i only when his transfer coincides with his reported value, and the values reported by the two bidders are associated to the same quantile, $q_A(v_A) = q_B(v_B)$. In minimally divisive designs, $F_i(v_i) = H(v_i)/2$ for all i, surplus simply amounts to the expected value of one of the two agents, $\mathbb{E}_{F_i}[v_i]$, which is equal to 1. In such joint design settings, the maximally divisive value designs, discussed in Proposition 5, remain optimal irrespective of correlation. But other designs may also be optimal for specific correlation structure. For instance in the example discussed above, when quantiles are discordant and $q_A = 1 - q_B$, the minimally divisive design yields the same surplus as the maximally divisive one. But discordance generally benefits surplus in the minimally divisive design, as the principal is certain to face an agent with values in excess of the median in such settings.

B Cost of Value Design

Our approach focuses on design constraints rather than costs to highlight the effects of different designs on the principal's revenue (net of design costs). Nevertheless, our two main results are robust to costly design.

Value Dispersion Conceptually, constraint (6) posits that any distribution with a mean at or below k costs the same and we characterise the optimal distribution under this assumption. It may be sensible to posit that value distributions with lower means should have a lower design cost. In such settings, there would be a force pushing towards distributions that have a lower mean and with it a lower cost. The shape of the cost function would determine the mean set by the principal, which would be positive unless costs are prohibitively high. Setting a mean at zero would lead to zero valuation among agents and nobody would compete for the good.

Therefore, it seems sensible to assume that the principal will select some positive k. Given a positive mean chosen by the principal, he would still choose distributions that display maximal dispersion. We also explicitly consider the case where the principal incurs a fixed cost from adjusting the distribution in our analysis, see Remark 3.

Value Reallocation We then turn to the case of reallocating values across agents. Although the assumption on redistribution implicit in this model may seem strong, similar insights would naturally arise in settings in which designing is costly, and costs satisfy a weak linearity property requiring that any two value distributions having the same sum to cost the same. This assumption would be fulfilled by common design cost functions such as entropy, and would be met for instance by integrable cost functions satisfying

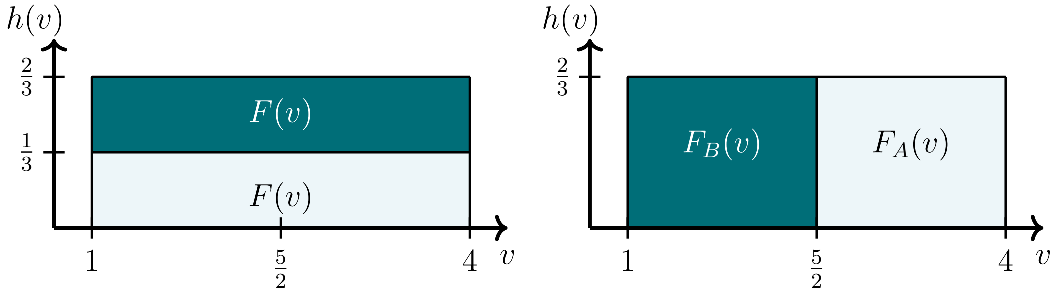
$$\int_0^{\overline{v}} c(v)d[F_A(v) + F_B(v)]. \tag{19}$$

for some function $c : [\underline{v}, \overline{v}] \to \mathbb{R}_+$.

Thus, even in a model of costly design, we would get maximally divisive products and maximal dispersion as features of optimal designs.

C An Example Of Jointly Optimal Distributions

For the purpose of this example, we allow the principal to generate the value distributions for two agents, A and B, adding up to $H(v) = \frac{2(v-1)}{3}$ with $v \in [1, 4]$. We contrast the revenue if the principal treats their two agents equally and generates cumulative distribution functions $F(v) = \frac{(v-1)}{3}$ for both agents, with revenue in the maximally divisive design. The latter design consists of two maximally distinct distributions, where one agent receives all the values below the median of $H(v)$, his distribution equals $F_B(v) = H(v)$ for $v < \frac{5}{2}$, while the other agent's distribution is given by $F_A(v) = H(v) - 1$ for $v > \frac{5}{2}$. These distributions are depicted in Figure 3. The principal implements a mechanism which specifies the probability of good as well as a transfer level for each induced valuation such that agents find it optimal to



(a) Minimally Divisive Design

(b) Maximally Divisive Design

Figure 3: Minimally and Maximally Divisive Designs

reveal their net value for the good (instead of pretending they have a different induced value). We follow Myerson (1981) in mapping value distributions into transfer and good probabilities. He shows that maximizing the sum of transfers is equivalent to maximizing the *expected virtual surplus*, which is a function of allocation probabilities and *virtual values*. The virtual value for distributions with defined densities is given by $\psi(v) = v - \frac{1-F(v)}{f(v)}$. The virtual surplus for a given profile of values v_A, v_B is then $\psi_A(v_A)x_A(v_A, v_B) + \psi_B(v_B)x_B(v_A, v_B)$, where x_A and x_B specify the expected good probabilities for A and B, respectively.[17] The agent with the highest realised virtual value receives the good, if his virtual value is positive. Otherwise, the good is not allocated.

We depict the expected virtual surplus of each agent, given his realised value, in Figure 4a. We calculate for each value the virtual value and the probability that the agent obtains the good. The expected virtual surplus for each value with the maximally divisive design is above the expected virtual surplus if distributions are equal. However, there are *two* individuals with distribution F, but only one with F_A and F_B, each.

Two agents with distribution F and realised values close to four generate a higher expected surplus, compared to two agents whose valuations are drawn from F_A and F_B. Nevertheless, the maximally divisive design yields a higher expected surplus if one takes the densities of the different values into account, determined by their

[17]When calculating the probability of obtaining the good, the agent knows his valuation, but only the distribution of the other agent's values. Note that the expected virtual surplus here is calculated for a *given* value v. That is, the expectation is taken over the other agent's value distribution. In our main analysis, we refer to expected virtual surplus when taking expectations over *all* value distributions.

respective distributions. Formally, we want to show the following:

$$\int_{\frac{5}{2}}^{4} \psi_A(v)x_A(v)h(v)dv + \int_{1}^{\frac{5}{2}} \psi_B(v)x_B(v)h(v)dv > 2\int_{1}^{4} \psi(v)x(v)f(v)dv.$$

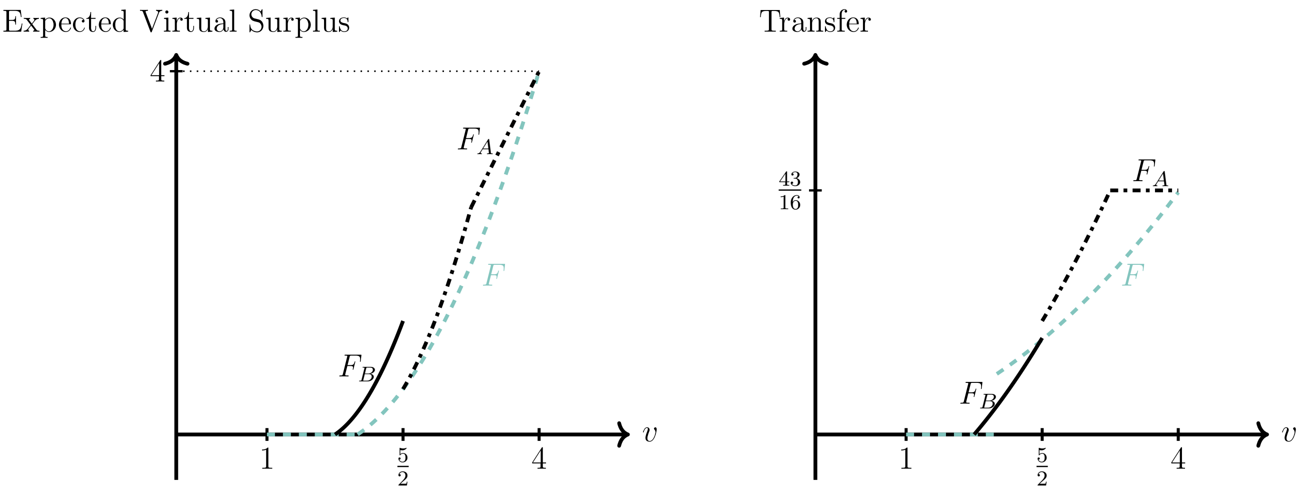
Note that $2f(v) = h(v) = H'(v)$ by definition and so the previous inequality simplifies to

$$\int_{\frac{5}{2}}^{4} \psi_A(v)x_A(v)h(v)dv + \int_{1}^{\frac{5}{2}} \psi_B(v)x_B(v)h(v)dv > \int_{1}^{4} \psi(v)x(v)h(v)dv.$$

It is as if one is comparing both F_A and F_B agents with *one* agent with distribution H as each value is associated with the same density. Therefore, 4a does in fact reveal that the maximally divisive design is associated with a higher virtual surplus and in turn a higher transfer.

Figure 4: Expected Virtual Surplus vs Transfer

(a) Expected Virtual Surplus (b) Transfer



Note: Expected Virtual Surplus of agent i: $\psi_i(v_i)x_i(v_i)$. The light, dashed line represents F, the solid line F_B and the dashed dotted line F_A.

The transfer is made explicit in Figure 4b. The maximally divisive design generates a higher transfer from those with high valuations, a lower transfer for intermediate values and a higher transfer for low values. The last effect is due to different exclusion thresholds. The maximally divisive design leads to fewer values being excluded, as the information rent is lower for each value. This is evident from Figure 4a: the value for which the expected virtual surplus becomes positive is higher if both agents have the same distribution compared to maximally distinct distributions. If there is no exclusion, the maximally divisive design reduces the transfer of the agent with lower values. However, this reduction is more than compensated for by the increase in transfer by the high value agent, making divisive designs optimal.

D Mathematical Proofs

Proof of Remark 1: No Constraints Note first that it is not possible to obtain a revenue higher than $\overline{v}$:

$$\mathbb{E}\left[\phi_A(q)y_A(q)\right] + \mathbb{E}\left[\phi_B(q)y_B(q)\right] \leq \overline{v}, \tag{20}$$

where the inequality follows from $\phi_i(q) \leq \overline{v}$ for all $i \in \{A, B\}$, and $E\left[y_A(q)\right] + \mathbb{E}\left[y_B(q)\right] \leq 1$ for any allocation rule. Thus, as long as at least one agent has value $\overline{v}$ with certainty, the principal extracts the highest possible revenue. ■

Proof of Proposition 1: Two-Agent Value Dispersion The proof proceeds in three steps: first it finds value designs maximizing surplus; then, it argues that the optimal value designs proposed in the statement indeed maximize surplus and consequently revenue; and finally, it establishes why no other value design can be optimal.

Consider any design (F_A, F_B) fulfilling the constraint. Total surplus amounts to

$$S = \int_{\underline{v}_A}^{\overline{v}_A} \int_{\underline{v}_B}^{\overline{v}_B} \max\{v_A, v_B\} dF_B(v_B) dF_A(v_A), \tag{21}$$

where Riemann-Stijltes integrals are used to calculate the expectations implicitly.

Conditional on any realized value of $v_A \in [\underline{v}_B, \overline{v}_B]$, surplus amounts to

$$S(v_A) = \int_{\underline{v}_B}^{\overline{v}_B} \max\{v_A, v_B\} dF_B(v_B) = \int_{\underline{v}_B}^{v_A} v_A dF_B(v_B) + \int_{v_A}^{\overline{v}_B} v_B dF_B(v_B) \tag{22}$$

$$= v_A F_B(v_A) + [v_B F_B(v_B)]_{v_A}^{\overline{v}_B} - \int_{v_A}^{\overline{v}_B} F_B(v_B) dv_B \tag{23}$$

$$= \overline{v}_B - \int_{v_A}^{\overline{v}_B} F_B(v_B) dv_B, \tag{24}$$

where the first equality is algebraic, the second follows from calculations and integration by parts, and the third equality simply factors redundant terms. For $v_A \geq \overline{v}_B$, surplus amounts to $S_A(v_A) = v_A$ since $\max\{v_A, \overline{v}_B\} = v_A$. While for $v_A \leq \underline{v}_B$, surplus amounts to $S_A(v_A) = k$ since $\max\{v_A, \underline{v}_B\} = \underline{v}_B$. Observe that $S(v_A)$ is differentiable in $v_A \in [\underline{v}_A, \overline{v}_A]$, and is increasing since for $v_A \in [\underline{v}_B, \overline{v}_B]$,

$$S'(v_A) = -\frac{d}{dv_A} \int_{v_A}^{\overline{v}_B} F_B(v_B) dv_B = F_B(v_A) \tag{25}$$

by Leibniz rule; while $S'(v_A) = 1$ for $v_A > \overline{v}_B$ and $S'(v_A) = 0$ for $v_A < \underline{v}_B$. Finally, $S(v_A)$ is convex since $S(v_A) - S(v'_A) \geq (v_A - v'_A)S'(v'_A)$ for any values $v_A, v'_A \in [\underline{v}_A, \overline{v}_A]$. To show this, assume without loss that $v_A \geq v'_A$ and observe that: when $v'_A > \overline{v}_B$

$$S(v_A) - S(v'_A) = v_A - v'_A = (v_A - v'_A)S'(v'_A), \tag{26}$$

given that A's values always exceed B's; when $\overline{v}_B \geq v_A \geq v'_A \geq \underline{v}_B$

$$S(v_A) - S(v'_A) = \int_{v'_A}^{v_A} F_B(v_B) dv_B \geq \int_{v'_A}^{v_A} F_B(v'_A) dv_B \tag{27}$$

$$= (v_A - v'_A)F_B(v'_A) = (v_A - v'_A)S'(v'_A), \tag{28}$$

where the inequality follows because F_B is increasing; when $v_A > \overline{v}_B \geq v'_A \geq \underline{v}_B$

$$S(v_A) - S(v'_A) = v_A - \overline{v}_B + S(\overline{v}_B) - S(v'_A) \geq v_A - \overline{v}_B + (\overline{v}_B - v'_A)S'(v'_A) \tag{29}$$

$$\geq (v_A - v'_A)S'(v'_A), \tag{30}$$

where the equality follows because $\overline{v}_B = S(\overline{v}_B)$, the first inequality follows from (27), and the second inequality follows from $S'(v_A) \leq 1$ and $v_A > \overline{v}_B$; when $v_A > \overline{v}_B$ and $\underline{v}_B > v'_A$

$$S(v_A) - S(v'_A) = v_A - k \geq 0 = (v_A - v'_A)S'(v'_A), \tag{31}$$

where then inequality follows as $k \in [\underline{v}_B, \overline{v}_B]$ and $V_A > \overline{v}_B$; when $\overline{v}_B \geq v_A \geq \underline{v}_B > v'_A$

$$S(v_A) - S(v'_A) \geq 0 = (v_A - v'_A)S'(v'_A), \tag{32}$$

where the inequality follows because surplus is increasing; while for $\underline{v}_B > v_A$ we have that

$$S(v_A) - S(v'_A) = (v_A - v'_A)S'(v'_A) = 0, \tag{33}$$

given that B's values always exceed A's. Thus, surplus, S, conditional on v_A is globally convex in v_A.

The surplus maximizing design F_A given F_B must solve

$$\max_{F_A} \int_{\underline{v}}^{\overline{v}} S(v_A) dF_A(v_A) \quad \text{subject to} \quad \int_{\underline{v}}^{\overline{v}} v_A dF_A(v_A) \leq k. \tag{34}$$

Because the objective function of the problem is convex and increasing in v_A while the constraint is linear, there always exists a surplus-maximizing design which is maximally spread and has mean k – meaning that setting $F_A = F^{\overline{v}}$ maximizes surplus for any F_B. In particular, surplus will be maximized by any design (F_A, F_B) satisfying $F_i = F^{\overline{v}}$ for some $i \in \{A, B\}$ and $\int_{\underline{v}}^{\overline{v}} v_j dF_j(v_j) = k$ for $j \neq i$. For any such design, surplus would amount to

$$S^* = P^{\overline{v}}(\overline{v}) \int_{\underline{v}_j}^{\overline{v}_j} \max\{\overline{v}, v_j\} dF_j(v_j) + P^{\overline{v}}(\underline{v}) \int_{\underline{v}_j}^{\overline{v}_j} \max\{\underline{v}, v_j\} dF_B(v_j) \tag{35}$$

$$= P^{\overline{v}}(\overline{v}) \int_{\underline{v}_j}^{\overline{v}_j} \overline{v} dF_j(v_j) + P^{\overline{v}}(\underline{v}) \int_{\underline{v}_j}^{\overline{v}_j} v_j dF_j(v_j) \tag{36}$$

$$= \frac{k - \underline{v}}{\overline{v} - \underline{v}} \overline{v} + \frac{\overline{v} - k}{\overline{v} - \underline{v}} k, \tag{37}$$

where the first equality follows from the definition of surplus, the second follows because $v_B \in [\underline{v}, \overline{v}]$, and the third as $\int_{\underline{v}_j}^{\overline{v}_j} v_j dF_j(v_j) = k$.

Next note that for any value design surplus bounds revenue from above since

$$S(F_A, F_B) = R(F_A, F_B) + U_A(F_A, F_B) + U_B(F_A, F_B) \geq R(F_A, F_B), \tag{38}$$

where the inequality follow from individual rationality of the two players. So, the revenue never strictly exceeds the value of surplus in a surplus-maximizing design. But for any proposed optimal design in the statement of the result, we have that $R(F_A, F_B) = S^*$, and thus these designs must indeed be optimal.

We conclude by arguing why no other design can be optimal. Consider any surplus maximizing design satisfying $F_i = F^*$ for some $i \in A, B$ and $\int_{\underline{v}}^{\overline{v}} v_j dF_j(v_j) = k$ for $j \neq i$. Observe that revenue at such design amounts to

$$R(F_A, F_B) = P^{\overline{v}}(\overline{v})\psi_i(\overline{v}) + P^{\overline{v}}(\underline{v}) \int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), \psi_i(\underline{v}), 0\} dF_j(v_j) \tag{39}$$

$$= \frac{k - \underline{v}}{\overline{v} - \underline{v}} \psi_i(\overline{v}) + \frac{\overline{v} - k}{\overline{v} - \underline{v}} \int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), \psi_i(\underline{v}), 0\} dF_j(v_j). \tag{40}$$

Further, for any of the optimal designs in the statement of the result, we have that $R(F_A, F_B) = S^*$: because $\psi_i(\overline{v}) = \overline{v}$; because for $\underline{v} > 0$, in the unique optimal design F^k, we have that

$$\int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), \psi_i(\underline{v}), 0\} dF_j(v_j) = \int_{\underline{v}_j}^{\overline{v}_j} \max\{k, \psi_i(\underline{v}), 0\} dF_j(v_j) = k, \tag{41}$$

where the first equality follows as $\psi_j(v_j) = v_j = k$ and the second equality follows since $k \geq \underline{v} \geq \max\{\psi_i(\underline{v}), 0\}$; and because for $\underline{v} = 0$, in any optimal design F^h, we have that

$$\int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), \psi_i(\underline{v}), 0\} dF_j(v_j) = \int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), 0\} dF_j(v_j) \tag{42}$$

$$= P^h(h)h + P^h(0)0 = k, \tag{43}$$

where the first equality follows as $\psi_i(\underline{v}) \leq \underline{v} = 0$, the second one holds as $\psi_j(\overline{v}_j) =$

$\overline{v}_j = h$ and $\psi_j(0) \leq 0$, while the third one hold by definition.

However, no other surplus-maximizing design can be optimal, because the only way to secure a revenue equal to S^* requires guaranteeing that $\psi_i(\overline{v}) = \overline{v}$ and

$$\int_{\underline{v}_j}^{\overline{v}_j} \max\{\psi_j(v_j), \psi_i(\underline{v}), 0\} dF_j(v_j) = \int_{\underline{v}_j}^{\overline{v}_j} v_j dF_j(v_j) = k. \tag{44}$$

But for the latter to hold, F_j can have either a single atom at the mean k, or only two atoms with one atom at 0 and the other above the mean k – which is only possible if $\underline{v} = 0$. In any other surplus-maximizing design, the designer would need to pay an information rent to bidder j to get them to reveal their type truthfully. ■

Proof of Proposition 2: One-Agent Value Dispersion We compare revenue under F^* to revenue when setting some other distribution F_B such that $\mathbb{E}_{F_B}[v] = k$. In a quantile setting, for $\overline{q} = 1 - F^*(0)$, the virtual value for the distribution F^* amounts to

$$\phi_B^*(q) = \begin{cases} \overline{v} & \text{if } q < \overline{q} \\ 0 & \text{if } q > \overline{q} \end{cases}. \tag{45}$$

revenue under distribution F^* is given by

$$\mathbb{E}\left[\phi_A(q) y_A^*(q)\right] + \mathbb{E}\left[\phi_B^*(q) y_B^*(q)\right], \tag{46}$$

while the transfer under any alternative F_B amounts to

$$\mathbb{E}\left[\phi_A(q) y_A(q)\right] + \mathbb{E}\left[\phi_B(q) y_B(q)\right]. \tag{47}$$

The change from F_B to F^* has two effects: (i) it affects the virtual valuation of agent B; and (ii) it affects the optimal allocation rule for the good. We begin by keeping the allocation rule fixed in the quantile space and show that a change from F_B to F^* increases revenue under the optimal allocation rule for distribution F_B.

Such insight then immediately delivers the result, since optimality implies that

$$\mathbb{E}\left[\phi_A(q)y_A^*(q)\right] + \mathbb{E}\left[\phi_B^*(q)y_B^*(q)\right] \geq \mathbb{E}\left[\phi_A(q)y_A(q)\right] + \mathbb{E}\left[\phi_B^*(q)y_B(q)\right]. \tag{48}$$

Note that fixing the allocation rule in the quantile space to $\mathbf{y}(\mathbf{q}) = \mathbf{x}(\mathbf{v}(\mathbf{q}))$ implies that interim allocation rules are also unchanged in the quantile space, since $y_i(q_i) = \int_0^1 \mathbf{y}_i(\mathbf{q})dq_j$, which was implicitly assumed in the previous expression.
Therefore, to prove the result, it suffices to establish that

$$\mathbb{E}\left[\phi_B^*(q)y_B(q)\right] - \mathbb{E}\left[\phi_B(q)y_B(q)\right] \geq 0. \tag{49}$$

For any $q < \overline{q}$, it must be that $\phi_B^*(q) = \overline{v} \geq \phi_B(q)$, since $\phi_B(q) = v_i(q) + v_i'(q)q \leq v_i(q) \leq \overline{v}$. Instead, for $q \geq \overline{q}$ and $y_B(q) > 0$, it must be that $\phi_B(q) \geq \phi_B^*(q) = 0$, since the good will only be given to a agent with a non-negative virtual value. As incentive compatibility requires $y_B(q)$ to be non-increasing, we have that if

$$\mathbb{E}\left[\phi_B^*(q) - \max\{\phi_B(q), 0\}\right] \geq 0 \quad \Rightarrow \quad \mathbb{E}\left[(\phi_B^*(q) - \phi_B(q))y_B(q)\right] \geq 0. \tag{50}$$

Moreover, by construction, the expected virtual values satisfy

$$\mathbb{E}\left[\phi_B^*(q)\right] = \mathbb{E}\left[v_B^*(q)\right] = \overline{q}\overline{v} = \mathbb{E}\left[v_A(q)\right] = \mathbb{E}\left[v_B(q)\right]. \tag{51}$$

The result then obtains, because expected virtual value for F_B satisfies

$$\mathbb{E}\left[v_B(q)\right] \geq \mathbb{E}\left[\max\left\{\phi_B(q), 0\right\}\right], \tag{52}$$

given that $\max\left\{\phi_B(q), 0\right\} \leq v_B(q)$ for all $q > 0$ since $v_B(q) \geq 0$ and $\phi_B(q) - v_B(q) = qv_B'(q) \leq 0$.
This establishes that it is never optimal to select a distribution for B with an expected virtual value which is strictly smaller than the expected value. However, there are multiple distribution that allow for the expected value to be equal to the expected virtual value, $\mathbb{E}\left[v_B(q)\right] = \mathbb{E}\left[\phi_B(q)\right]$. Note that it is never optimal to allocate mass to more than two values $v > 0$. Suppose to the contrary, the

principal chose such a distribution. Then, the allocation probability would need to differ across the different valuations for the mechanism to be incentive compatible. Otherwise a agent with a higher valuation would pretend to be a agent with a lower valuation and still obtain the good with the same probability. It follows that with such a distribution it would not be feasible to extract the entire valuation, yielding the contradiction.

To show that F^* is the unique optimum when $F_A \neq F^*$, consider any other distribution F_B with mean k and a single atom on positive values at $\overline{v}_B$, meaning that

$$F_B(v) = \begin{cases} 1 & \text{if} \quad v = \overline{v}_B \\ 1 - \frac{k}{\overline{v}_B} & \text{if} \quad v < \overline{v}_B \end{cases} . \tag{53}$$

Letting $p_B = k/\overline{v}_B$ denote the probability of having value $\overline{v}_B$, revenue can be rewritten as

$$\mathbb{E}[\phi_A(q)y_A(q)] + \mathbb{E}[\phi_B(q)y_B(q)] \tag{54}$$

$$= (1 - p_B) \int_{V_A} \max\{\psi_A(v_A), 0\} dF_A(v_A) + p_B \int_{V_A} \max\{\psi_A(v_A), \overline{v}_B\} dF_A(v_A) \tag{55}$$

$$= (1 - p_B) \int_{V_A} \max\{\psi_A(v_A), 0\} dF_A(v_A) + \int_{V_A} \max\{p_B \psi_A(v_A), k\} dF_A(v_A). \tag{56}$$

The first equality follows as the optimal mechanism allocates the good to A when their virtual valuation is positive and $v_B = 0$, and when their virtual valuation exceeds $\overline{v}_B$ if $v_B = \overline{v}_B$.[18]

The last expression is differentiable in p_B. Next we differentiate such expression with respect to p_B and establish that revenue decreases in p_B, meaning that the the optimal distribution will set p_B to be as small as possible, or equivalently $\overline{v}_B$ as large as possible. Letting $V_+(k) = \{v_A \in [0, \overline{v}] | \psi_A(v_A) \geq k\}$ for $k \geq 0$ denote the set

[18]If F_A was not regular, the previous expression for revenue would still apply. In such scenarios, F_A would denote the ironed distribution of values yielding the same revenue, rather than F_A itself – see Hartline (2013), Theorem 3.14, p.78.

of values for which A's virtual valuation weakly exceeds k, we find that

$$\frac{\partial \mathbb{E}[\phi_A(q)y_A(q) + \phi_B(q)y_B(q)]}{\partial p_B} = -\int_{V_+(0)} \psi_A(v_A)dF_A(v_A) + \int_{V_+(\overline{v}_B)} \psi_A(v_A)dF_A(v_A) \leq 0, \tag{57}$$

where the inequality holds, since $V_+(\overline{v}_B) \subseteq V_+(0)$. Moreover, the inequality is strict whenever $\psi_A(v_A) \in (0, \overline{v}_B)$ for some $v_A \in V_A$. Thus, it is optimal to minimize p_B which is accomplished by setting $\overline{v}_B = \overline{v} \geq \overline{v}_A$. This establishes the uniqueness result when $F_A \neq F^*$. To prove tho optimality of designs F^K for $K \in [k, \overline{v}]$ when $F_A = F^*$, it suffices to show that any design F^K yields the same revenue as F^*, which is immediate and thus omitted. ■

Proof of Corollary 2.1: Second Order Stochastic Dominance We show that F^* is second order stochastically dominated by G. If $G = F^*$, then F^* is second order stochastically dominated by G trivially. If $G \neq F^*$, it must be that for some $\overline{v}$ in the support of G and all $v \in [0, \overline{v})$,

$$\Delta(v) = G(v) - F^*(v) < 0, \tag{58}$$

since F^* places the maximal possible measure on $v = 0$ amongst all distributions with mean equal to k. It follows that for any $v \in [0, \overline{v})$, we have that

$$\int_0^v \Delta(s)ds = \int_0^v G(s) - F^*(s)ds < 0. \tag{59}$$

Moreover at $\overline{v}$, Riemann Stieltjes integration by parts yields

$$\int_0^{\overline{v}} \Delta(s)ds = [s\,(G(s) - F^*(s))]_0^{\overline{v}} - \int_0^{\overline{v}} sd\,(G(s) - F^*(s)) = 0. \tag{60}$$

To establish second order stochastic dominance, it then suffices to show that $\Delta(v)$ is non-decreasing, since that implies $\int_0^v \Delta(s)ds < 0$ for all $v \in [0, \overline{v})$. This holds as G is non-decreasing and F^* is constant up until $\overline{v}$. Therefore, G second order stochastically dominates F^*. The proof then follows because any distribution that is second order stochastically dominated by G must have a lower mean than G, and because F^* was optimal among all distributions with mean lower than G. ■

Proof of Proposition 3: First Order Stochastic Dominance Suppose by contradiction that the principal found it optimal to set $F_B \neq G$, which implies that $F_B(v) > G(v)$ for some v. We show that in this case, there exists a profitable deviation to a distribution $\hat{F}_B$, such that $F_B(v) \geq \hat{F}_B(v)$ with strict inequality for some v and $\hat{F}_B(v) \geq G(v)$.

As in the proof for value dispersion, the change from F_B to $\hat{F}_B$ has two effects: (i) it affects the virtual valuation of agent B and (ii) it affects the optimal allocation rule for the good. We again begin by keeping the allocation rule fixed and show that a change from F_B to $\hat{F}_B$ increases revenue under the same allocation rule. Recall that optimal revenue is given by[19]

$$\mathbb{E}[\phi_A(q)y_A(q)] + \mathbb{E}[\phi_B(q)y_B(q)] = \int_0^1 \phi_A(q)y_A(q)dq + \int_0^1 \phi_B(q)y_B(q)dq. \tag{61}$$

As the allocation rule is unchanged, $\mathbb{E}[\phi_A(q)y_A(q)]$ is not affected, and we only need to establish that

$$\mathbb{E}[\hat{\phi}_B(q)y_B(q)] \geq \mathbb{E}[\phi_B(q)y_B(q)] \quad \Leftrightarrow \quad \int_0^1 \left(\hat{\phi}_B(q) - \phi_B(q)\right) y_B(q)dq \geq 0. \tag{62}$$

Denote by $D \subseteq [0,1]$ the set of points at which y_B in non-differentiable, and by $C = [0,1] \setminus D$. The interim allocation rule y_B can be discontinuous, when the distribution F_B is not continuously differentiable and if there are gaps in the support. We first consider the case where only the allocation rule is discontinuous, before turning to gaps in valuations. Integration by parts using Riemann Stieltjes integrals implies that

$$\mathbb{E}[\phi_B(q)y_B(q)] = \underline{v}_B y_B(1) + \sum_{q\in D} qv_B(q)\left(-J(q)\right) - \int_C qv_B(q)dy_B(q), \tag{63}$$

where $J(q) = y_B^+(q) - y_B^-(q) \equiv \lim_{\epsilon\to 0}[y_B(q+\epsilon) - y_B(q-\epsilon)] < 0$ and the inequality

[19] If there are gaps in the support, the following expression requires a slight amendment, which does not affect results, but adds additional notation.

follows from y_B being decreasing. Exploiting the latter, we find that

$$\mathbb{E}[(\hat{\phi}_B(q) - \phi_B(q))y_B(q)] = \int_0^1 (\hat{\phi}_B(q) - \phi_B(q))y_B(q)dq \tag{64}$$

$$= (\underline{\hat{v}}_B - \underline{v}_B)y_B(1) + \sum_{q\in D} q(\hat{v}_B(q) - v_B(q))\,(-J(q)) + \int_C (q\hat{v}_B(q) - qv_B(q))\,(-y'_B(q))\,dq. \tag{65}$$

As $\hat{F}_B$ first order stochastically dominates F_B, we know that $\underline{\hat{v}}_B \geq \underline{v}_B$ and that $\hat{v}_B(q) \geq v_B(q)$. Thus, since the allocation probability is decreasing in q by incentive compatibility, $y'_B(q) \leq 0$, it follows that

$$\mathbb{E}[\phi_A(q)y_A(q)] + \mathbb{E}[\hat{\phi}_B(q)y_B(q)] \geq \mathbb{E}[\phi_A(q)y_A(q)] + \mathbb{E}[\phi_B(q)y_B(q)]. \tag{66}$$

Moreover, if the principal was allowed to set the allocation rule $\hat{y}(q)$ optimally, revenue would further increase

$$\mathbb{E}[\phi_A(q)\hat{y}_A(q)] + \mathbb{E}[\hat{\phi}_B(q)\hat{y}_B(q)] \geq \mathbb{E}[\phi_A(q)y_A(q)] + \mathbb{E}[\hat{\phi}_B(q)y_B(q)], \tag{67}$$

or else the principal would prefer to leave allocation rule unchanged.

Suppose now $F_B(v)$ has gaps in its support. Then, for each of these discontinuities in D, we need to show that

$$v_B^-(q)y_B^-(q) - v_B^+(q)y_B^+(q) \leq \hat{v}_B^-(q)y_B^-(q) - \hat{v}_B^+(q)y_B^+(q), \tag{68}$$

where $v_B^-(q)$ and $v_B^+(q)$ are defined in line with the allocation probabilities. If $G(v)$ does not have gaps in its support, there always exists a distribution $\hat{F}_B$ that first order stochastically dominates F_B and has continuous support. In this case, $\hat{v}_B^-(q)y_B^-(q) - \hat{v}_B^+(q)y_B^+(q) = \hat{v}_B^-(q)\,(-J(q))$ and as values are decreasing in q, $\hat{v}_B^-(q)\,(-J(q)) > v_B^-(q)y_B^-(q) - v_B^+(q)y_B^+(q)$. Thus, there exists once again a deviation that leads to a higher revenue.

Last, suppose the support of the distribution G has gaps. In this case, for a given q, there can be a jump for F_B, a jump for $\hat{F}_B$ or a jump for both. If there is a jump for exactly one distribution, first order stochastic dominance implies that $\hat{v}$

lies above v ($\hat{v}^+, \hat{v}^- > v$ or $v^+, v^- < \hat{v}$). If there is a jump for both distributions at the same q with the jump in values is larger under F_B, then inequality (68) is violated. We therefore require a more sophisticated approach. In this case, we need to amend expression (63) to

$$\mathbb{E}[\phi_B(q)y_B(q)] = \underline{v}_B y_B(1) + \sum_{q\in D} q\left[v_B^-(q)y_B^-(q) - v_B^+(q)y_B^+(q)\right] - \int_C qv_B(q)dy_B(q). \tag{69}$$

We compare distribution F_B to distribution $\hat{F}_B(v) = F_B(v)$ for $v \notin [v^+, \hat{v}^+]$, where $v^+ < \hat{v}^+$. This implies that $\hat{F}_B(\hat{v}^+) = F_B(v^+)$. Such a F_B yields a higher revenue than another distribution that displays a gap between v^+ and $\hat{v}^+$.

We can now flip this expression into the value-quantile-space, which yields

$$\underline{v}_B x_B(1) + \sum_{v\in D} vq_B(v)\left[x_B^+(v) - x_B^-(v)\right] + \int_{V_A\setminus D} q_B(v)vx_B'(v)dv, \tag{70}$$

Integrating over values allows to capture gaps in the support directly and thus a correction term is no longer needed. Thus, comparing revenue under $\hat{F}_B$ and F_B amounts to showing that

$$\int_{v^+}^{\hat{v}^+} \hat{q}_B(v)v\hat{x}_B'(v)dv \geq 0, \tag{71}$$

which always holds. Note that we keep here $\hat{x}_B'(v)$ to emphasise that this is not the allocation probability $x_B(v)$, but rather the allocation probability at the q associated with v_B, which differs from $\hat{v}_B$. ■

Proof of Proposition 4: Reallocating Value, Fixed Mean To begin, we characterize the optimal designs when $\mathbb{E}_{F_i}[v] = k_i$ for some $i \in \{A, B\}$ and $\mathbb{E}_{F_j}[v] = k - k_i$ for $j \neq i$. We will then let the designer choose $k_i \in [\underline{v}, k - \underline{v}]$ to maximize revenue.

As shown in the first part of proof of result 1, the surplus maximizing design F_i

given F_j must solve

$$\max_{F_i} \int_{\underline{v}}^{\overline{v}} S(v_i) dF_A(v_i) \quad \text{subject to} \quad \int_{\underline{v}}^{\overline{v}} v_i dF_i(v_i) = k_i. \tag{72}$$

As before, because the objective function of the problem is convex and increasing in v_i while the constraint is linear, there always exists a surplus-maximizing design which is maximally spread and has mean k_i – meaning that setting $F_i = F_{k_i}^{\overline{v}}$ maximizes surplus for any F_j. In particular, surplus will be maximized by any design (F_A, F_B) satisfying $F_i = F_{k_i}^{\overline{v}}$ for some $i \in \{A, B\}$ and $\int_{\underline{v}}^{\overline{v}} v_j dF_j(v_j) = k - k_i$ for $j \neq i$. For any such design, surplus would amount to

$$S^*(k_i) = \frac{k_i - \underline{v}}{\overline{v} - \underline{v}} \overline{v} + \frac{\overline{v} - k_i}{\overline{v} - \underline{v}} (k - k_i). \tag{73}$$

For such designs, surplus $S^*(k_i)$ is convex in k_i, it is decreasing for $k_i < k/2$ and increasing thereafter. Maximizing $S^*(k_i)$ with respect to k_i then establishes that in any surplus-maximizing design k_i is either equal to $\underline{v}$ or to $k - \underline{v}$. For both of these values of k_i, surplus is maximized and equals $S^+ = k - \underline{v}$.

The latter observation essentially completes the proof as for any of the designs in the statement of the result $R(F_A^*, F_B^*) = S^+$, given that the principal either trades with j at price h or sells to i at price $\underline{v}$. ■

Proof of Proposition 5: Reallocating Value, Fixed Distribution We want to show that revenue is maximized by selecting the maximally divisive design

$$F_B^*(v) = H(v) \qquad \text{if } v \in [\underline{v}, v^M) \tag{74}$$

$$F_A^*(v) = H(v) - 1 \qquad \text{if } v \in [v^M, \overline{v}]. \tag{75}$$

To simplify notation, we drop the star to indicate the optimal adjustment and simply refer to these two distributions by $F_A(v)$ and $F_B(v)$. We proceed as follows:
(i) We first assume that each value v admits a density and show that $F_A(v)$ and $F_B(v)$ yield higher transfer than any other two distributions, $\hat{F}_A(v)$ and $\hat{F}_B(v)$, that allocate strictly positive density at each $v \in [\underline{v}, \overline{v}]$ – meaning that $\min\{\hat{f}_A(v), \hat{f}_B(v)\} > 0$ for

all v.

(ii) We prove that $F_A(v)$ and $F_B(v)$ lead to higher revenue than any other two distributions with disjoint support. This allows to account for mass points.

(iii) We combine these insights to show that $F_A(v)$ and $F_B(v)$ yield a higher revenue than any other two distributions.

Splitting Densities We want to compare revenue with the maximally divisive design to revenue with split densities, where $\min\{\hat{f}_A(v), \hat{f}_B(v)\} > 0$ for all $v \in [\underline{v}, \overline{v}]$. Denote by $a_i(v)$ the share of the density $h(v)$ assigned to agent $i \in \{A, B\}$. We want to show that

$$\int_{v^M}^{\overline{v}} \psi_A(v)x_A(v)h(v)dv + \int_{\underline{v}}^{v^M} \psi_B(v)x_B(v)h(v)dv \tag{76}$$

$$\geq \int_{\underline{v}}^{\overline{v}} \hat{\psi}_A(v)\hat{x}_A(v)a_A(v)h(v)dv + \int_{\underline{v}}^{\overline{v}} \hat{\psi}_B(v)\hat{x}_B(v)a_B(v)h(v)dv. \tag{77}$$

For convenience, define

$$\overline{x}_A(v) = \begin{cases} \max\{\hat{x}_A(v), \hat{x}_B(v)\} & \text{if } v \geq v^M \\ \min\{\hat{x}_A(v), \hat{x}_B(v)\} & \text{if } v < v^M \end{cases}, \tag{78}$$

$$\overline{x}_B(v) = \{\hat{x}_A(v), \hat{x}_B(v)\} \setminus \overline{x}_A(v). \tag{79}$$

Similarly, define for any $i, j \in \{A, B\}$ such that $i \neq j$

$$\overline{\psi}_i(v) = \begin{cases} \hat{\psi}_i(v) & \text{if } \hat{x}_i(v) = \overline{x}_i(v) \\ \hat{\psi}_j(v) & \text{if } \hat{x}_i(v) \neq \overline{x}_i(v) \end{cases}, \tag{80}$$

$$\overline{a}_i(v) = \begin{cases} a_i(v) & \text{if } \hat{x}_i(v) = \overline{x}_i(v) \\ a_j(v) & \text{if } \hat{x}_i(v) \neq \overline{x}_i(v) \end{cases}. \tag{81}$$

These definitions immediately imply that

$$\int_{\underline{v}}^{\overline{v}} \hat{\psi}_A(v)\hat{x}_A(v)a_A(v)h(v)dv + \int_{\underline{v}}^{\overline{v}} \hat{\psi}_B(v)\hat{x}_B(v)a_B(v)h(v)dv \tag{82}$$

$$= \int_{\underline{v}}^{\overline{v}} \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v)h(v)dv + \int_{\underline{v}}^{\overline{v}} \overline{\psi}_B(v)\overline{x}_B(v)\overline{a}_B(v)h(v)dv. \tag{83}$$

As in previous proofs, we again compare revenue for distributions F_A and F_B with virtual values $\psi_A(v)$,$\psi_B(v)$ and allocation probabilities $\overline{x}_A$, $\overline{x}_B$ to revenue for any other distribution. To do so, we need to establish that $\overline{x}_A$ and $\overline{x}_B$ satisfy interim feasibility for distributions F_A and F_B.

Interim Feasibility Interim feasibility is satisfied if and only if

$$\int_{\max\{\overline{v},v^M\}}^{\overline{v}} \overline{x}_A(v)h(v)dv + \int_{\min\{\overline{v},v^M\}}^{v^M} \overline{x}_B(v)h(v)dv \leq 1 - F_A(\overline{v})F_B(\overline{v}), \tag{84}$$

see Border (1991). First, let $\overline{v} > v^M$. Then, $F_B(\overline{v}) = 1$ and the problem simplifies to

$$\int_{\overline{v}}^{\overline{v}} \overline{x}_A(v)h(v)dv \leq 2 - H(v), \tag{85}$$

which trivially holds. Next, let $\overline{v} < v^M$. In this case,

$$\int_{v^m}^{\overline{v}} \overline{x}_A(v)h(v)dv + \int_{\overline{v}}^{v^M} \overline{x}_B(v)h(v)dv \leq 1 \tag{86}$$

as $F_A(\overline{v}) = 0$. Therefore, it suffices to show that

$$\int_{v^m}^{\overline{v}} \overline{x}_A(v)h(v)dv + \int_{\underline{v}}^{v^M} \overline{x}_B(v)h(v)dv \leq 1 \tag{87}$$

Inequality (87) corresponds to the following inequality in the quantile-probability space:

$$\int_q \overline{y}_A(q)dq + \int_q \overline{y}_B(q)dq \leq 1 \tag{88}$$

We can alternatively integrate over allocation probabilities, which transforms the inequality to

$$\overline{y}_A(1) + \int_{\overline{y}_A} q_A(y)dy + \overline{y}_B(1) + \int_{\overline{y}_B} q_B(y)dy \leq 1 \tag{89}$$

The allocation probability for A, $\overline{y}_A$, lies between 1 and the allocation probability at the median value $\overline{y}_A(1)$. Inequality (89) takes into account that $\overline{y}_A(1) > 0$. This is as if the allocation probability displays a jump, in which case the quantile does not change. Formally, if there is a jump at some q we define

$$y^- \equiv \lim_{\epsilon \to 0} y(q - \epsilon) \tag{90}$$

$$y^+ \equiv \lim_{\epsilon \to 0} y(q + \epsilon) \tag{91}$$

and for every $y \in [y^-, y^+]$, $q(y)$ remains constant. The probability of good for B at the median is denoted by $\overline{y}_B(0)$ and it goes down to $\overline{y}_B(1)$.

We can express any $q_A = \hat{q}_A + \hat{q}_B$ and $q_B = \hat{q}_A + \hat{q}_B - 1$. To see this note that $q_A(v) = 1 - F_A(v) = 2 - H(v)$, $q_B(v) = 1 - F_B(v) = 1 - H(v)$ and $\hat{q}_A(v) + \hat{q}_B(v) = 2 - \hat{F}_A(v) + \hat{F}_B(v) = 2 - H(v)$. Note that in general, it will not hold that $q_A(y) = \hat{q}_A(y) + \hat{q}_B(y)$. However, there always exists a y' such that $q_A(y) = \hat{q}_A(y) + \hat{q}_B(y')$. As we integrate over all y and thus all y', we omit the dependence on y'. We can therefore replace q_A and q_B by $\hat{q}_A$ and $\hat{q}_B$ as follows:

$$\overline{y}_A(1) + \int_{\overline{y}_A} \left(\hat{q}_A(y) + \hat{q}_B(y)\right) dy + \overline{y}_B(1) + \int_{\overline{y}_B} \left(\hat{q}_A(y) + \hat{q}_B(y) - 1\right) dy \le 1 \tag{92}$$

Note that $\int_{\overline{y}_B} dy = \overline{y}_B(0) - \overline{y}_B(1)$. This implies

$$\overline{y}_A(1) - \overline{y}_B(0) + 2\overline{y}_B(1) + \int_{\overline{y}_A} \left(\hat{q}_A(y) + \hat{q}_B(y)\right) dy + \int_{\overline{y}_B} \left(\hat{q}_A(y) + \hat{q}_B(y)\right) dy \le 1 \tag{93}$$

If $\overline{y}_A(1) = \overline{y}_B(0)$, that is there is no jump at the median value, then we can rewrite inequality (93) as

$$2\overline{y}_B(1) + \int_{\overline{y}_B(1)}^{\overline{y}_A(0)} \hat{q}_A(y)dy + \int_{\overline{y}_B(1)}^{\overline{y}_A(0)} \hat{q}_B(y)dy \le 1 \tag{94}$$

$$\Leftrightarrow \qquad \int_q \hat{y}_A(q)dq + \int_y \hat{y}_B(q)dq \le 1, \tag{95}$$

where the latter holds as $\hat{y}_A$ and $\hat{y}_B$ are the allocation probabilities for distributions $\hat{F}_A$ and $\hat{F}_B$. If $\overline{y}_A(1) > \overline{y}_B(0)$, that is there is a jump at the median value, inequality

(93) can be expressed as

$$\overline{y}_A(1) - \overline{y}_B(0) + 2\overline{y}_B(1) + \int_{\overline{y}_B(1)}^{\overline{y}_B(0)} \hat{q}_A(y)dy + \int_{\overline{y}_A(1)}^{\overline{y}_A(0)} \hat{q}_A(y)dy$$
$$+ \int_{\overline{y}_B(1)}^{\overline{y}_B(0)} \hat{q}_B(y)dy + \int_{\overline{y}_A(1)}^{\overline{y}_A(0)} \hat{q}_B(y)dy \leq 1 \tag{96}$$

Note that $\overline{y}_A(1) - \overline{y}_B(0) = \int_{\overline{y}_B(0)}^{\overline{y}_A(1)} q dy$, with $q = 1 = \hat{q}_A(y) + \hat{q}_B(y)$, for all $y \in [\overline{y}_B(0), \overline{y}_A(1)]$. We can then amend inequality (96) to

$$2\overline{y}_B(1) + \int_{\overline{y}_B(1)}^{\overline{y}_A(0)} \hat{q}_A(y)dy + \int_{\overline{y}_B(1)}^{\overline{y}_A(0)} \hat{q}_B(y)dy \leq 1 \tag{97}$$
$$\Leftrightarrow \quad \int_q \hat{y}_A(q)dq + \int_y \hat{y}_B(q)dq \leq 1, \tag{98}$$

where the latter holds once again as we fixed an interim feasible allocation for $\hat{F}_A$ and $\hat{F}_B$.

As $\overline{x}_A$, $\overline{x}_B$ satisfy interim feasibility, it is sufficient to show that

$$\int_{v^M}^{\overline{v}} \psi_A(v)\overline{x}_A(v)h(v)dv + \int_{\underline{v}}^{v^M} \psi_B(v)\overline{x}_B(v)h(v)dv \tag{99}$$
$$\geq \int_{\underline{v}}^{\overline{v}} \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v)h(v)dv + \int_{\underline{v}}^{\overline{v}} \overline{\psi}_B(v)\overline{x}_B(v)\overline{a}_B(v)h(v)dv. \tag{100}$$

This is equivalent to establishing that

$$\underbrace{\int_{v^M}^{\overline{v}} \left(\left(\psi_A(v) - \overline{\psi}_A(v)\overline{a}_A(v)\right)\overline{x}_A(v) - \overline{\psi}_B(v)\overline{x}_B(v)\overline{a}_B(v)\right)h(v)dv}_{\text{Part 1}} \tag{101}$$
$$+ \underbrace{\int_{\underline{v}}^{v^M} \left(\left(\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v)\right)h(v)dv}_{\text{Part 2}} \geq 0. \tag{102}$$

We first focus on the case of virtual values being weakly positive, both under the maximally divisive design and the competing design, for all values, and establish the inequality by signing the two parts in turn.

Part 1 Note that since $\overline{x}_A(v) \geq \overline{x}_B(v)$ when $v \geq v^M$, we have that

$$\int_{v^M}^{\overline{v}} \left(\left(\psi_A(v) - \overline{\psi}_A(v)\overline{a}_A(v)\right) \overline{x}_A(v) - \overline{\psi}_B(v)\overline{x}_B(v)\overline{a}_B(v)\right) h(v)dv \tag{103}$$

$$\geq \int_{v^M}^{\overline{v}} \left(\psi_A(v) - \overline{\psi}_A(v)\overline{a}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right) \overline{x}_A(v)h(v)dv = 0. \tag{104}$$

The right hand side of the last inequality is equal to zero because

$$\psi_A(v) - \overline{\psi}_A(v)\overline{a}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v) = \psi_A(v) - \hat{\psi}_A(v)\hat{a}_A(v) - \hat{\psi}_B(v)\hat{a}_B(v) \tag{105}$$

$$= v - \frac{1 - F_A(v)}{h(v)} - \hat{a}_A(v)\left(v - \frac{1 - \hat{F}_A(v)}{\hat{a}_A(v)h(v)}\right) - \hat{a}_B(v)\left(v - \frac{1 - \hat{F}_B(v)}{\hat{a}_B(v)h(v)}\right) \tag{106}$$

$$= \frac{1}{h(v)}\left(1 + F_A(v) - \hat{F}_A(v) - \hat{F}_B(v)\right) = 0, \tag{107}$$

where the final equality follows from $\hat{F}_A(v) + \hat{F}_B(v) = H(v)$ and $F_A(v) = H(v) - 1$. Therefore, the integral in Part 1 is necessarily non-negative as it is bounded below by zero. Moreover, the integral is strictly positive provided that $\overline{x}_A(v) \neq \overline{x}_B(v)$ for a positive measure of $v \geq v^M$.

Part 2 As in Part 1, note that since $\overline{x}_A(v) \leq \overline{x}_B(v)$ when $v < v^M$, we have that

$$\int_{\underline{v}}^{v^M} \left(\left(\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right) \overline{x}_B(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v)\right) h(v)dv \tag{108}$$

$$\geq \int_{\underline{v}}^{v^M} \left(\psi_B(v) - \overline{\psi}_A(v)\overline{a}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right) \overline{x}_B(v)h(v)dv = \int_{\underline{v}}^{v^M} \overline{x}_B(v)dv \geq 0. \tag{109}$$

As in the previous part, the equality in the previous expression follows because

$$\psi_B(v) - \overline{\psi}_A(v)\overline{a}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v) = \frac{1}{h(v)}\left(1 + F_B(v) - \hat{F}_A(v) - \hat{F}_B(v)\right) = \frac{1}{h(v)}, \tag{110}$$

where the final equality follows from $\hat{F}_A(v) + \hat{F}_B(v) = H(v)$ and $F_B(v) = H(v)$. Therefore, the integral in Part 2 is also non-negative as it is bounded below by zero. Moreover, the integral is strictly positive whenever either $\overline{x}_A(v) \neq \overline{x}_B(v)$ or

$\overline{x}_B(v) > 0$ for a positive measure of $v < v^M$.

The two parts establish that splitting densities is strictly worse than the maximally divisive design when $\overline{x}_B(v) > 0$ for some $v < v^M$, since $\int_{\underline{v}}^{v^M} \overline{x}_B(v)dv > 0$ given that $\overline{x}_B(v)$ is increasing by incentive compatibility. However, splitting densities is strictly worse than the maximally divisive design even when $\overline{x}_B(v) = 0$ for all $v < v^M$, because the optimal allocation under the maximally divisive design must differ from $\overline{x}$ given that agent B must be promoted with positive probability when v_B is smaller but close v^M – meaning that for such values v_B we have that $x_B(v_B) > \overline{x}_B(v_B) = 0$.

So far we assumed that all virtual values are weakly positive. We relax this assumption and show that even with negative virtual values our result continues to hold.

Case 1 $\psi_A(v), \psi_B(v) > 0$ for all v, $\overline{\psi}_i(v) \geq 0$ for all $v \geq v^M$, $\overline{\psi}_i(v) < 0$ for some $v < v^M$

First, note that $\psi_B(v) > \overline{\psi}_i(v)$ for all $v < v^M$. In this case, Part 1 remains unchanged, while Part 2 needs to be amended. If both $\overline{\psi}_A(v), \overline{\psi}_B(v) < 0$ for some v, the difference in virtual values for each v is trivially positive. If $\overline{\psi}_B(v) > 0 > \overline{\psi}_A(v)$ for some v, then for these values

$$\left(\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v) > 0, \tag{111}$$

as $\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v) > 0$ and $-\overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v) \geq 0$. If for some v, $\overline{\psi}_A(v) > 0 > \overline{\psi}_B(v)$

$$\left(\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v) \tag{112}$$

$$> \left(\psi_B(v) - \overline{\psi}_B(v)\overline{a}_B(v) - \overline{\psi}_A(v)\overline{a}_A(v)\right)\overline{x}_B(v) > 0, \tag{113}$$

where the last inequality holds as $\psi_B(v) > \overline{\psi}_A(v)\overline{a}_A(v)$.

Case 2 $\psi_A(v), \overline{\psi}_i(v) > 0$ for $v \geq v^M$, $\psi_B(v) < 0$ for some $v < v^M$

If the virtual value is negative under the maximally divisive design for some v, it holds that $0 > \psi_B(v) > \overline{\psi}_i(v)$. Note that the allocation probability must not necessarily be zero, as regularity is not assumed. Now instead of assigning $\overline{x}_B$

to $\psi_B(v)$, assign $\overline{x}_A$ for all values for which $\psi_B(v) < 0$. Given that $\overline{x}_A(v) < \overline{x}_B$ interim feasibility continues to hold. Then, for these v

$$\psi_B(v)\overline{x}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v) \tag{114}$$
$$\geq \psi_B(v)\overline{x}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v)\overline{x}_A(v) - \overline{\psi}_A(v)\overline{x}_A(v)\overline{a}_A(v), \tag{115}$$

which, by the same logic as in Part 2 is positive.

Case 3 $\psi_A(v) \geq 0$ for all $v \geq v^M$, but $\overline{\psi}_i(v) < 0$ for some $v \geq v^M$. Suppose first that both $\overline{\psi}_A(v), \overline{\psi}_B(v) < 0$. In this case, Part 1 is trivially fulfilled. Assume next that there exists some v such that $\overline{\psi}_A(v) < 0 < \overline{\psi}_B(v)$. In this case, we construct a profitable deviation, that coincides with $\overline{F}_A, \overline{F}_B$ for all values below some threshold $\underline{v}$ and with the maximally divisive design above $\underline{v}$. We proceed with this approach until the candidate for the profitable deviation does not contain negative virtual values above the median anymore, in which case either Case 1 or Case 2 apply.

To construct such a deviation, note that $\overline{\psi}_A(\overline{v}) = \overline{\psi}_B(\overline{v}) = \overline{v}$, by assumption. This implies that there exists a $\overline{v}$, such that for all $v > \overline{v}$, $\overline{\psi}_A(v), \overline{\psi}_B(v) > 0$. The mass between $\overline{v}$ and $\overline{v}$ is given by $m = 2 - H(\overline{v}) > 0$. Construct another cutoff $\underline{v}$ such that $H(\overline{v}) - H(\underline{v}) = m$. Now consider a distribution $\tilde{F}_A(v), \tilde{F}_B(v)$ which corresponds to $\overline{F}_A(v), \overline{F}_B(v)$ for $v < \underline{v}$ and

$$\tilde{F}_A(v) = \begin{cases} \overline{F}_A(v) & \forall \quad v \leq \underline{v} \\ \overline{F}_A(\underline{v}) & \forall \quad \underline{v} < v \leq \overline{v} \\ H(v) - \overline{F}_A(\underline{v}) & \forall \quad \overline{v} < v, \end{cases} \tag{116}$$

$$\tilde{F}_B(v) = \begin{cases} \overline{F}_B(v) & \forall \quad v \leq \underline{v} \\ H(v) - \overline{F}_B(\underline{v}) & \forall \quad \underline{v} < v \leq \overline{v} \\ 1 & \forall \quad \overline{v} < v, \end{cases} \tag{117}$$

As $\tilde{F}$ and $\overline{F}$ coincide for values below $\underline{v}$, we focus on $v > \underline{v}$. Applying the same approach as in Part 1, for $v > \overline{v}$ it must hold that

$$1 + \tilde{F}_A(v) - \hat{F}_A(v) - \hat{F}_B(v) = 1 + H(v) - \overline{F}_A(\underline{v}) - H(v) > 0. \tag{118}$$

Part 2 can be similarly amended. We have now a new candidate for a profitable deviation $\tilde{F}$. Repeating the same steps if virtual values are negative for some values $v > v^M$, $\tilde{\psi}_i(v) < 0$ leads again to the maximally divisive design being optimal.

Case 4 Last, suppose that $\psi_A(v) < 0$ for some $v \geq v^M$. In this case, $\overline{\psi}_i(v) < 0$ for all i. Assign $\overline{x}_B(v)$ to F_A. Then, we obtain

$$\psi_A(v)\overline{x}_B(v) - \overline{\psi}_A(v)\overline{a}_A(v)\overline{x}_A(v) - \overline{\psi}_B(v)\overline{x}_B(v)\overline{a}_B(v) \tag{119}$$

$$> \left(\psi_A(v) - \overline{\psi}_A(v)\overline{a}_A(v) - \overline{\psi}_B(v)\overline{a}_B(v)\right)\overline{x}_B(v) = 0, \tag{120}$$

as before.

Disjoint Support This proof allows for distributions $H(v)$ for which densities are not defined. If the density is defined, we assign each density to one distribution. This implies that the support of at least one distribution is disjoint. We restrict attention to alternative distributions with $\hat{v}_B \in [\underline{v}, s_1] \cup [s_2, s_3]$ and $\hat{v}_A \in [s_1, s_2] \cup [s_3, \overline{v}]$, with $\underline{v} < s_1 < s_2 < s_3 \leq \overline{v}$. If these distributions do not yield higher revenue than the maximally divisive design, then splitting the distribution further cannot be optimal either. Note that $s_1 < v^M < s_3$.

Recall that revenue in the quantile space if given by

$$\int_q \phi_A(q)y_A(q)dq + \int_q \phi_B(q)y_B(q)dq \tag{121}$$

Integrating each component by parts yields

$$v_i(q)qy_i(q)\Big|_0^1 + \int_q v_i(q)\left(-y_i'(q)\right)qdq \tag{122}$$

$$= v_i(1)y_i(1) + \int_C v_i(q)\left(-y_i'(q)\right)qdq + \sum_{q\in D} q\left[v^-(q)y^-(q) - v^+(q)y^+(q)\right] \tag{123}$$

Flipping this to a $v-q-$ space yields

$$\underline{v}_A x_A(\underline{v}_A) + \int_{V_A \setminus D} v q_A(v) x'_A(v) dv \tag{124}$$

$$+\underline{v}_B x_B(\underline{v}_B) + \int_{V_B \setminus D} v q_B(v) x'_B(v) dv + \sum_{v \in D} q_{A\setminus B}(v) v \left[x^+_{A\setminus B}(v) - x^-_{A\setminus B}(v) \right], \tag{125}$$

which allows us once more to directly account for gaps in the support. We use $A \setminus B$ to signify that the variable may either belong to distribution A or B.

We consider three distinct cases:

1. both distributions with disjoint support, cutoff s_2 below v^M, $s_2 \le v^M$
2. both distributions with disjoint support, cutoff s_2 above v^M, $s_2 > v^M$
3. one distribution with disjoint support: $\hat{v}_B \in [\underline{v}, s_1] \cup [s_2, \overline{v}]$ and $\hat{v}_A \in [s_1, s_2]$

We relabel $\hat{x}_i(v)$, $\hat{q}_i(v)$ as $\tilde{x}_A(v)$, $\tilde{q}_A(v)$ if $v > v^M$ and as $\tilde{x}_B(v)$, $\tilde{q}_B(v)$ if $v < v^M$. Then, it is sufficient to show that

$$v^M \tilde{x}_A(v^M) + \int_{v^M}^{\overline{v}} q_A(v) \tilde{x}'_A(v) v dv + \int_{\underline{v}}^{v^M} q_B(v) \tilde{x}'_B(v) v dv + \sum_{v \in D} q_{A\setminus B}(v) v \left[x^+_{A\setminus B}(v) - x^-_{A\setminus B}(v) \right], \tag{126}$$

$$> s_1 \tilde{x}_B(s_1) + \int_{v^M}^{\overline{v}} \tilde{q}_A(v) \tilde{x}'_A(v) v dv + \int_{\underline{v}}^{v^M} \tilde{q}_B(v) \tilde{x}'_B(v) v dv + \sum_{v \in D} \tilde{q}(v) v \left[\tilde{x}^+_{A\setminus B}(v) - \tilde{x}^-_{A\setminus B}(v) \right]. \tag{127}$$

Disjoint Support for $\hat{F}_A, \hat{F}_B$: $s_2 \le v^M$ In this case inequality (126) can be amended to

$$v^M \tilde{x}_A(v^M) + \int_{v^M}^{s_3} q_A(v) \tilde{x}'_A(v) v dv + \int_{s_1}^{v^M} q_B(v) \tilde{x}'_B(v) v dv + \sum_{v \in D} q_{A\setminus B}(v) v \left[x^+_{A\setminus B}(v) - x^-_{A\setminus B}(v) \right], \tag{128}$$

$$> s_1 \tilde{x}_B(s_1) + \int_{v^M}^{s_3} \tilde{q}_A(v) \tilde{x}'_A(v) v dv + \int_{s_1}^{v^M} \tilde{q}_B(v) \tilde{x}'_B(v) v dv + \sum_{v \in D} \tilde{q}_{A\setminus B}(v) v \left[\tilde{x}^+_{A\setminus B}(v) - \tilde{x}^-_{A\setminus B}(v) \right], \tag{129}$$

as the distributions are identical below s_1 and above s_3. Note that $\tilde{q}_A(v) = \hat{q}_B(v)$

for $v \in [v^M, s_3]$. Taking the difference yields

$$\int_{v^M}^{s_3} \left(q_A(v) - \hat{q}_B(v)\right) \tilde{x}'_A(v) v dv \tag{130}$$

where

$$q_A(v) = 2 - H(v) \tag{131}$$

while

$$\hat{q}_B(v) = 1 - \hat{F}_B = 1 - (H(v) - H(s_2) + H(s_1)) \tag{132}$$

Then, for a given $v \in [v^M, s_3]$

$$q_A(v) - \hat{q}_B(v) = 1 - H(v) + (H(v) - H(s_2) + H(s_1)) = 1 - H(s_2) + H(s_1) > 0 \tag{133}$$

implying that the difference is constant. Expression (130) then becomes:

$$(1 - H(s_2) + H(s_1)) \int_{v^M}^{s_3} \tilde{x}'_A(v) v dv \tag{134}$$

We then turn to

$$\int_{s_1}^{v^M} \left(q_B(v) - \tilde{q}_B(v)\right) \tilde{x}'_B(v) v dv \tag{135}$$

Between s_2 and v^M, $\tilde{q}_B(v) = \hat{q}_B(v)$:

$$q_B(v) - \hat{q}_B(v) = 1 - H(v) - (1 - (H(v) - H(s_2) + H(s_1))) \tag{136}$$

$$= 1 - H(v) - 1 + (H(v) - H(s_2) + H(s_1)) = -\left(H(s_2) - H(s_1)\right) < 0 \tag{137}$$

which is again constant. For $v \in [s_1, s_2]$, this difference is given by

$$q_B(v) - \hat{q}_A(v) = 1 - H(v) - (1 - (H(v) - H(s_1))) \tag{138}$$

$$= -H(v) + (H(v) - H(s_1)) = -H(s_1) \tag{139}$$

Expression (135) is then equivalent to

$$\int_{s_1}^{v^M} (q_B(v) - \tilde{q}_B(v))\, \tilde{x}'_B(v) v dv = -(H(s_2) - H(s_1)) \int_{s_2}^{v^M} \tilde{x}'_B(v) v dv - H(s_1) \int_{s_1}^{s_2} \tilde{x}'_B(v) v dv \tag{140}$$

Collecting all terms leads to the following comparison

$$v^M \tilde{x}_A(v^M) + (1 - H(s_2) + H(s_1)) \int_{v^M}^{s_3} \tilde{x}'_A(v) v dv \tag{141}$$

$$-(H(s_2) - H(s_1)) \int_{s_2}^{v^M} \tilde{x}'_B(v) v dv - H(s_1) \int_{s_1}^{s_2} \tilde{x}'_B(v) v dv \tag{142}$$

$$+(1 - H(s_2) + H(s_1)) \sum_{v \in D \cap [v^M, s_3]} v \left[\tilde{x}^+_A(v) - \tilde{x}^-_A(v)\right] \tag{143}$$

$$-(H(s_2) - H(s_1)) \sum_{v \in D \cap [s_2, v^M]} v \left[\tilde{x}^+_B(v) - \tilde{x}^-_B(v)\right] \tag{144}$$

$$-H(s_1) \sum_{v \in D \cap [s_1, s_2]} v \left[\tilde{x}^+_B(v) - \tilde{x}^-_B(v)\right] > s_1 \tilde{x}_B(s_1) \tag{145}$$

Rearranging yields

$$v^M \tilde{x}_A(v^M) + \int_{v^M}^{s_3} \tilde{x}'_A(v) v dv - (H(s_2) - H(s_1)) \int_{s_2}^{s_3} \tilde{x}'_{A \setminus B}(v) v dv - H(s_1) \int_{s_1}^{s_2} \tilde{x}'_B(v) v dv \tag{146}$$

$$+ \sum_{v \in D \cap [v^M, s_3]} v \left[\tilde{x}^+_A(v) - \tilde{x}^-_A(v)\right] - (H(s_2) - H(s_1)) \sum_{v \in D \cap [s_2, s_3]} v \left[\tilde{x}^+_{A \setminus B}(v) - \tilde{x}^-_{A \setminus B}(v)\right] \tag{147}$$

$$-H(s_1) \sum_{v \in D \cap [s_1, s_2]} v \left[\tilde{x}^+_B(v) - \tilde{x}^-_B(v)\right] > s_1 \tilde{x}_B(s_1) \tag{148}$$

As $H(s_1), H(s_2) - H(s_1) < 1$, it suffices to show that

$$s_3\tilde{x}_A(s_3) - \int_{v^M}^{s_3} \tilde{x}_A(v)dv - (s_3\tilde{x}_A(s_3) - s_1\tilde{x}_B(s_1)) + \int_{s_1}^{s_3} \tilde{x}(v)dv > \tilde{x}_B(s_1)s_1 \quad (149)$$

$$\Leftrightarrow \quad \int_{s_1}^{v^M} \tilde{x}_B(v)dv > 0, \quad (150)$$

which always holds and establishes that any other disjoint distributions do not yield a higher revenue. As we integrate over allocation probabilities, potential discontinuities are then once again subsumed.

Disjoint Support for $\hat{F}_A, \hat{F}_B$: $s_2 > v^M$ By the same logic as in the previous case, it is sufficient to show

$$v^M\tilde{x}_A(v^M) + \int_{v^M}^{s_3} q_A(v)\tilde{x}'_A(v)vdv + \int_{s_1}^{v^M} q_B(v)\tilde{x}'_B(v)vdv \quad (151)$$

$$> s_1\tilde{x}_B(s_1) + \int_{v^M}^{s_3} \tilde{q}_A(v)\tilde{x}'_A(v)vdv + \int_{s_1}^{v^M} \tilde{q}_B(v)\tilde{x}'_B(v)vdv \quad (152)$$

We can ignore any potential discontinuities in allocation probabilities here, as we integrate at the end, meaning the discontinuities will vanish once again. As before, with a different integration bound below

$$\int_{s_2}^{s_3} q_A(v) - \hat{q}_B(v)\tilde{x}'_A(v)vdv = (1 - H(s_2) + H(s_1)) \int_{s_2}^{s_3} \tilde{x}'_A(v)vdv \quad (153)$$

We turn to

$$\int_{v^M}^{s_2} (q_A(v) - \hat{q}_A(v))\, \tilde{x}'_A(v)vdv \quad (154)$$

$$q_A(v) - \hat{q}_A(v) = 2 - H(v) - (1 - (H(v) - H(s_1))) \quad (155)$$

$$= 1 - H(v) + (H(v) - H(s_1)) = 1 - H(s_1) \quad (156)$$

Then,

$$\int_{v^M}^{s_2} (q_A(v) - \hat{q}_A(v))\, \tilde{x}'_A(v)vdv = (1 - H(s_1)) \int_{v^M}^{s_2} \tilde{x}'_A(v)vdv \quad (157)$$

Last,

$$\int_{s_1}^{v^M} \left(q_B(v) - \hat{q}_A(v)\right) \tilde{x}'_A(v) v dv = -H(s_1) \int_{s_1}^{v^M} \tilde{x}'_B(v) v dv \tag{158}$$

Collecting terms yields

$$v^M \tilde{x}_A(v^M) + (1 - H(s_2) + H(s_1)) \int_{s_2}^{s_3} \tilde{x}'_A(v) v dv + (1 - H(s_1)) \int_{v^M}^{s_2} \tilde{x}'_A(v) v dv \tag{159}$$

$$- H(s_1) \int_{s_1}^{v^M} \tilde{x}'_B(v) v dv > s_1 \tilde{x}_B(s_1) \tag{160}$$

As before it is sufficient to show that

$$s_3 \tilde{x}_A(s_3) - \int_{v^M}^{s_3} \tilde{x}_A(v) dv - (s_3 \tilde{x}_A(s_3) - s_1 \tilde{x}_B(s_1)) + \int_{s_1}^{s_3} \tilde{x}_{A \setminus B}(v) dv > \tilde{x}_B(s_1) s_1 \tag{161}$$

$$\Leftrightarrow \qquad \int_{s_1}^{v^M} \tilde{x}_B(v) dv > 0 \tag{162}$$

This establishes that also in this case, the maximally divisive design yields the highest transfer.

Disjoint Support for $\hat{F}_B$: $s_2 > v^M$ Similar to the previous case, we can ignore discontinuities in the allocation probabilities as they will once again vanish in the end. It is sufficient to show

$$v^M \tilde{x}_A(v^M) + \int_{v^M}^{s_2} q_A(v) \tilde{x}'_A(v) v dv + \int_{s_1}^{v^M} q_B(v) \tilde{x}'_B(v) v dv \tag{163}$$

$$> s_1 \tilde{x}_B(s_1) + \int_{v^M}^{s_2} \tilde{q}_A(v) \tilde{x}'_A(v) v dv + \int_{s_1}^{v^M} \tilde{q}_B(v) \tilde{x}'_B(v) v dv \tag{164}$$

As before,

$$\int_{v^M}^{s_2} q_A(v) - \hat{q}_A(v) \tilde{x}'_A(v) v dv = (1 - H(s_1)) \int_{v^M}^{s_2} \tilde{x}'_A(v) v dv \tag{165}$$

$$\int_{s_1}^{v^M} q_B(v) - \hat{q}_A(v) \tilde{x}'_A(v) v dv = -H(s_1) \int_{s_1}^{v^M} \tilde{x}'_A(v) v dv \tag{166}$$

Following the same steps as before establishes that the maximally divisive design yields higher revenue than this candidate.

Mixing Split Densities and Disjoint Supports It follows that there cannot be any other design combining mixing split densities and disjoint supports that yields higher revenue compared to the maximally divisive design. We can always treat subsets of the distribution as the entire distribution. This just requires an adjustment of the mass in a certain subset. Then we can perform the same analysis as we did for a subset, which yields lower revenue for this subset by the same logic as above.

Therefore, we have established that maximally divisive designs is optimal and yields highest revenue among all possible distributions.■

References

Bergemann, D., B. Brooks, and S. Morris (2015). The limits of price discrimination. *American Economic Review 5*(105), 921–957.

Bergemann, D. and M. Pesendorfer (2007). Information structures in optimal auctions. *Journal of economic theory 137*(1), 580–609.

Bobkova, N. (2019). Information choice in auctions. Technical report, Working paper.

Border, K. C. (1991). Implementation of reduced form auctions: A geometric approach. *Econometrica: Journal of the Econometric Society*, 1175–1187.

Bulow, J. and J. Roberts (1989). The simple economics of optimal auctions. *The Journal of Political Economy 97, No. 5*, 1060–1090.

Cantillon, E. (2008). The effect of bidders' asymmetries on expected revenue in auctions. *Games and Economic Behavior 62*(1), 1–25.

Condorelli, D. and B. Szentes (2020). Information design in the holdup problem. *Journal of Political Economy 128*(2), 681–709.

Condorelli, D. and B. Szentes (2022). Surplus sharing in cournot oligopoly. *Theoretical Economics 3*(17), 955–975.

Crémer, J. and R. P. McLean (1985). Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent. *Econometrica 53*, 345–361.

Deb, R. and M. M. Pai (2017). Discrimination via symmetric auctions. *American Economic Journal: Microeconomics 9*(1), 275–314.

Eso, P. and B. Szentes (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies 74*(3), 705–731.

Ganuza, J.-J. and J. Penalva (2019). Information disclosure in optimal auctions. *International Journal of Industrial Organization March*(63), 460–479.

Hart, S. and P. J. Reny (2015). Maximal revenue with multiple goods: nonmonotonicity and other observations. *Theoretical Economics 10*(3), 893–922.

Hartline, J. D. (2013). Mechanism design and approximation. *Book draft. October 122.*

Johnson, J. P. and D. P. Myatt (2006). On the simple economics of advertising, marketing, and product design. *American Economic Review 96*(3), 756–784.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Kleiner, A., B. Moldovanu, and P. Strack (2021). Extreme points and majorization: Economic applications. *Econometrica 4*(89), 1557–1593.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research 6*(1), 58–73.

Roesler, A.-K. and B. Szentes (2017). Buyer-optimal learning and monopoly pricing. *American Economic Review 107*(7), 2072–80.

Sorokin, C. and E. Winter (2018). Pure information design in classical auctions. Technical report, Working paper.