Screening with damages and ordeals*

Filip Tokarski
Stanford GSB

August 5, 2025

Abstract

A welfare-maximizing designer allocates two kinds of goods using two wasteful screening instruments: *ordeals*, which enter agents' utilities additively, and *damages*, which harm agents in proportion to their values for the goods. If agents have common valuations for one of the goods, damages always lead to Pareto-dominated mechanisms: any allocation using damages can also be implemented with ordeals alone, while also leaving greater rents to inframarginal types. However, using damages can be optimal when agents' valuations for both goods are heterogeneous: with multidimensional types, the two devices differ in how they sort agents into available options, with the optimal sorting sometimes requiring the use of damages. I nevertheless identify distributional conditions under which using damages is not optimal. In those cases, the optimal mechanism produces an efficient allocation by posting "market-clearing" ordeals for each type of good.

1 Introduction

Affordable housing programs offer units that vary in location and size, with households' preferences over them having a strong horizontal component (Waldinger, 2021). These apartments are often allocated via waitlist, where wait-times differ substantially between developments. Participating households therefore trade off their values for specific units against expected wait-time (Van Ommeren and Van der Vlist, 2016).[1] For example, applicants seeking larger or more centrally-located apartments must often endure waits that are years longer than those for less popular ones (Van Dijk, 2019). Thus, in public housing programs, wait-times largely assume the role of prices, balancing supply and demand and "sorting" participants into the available units. However, wait-times have a curious screening property: they are *more costly* to

*I am grateful to Joey Feffer, Federico Llarena, Sam Wycherley, Benjamin Brooks, Rafael Berriel, Andrzej Skrzypacz, Piotr Dworczak, Jacob Leshno, Michael Ostrovsky, Ilya Segal, and Rebecca Diamond, Zi Yang Kang for their helpful comments and suggestions.

[1] While affordable housing programs around the world differ in terms of the exact mechanism or its implementation, the same core trade-off is often present.

households whose values for the allocated units are higher: indeed, each period during which a household waits deprives it of the flow value of the apartment. Consequently, the cost of delaying receipt is *multiplicative* with one's value for the good. This is in contrast to other screening devices, such as differences in rent subsidies between apartments or bureaucratic hassles, whose disutilities are *separable* from the value of the allocated good.

In this paper, I compare the properties of these two kinds of screening instruments. I call the former class of instruments, whose costs to agents are proportional to their values for the good, *damages*. I refer to the latter class, whose costs are separable from agents' values, as *ordeals*. Both classes of instruments are common in practice. A prominent example of damages are delays in allocation.[2] Apart from the public housing context, screening through delays is also common in healthcare. Since the supply of different treatments is limited, patients requesting scarce or costly therapies are often mandated to secure referrals or demonstrate that less invasive methods have been attempted previously. Indeed, the World Health Organization (2023) states that the referral system is "aiming to ensure patient access to specialist health care when needed, while maintaining resource efficiency". While referral requirements may reduce the burden on the healthcare system by ensuring that only patients with a high need pursue scarce treatment options, they often also delay the therapy itself. For instance, the American Physical Therapy Association claims that "direct access restrictions cause unnecessary delays for people who would benefit from treatment by a PT."[3] Indeed, delays in treatment are more burdensome to more severely injured patients whose value for the scarce procedure is the largest. Other kinds of damages include choice and usage restrictions, and reductions in good quality (Deneckere and McAfee, 1996; Arnosti and Shi, 2020; Waldinger, 2021).

It is worth noting that the screening properties of delays (and thus waitlists) are different from those of *queues*, à la Nichols et al. (1971), where agents waste time standing in line. Since queueing does not affect the value derived from the good, under my definition, it is an *ordeal*. Other examples of wasteful screening devices that are ordeals include form-filling and bureaucratic processes, work requirements, or traveling to a distant office (Nichols and Zeckhauser, 1982; Besley and Coate, 1992; Kleven and Kopczuk, 2011; Dupas et al., 2016; Alatas et al., 2016; Deshpande and Li, 2019).

I study a mechanism design setting where a welfare-maximizing designer allocates fixed supplies of goods to agents. She chooses deterministic mechanisms that combine two wasteful screening instruments: ordeals and damages to the allocated goods. I first consider a setting where agents differ in their values for a single type of scarce good and have common values for an unlimited outside option. I show that in such cases, damages always lead to Pareto-dominated mechanisms: damaging the good is less efficient at screening out low-value agents, as it is *more costly* to high-value inframarginal recipients than it is to the marginal recipient. Since ordeals are equally costly to everyone, they could be used to produce the same allocation

[2] For flow goods that can be enjoyed in every period after receipt, delays deprive recipients of periods of usage. For consumable goods, they lead to temporal discounting of the good's value.

[3] https://www.apta.org/advocacy/issues/direct-access-advocacy

while preserving more rents for inframarginal recipients.

I then consider settings where the designer allocates two kinds of goods to agents with heterogeneous values for both of them. I show that in such settings, the designer *could* benefit from damaging one of the goods. Intuitively, in the previous case, any possible way of "sorting" agents into the two options could have been implemented using only ordeals. Moreover, implementing any sorting pattern this way was always optimal. However, this is no longer the case when types are multidimensional: many sorting patterns can now only be implemented by combining both instruments. Nevertheless, I identify distributional conditions under which using damages is not optimal. In those cases, the welfare-maximizing mechanism posts "market-clearing" ordeals for each type of the good, allocating the whole supply efficiently.

The core of my paper studies a stylized model illustrating the properties of the two classes of screening instruments. However, I subsequently consider extensions which show that my insights and methods survive when other practically important effects are also considered. I first relax the assumption that both screening devices are fully wasteful. This may, for instance, capture the case where ordeals are interpreted as monetary payments or reductions in subsidies, which generate revenue for the designer. In Section 5, I therefore extend my analysis to the case where ordeals are only *partially* wasteful, and show this does not affect my main result. Second, I consider the case where agents differ in how costly they find the ordeal. For instance, when screening is done with monetary payments, poorer agents whom the program tries to target may find them more burdensome. On the contrary, the same agents may be more willing to wait, travel or endure other inconveniences in order to get the good (Dupas et al., 2016). I show that my solution method directly extends to the case of heterogeneous ordeal costs; I provide modified statistical conditions under which my main result holds in this setting. Lastly, in Appendix A, I draw a connection between my baseline model and waitlist mechanisms described in the introduction. Indeed, in this environment wait-times are not a direct design choice but arise naturally as byproducts of the system's equilibrium dynamics. Nonetheless, I provide a microfoundation which shows that the steady state of such a system maps onto my baseline model. I also show how the restriction to deterministic mechanisms can be related to the assumption that the designer cannot ban agents from re-entering the mechanism.

I conclude the analysis by discussing its implications for public housing allocation. My results suggest that large imbalances in the lengths of different waitlists, which effectively correspond to screening using damages, may be undesirable. The policymaker should instead consider rebalancing the prices or subsidies between different housing developments in a way that brings their wait-times closer together.

From a technical perspective, my model is an instance of a tractable multidimensional screening problem. By restricting attention to deterministic mechanisms, I am able to characterize them as pairs of ordeal and damage menus for the two goods. This in turn allows me to represent two-dimensional mechanisms as (endogenously) interconnected single-dimensional screening problems. The interaction between them is summarized by a boundary in the type space that separates the sets of types who choose each good. The multidimensional problem can then be broken up into two stages: first, determining the optimal way to implement a given boundary, and second, solving an optimal control problem to select the best boundary among all implementable ones.

Apart from the aforementioned work on screening with ordeals in social programs, this paper relates closely to a more general literature on using costly screening and money-burning to maximize welfare (Hartline and Roughgarden, 2008; Condorelli, 2012). A related literature also studies the use of wasteful screening devices by a monopolist, with the seminal paper by Deneckere and McAfee (1996) establishing conditions under which the seller would want to damage goods. Yang (2021) considers a more general problem where a monopolist has access to both wasteful and non-wasteful instruments, and characterizes cases where the wasteful one should not be used. However, these literatures study the problem of allocating a single kind of good. By considering a model with heterogeneous goods, I explore an additional role of screening devices, where they are also used to "sort" agents into the available options. My comparison of the screening properties of damages and ordeals also relates to the work of Akbarpour et al. (2023) who ask when one screening device dominates another for a planner aiming to maximize a social welfare function. Unlike them, I allow the designer to *combine* instruments and show that, under certain distributional conditions, screening with ordeals alone dominates any mechanism using both devices. Finally, my paper relates to a literature on waitlist design. While no paper has studied combining waitlists with payments in settings with heterogeneous goods, a substantial literature examines designing such waitlists without transfers. Arnosti and Shi (2020) and Waldinger (2021) study the effects of restricting recipients' choice on targeting. Barzel (1974), Bloch and Cantala (2017), and Leshno (2022) observe that in environments with homogeneous waiting costs, wait-times may "act as prices", screening for agents with higher valuations. I refine this intuition by showing that the screening properties of wait-times are different when the cost of waiting stems from delayed receipt—in those cases, wait-times can only screen on agents' *relative* values for the offered goods.

2 Model

A designer distributes two types of goods, A and B. Their supplies are equal to $\mu_A, \mu_B > 0$. There is a unit mass of agents whose values for the two goods are given by a and b, respectively. Agents' values (a, b) are distributed according to F with a piece-wise Lipschitz continuous density f. The designer chooses a menu of qualities and ordeals for each of the goods. That is, an agent can choose which good she wants to get and then choose a quality and ordeal option from the relevant good's menu. She can also not participate, which gives her utility 0. When a type-(a, b) agent participates and receives good y, her utility is:

$$x \cdot a - c \;\text{ if }\; y = A,$$
$$x \cdot b - c \;\text{ if }\; y = B,$$

where $c \in \mathbb{R}_+$ is the ordeal the agent completes and $x \in [0,1]$ is her good's quality. Whenever $x < 1$, we say the good has been *damaged*. The designer chooses the menus to maximize the total welfare of agents. Thus, by the Revelation Principle, we can reduce her problem to picking allocation rules for ordeals, $c : [0,1]^2 \to \mathbb{R}_+$, qualities, $x : [0,1]^2 \to [0,1]$, and goods, $y : [0,1]^2 \to \{\varnothing, A, B\}$, to maximize welfare:

$$\int U[a,b,(c,x,y)(a,b)]\,\mathrm{d}F(a,b), \tag{W}$$

subject to (IC) and (IR) constraints, and the supply constraint (S):

$$\text{for all } (a,b),(a',b') \in [0,1]^2, \quad U[a,b,(c,x,y)(a,b)] \geq U[a,b,(c,x,y)(a',b')], \tag{IC}$$

$$\text{for all } (a,b) \in [0,1]^2, \quad U[a,b,(c,x,y)(a,b)] \geq 0, \tag{IR}$$

$$\int \mathbb{1}_{y(a,b)=A}\,\mathrm{d}F(a,b) \leq \mu_A, \quad \int \mathbb{1}_{y(a,b)=B}\,\mathrm{d}F(a,b) \leq \mu_B. \tag{S}$$

Here $U[a,b,(c,x,y)(a',b')]$ denotes the utility type (a,b) gets from reporting (a',b') in the mechanism (c,x,y). I call a mechanism (c,x,y) satisfying (IC),(IR), (S) *feasible*.

3 One good case

Let us first consider the case where only good A is scarce, and interpret good B as an unlimited outside option for which all agents have common values. I therefore assume $\mu_A < 1$ and $\mu_B = \infty$, and set $b > 0$ to be all agents' value for good B.

Proposition 1. *Any mechanism that uses damages, so features $x(a,b) < 1$, is Pareto-dominated by a mechanism that uses only ordeals, i.e. where $x(a,b) \equiv 1$.*

Proof. By single crossing, there is a cutoff $\underline{a}$ such that all and only agents with value a above it get good A:

$$\underline{a} = \inf\{a : \ y(a,b) = A \text{ or } a = 1\}.$$

Fix the value for the cutoff $\underline{a}$; I will show that among mechanisms with this cutoff, the one not using damages Pareto-dominates all the other ones. First, note that for the mechanism to satisfy (IC), all agents who do not receive A must be getting the same utility level $\underline{U}$. By the envelope theorem, we can then write an agent's indirect utility as:

$$U(a,b) = \underline{U} + \int_{\underline{a}}^{\max[\underline{a},a]} x(t,b)\,dt.$$

Now, notice that $\underline{U} \leq b$ and $x(a,b) \leq 1$, and thus $U(a,b)$ is at most $b + \max[\underline{a},a] - \underline{a}$. Indeed, this upper bound on the utility profile is implementable by offering a menu with two options: good B with $x = 1, c = 0$ and good A with $x = 1$ and $c = \underline{a}$. □

To understand the intuition behind this result, note that in the single-good case, good A will go to agents whose values a for it lie in some upper interval $[\underline{a}, 1]$ under any mechanism. The designer's problem therefore boils down to selecting a cutoff $\underline{a}$ and choosing how to enforce it. To do so, she has to deter some low-a agents from choosing good A. The burden imposed on the takers of A must then be chosen to make the cutoff type $\underline{a}$ exactly indifferent between A and the outside option. This can be done by imposing an ordeal on those choosing good A, or by damaging good A to make the outside option relatively more attractive. Note, however, that damages are more burdensome to *inframarginal types* than the types below $\underline{a}$ they are actually meant to deter. Ordeals, on the other hand, are equally burdensome to everyone. Thus, enforcing the cutoff $\underline{a}$ with ordeals leaves more rents to inframarginal takers of A (Figure 1).[4]

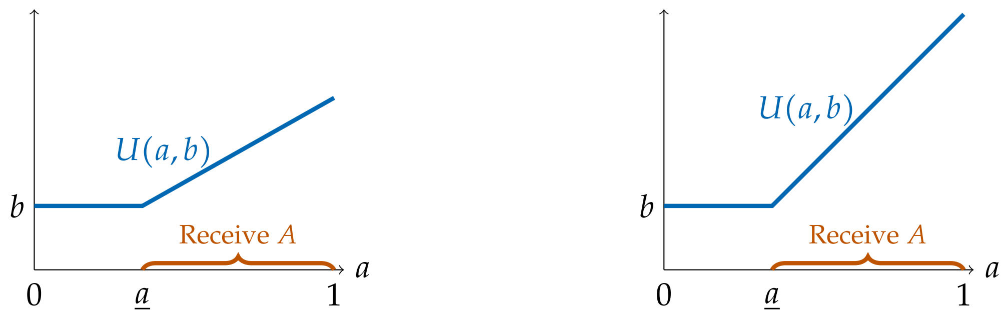

Figure 1: Indirect utilities $U(a,b) = \underline{U} + \int_{\underline{a}}^{\max[\underline{a},a]} x(t,b)\, dt$ for mechanisms enforcing the cutoff $\underline{a}$ with damages (left) and ordeals (right).

While Proposition 1 is simple, it provides economic insight. For example, it captures a key force in the model of congestion pricing of Vickrey (1973) and its subsequent generalization by Van Den Berg and Verhoef (2011). The model studies the use of congestion pricing to spread the amount of traffic passing through a capacity-constrained road. It shows that when drivers' values for time are identical, congestion pricing does not improve their utility. This is because the same allocation of driving times is implemented with or without tolls: the "prices" for traveling at specific times are then pinned down by market clearing, and it is irrelevant whether they are paid in money, through tolls, or in wasted time, through congestion. This conclusion is overturned, however, when agents' values for time are heterogeneous. Proposition 1 makes this clear: while tolls are an *ordeal*, waiting in traffic is a *damage*. When the road's capacity constraint at peak time is enforced through payments, everyone pays the same price as the marginal driver. When it is enforced through waiting in traffic, the marginal driver experiences the same disutility as she would from the payment, but the inframarginal drivers with the highest values for arriving early suffer strictly more.

[4]Proposition 1 was formulated in a linear model to maintain consistency with the two-good setup. However, its logic easily extends to the case of general screening instruments whose disutilities to agents are additive. We could then consider two wasteful screening instruments, where the cost of one is increasing more steeply in the value for a than the cost of the other. The result would then say that any mechanism using the former instrument is Pareto-dominated by one using only the latter instrument.

4 Two good case

I now consider the case where screening devices are used not only to exclude low-value agents, but also to direct recipients to choose types of goods in a socially-efficient way. I therefore consider the case where both goods are scarce, $\mu_A + \mu_B \le 1$, and F, the distribution of values (a, b), has full support on $[0,1]^2$. This two-dimensional heterogeneity makes the setting significantly richer. In contrast to the previous section, it does not feature a uniform ordering across agents that pins down the "sorting" into available goods. Indeed, as I explain below, the designer can now use damages and ordeals to sort agents into the two goods in rich ways.

4.1 Boundary structure of mechanisms

We can without loss restrict attention to mechanisms that allocate positive masses of both goods.[5] Let us then describe the properties of such mechanisms. It will be convenient to characterize them in terms of A, B-indirect utilities $U_A, U_B : [0,1] \to \mathbb{R}_+$, defined as follows:

$$U_A(a) = \max_{(a',b')} \left\{ x(a',b') \cdot a - c(a',b') : \quad y(a',b') \in \{A, \varnothing\} \right\},$$

$$U_B(b) = \max_{(a',b')} \left\{ x(a',b') \cdot b - c(a',b') : \quad y(a',b') \in \{B, \varnothing\} \right\}.$$

Intuitively, $U_A(a)$ and $U_B(b)$ represent the highest utility type (a, b) could get from selecting some quality and ordeal option for the A- and the B-goods, respectively (or not participating). Then agents for whom $U_A(a) > U_B(b)$ choose good A and those from whom $U_A(a) < U_B(b)$ choose good B. Note that U_A and U_B are convex and increasing, and that they depend only on one dimension of the type—an agent's value for good B does not affect her choice of quality and ordeal option if she chooses good A.

Definition 1. *Define a mechanism's* ***lowest participating values*** *as follows:*

$$\underline{a} = \sup\{a : \ U_A(a) = 0\}, \quad \underline{b} = \sup\{b : \ U_B(b) = 0\}.$$

Let a ***boundary*** *be a strictly increasing, continuous function $z : [\underline{a}, \overline{a}] \to [\underline{b}, \overline{b}]$ such that $\overline{a} \le 1$ and $\overline{b} \le 1$, with one of them holding with equality.*[6]

[5] Indeed, consider a mechanism allocating only A. Now, augment it by adding to its menu of options one which allocates good B with $x = 1$ with an ordeal of $c = 1 - \epsilon$. Since the distribution F of values (a, b) is full-support, for any $\epsilon > 0$ there will be a mass of agents who prefer the B-good option to any A-good option offered; that is, introducing this option increases their welfare. Note that adding the B-option only relaxes the supply constraint on A, and that the mass of the agents taking the B-option converges to 0 as $\epsilon \to 0$. Thus, the augmented mechanism satisfies the supply constraint for B for ϵ small enough.

[6] Note that despite the boundary $z(a)$ being written as a function of a, the setting is symmetric with respect to both goods. The characterization in the following lemma could equally well be formulated in terms of $z^{-1}(b)$.

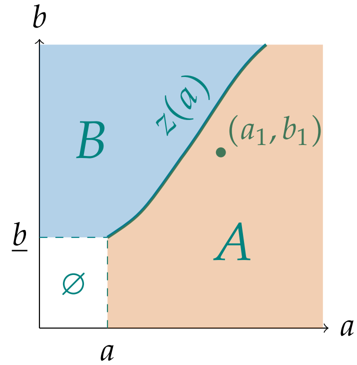

Figure 2: Types below the boundary (orange) choose good A and types above it (blue) choose good B.

Proposition 2. *Agents' choices of goods are characterized by the mechanism's lowest participating values* $\underline{a}, \underline{b}$ *and a boundary* z*:*

1. *Types* $(a, b) < (\underline{a}, \underline{b})$ *do not get either good.*[7]
2. *Types* $(a, b) > (\underline{a}, \underline{b})$ *get good* A *if* (a, b) *is below the boundary* z*, that is, if* $z(a) > b$.
3. *Types* $(a, b) > (\underline{a}, \underline{b})$ *get good* B *if* (a, b) *is above the boundary* z*, that is, if* $z(a) < b$.

Moreover, types on the boundary are indifferent between their favorite options for both goods, thus:

$$U_A(a) = U_B(z(a)) \quad \text{for all} \quad a \in [\underline{a}, \overline{a}]. \tag{I}$$

When all qualities of A- and B-goods come with ordeals, types with sufficiently low values for both of them, i.e. $(a, b) < (\underline{a}, \underline{b})$, do not participate. To understand the good choices of participating types, consider some (a_1, b_1) choosing good A (Figure 2). Then any type (a, b) with $a > a_1$ and $b < b_1$ will also choose good A—since she values the A-good even more than (a_1, b_1) and values the B-good even less, all the ordeal and quality options for good B are strictly less attractive to her than they were to (a_1, b_1). We can now notice that the types who are indifferent between their best options for either good lie on an increasing curve z originating from $(\underline{a}, \underline{b})$. By the above logic, all types below this curve choose A and pick some ordeal and quality option from its menu, while types above it choose B.

We can therefore think of our multidimensional mechanism design problem as two single-dimensional problems connected endogenously through the boundary z. While agents on its either side effectively face one-dimensional problems, making one of the goods more attractive invites more types to switch to it, effectively deforming the boundary.

[7]When comparing vectors, I will use $\geq$ and $>$ for point-wise comparisons.

4.2 Why damages can be optimal

In the single-good example, all patterns of sorting into good A and the outside option B could be implemented using both ordeals and damages; this is no longer the case when agents are heterogeneous in two dimensions. To see this, consider first the case in which the designer does not use damages, i.e. $x(a,b) \equiv 1$. In this case, (IC) requires that each good be given with a single ordeal, c_A or c_B. Agents of type (a,b) will then select good A if:

$$a - c_A > b - c_B,$$

and select good B otherwise. Note that this kind of screening can only lead to sorting patterns like the one illustrated in Figure 3a, where agents get good A if their types lie below a certain 45-degree line. It cannot create a sorting pattern like that in Figure 3b, where agents get good A if their types lie below a ray from the origin, that is, if:

$$\frac{b}{a} < q,$$

for some q, and get good B otherwise. Such a sorting pattern can be achieved with damages, however. Consider a mechanism offering A with a damage, $x = q < 1$, and B without it, $x = 1$, with no ordeals for either. The set of indifferent agents will then be given by:

$$a \cdot q = b \quad \Rightarrow \quad \frac{b}{a} = q.$$

Agents below and above this boundary will then choose goods A and B, respectively.

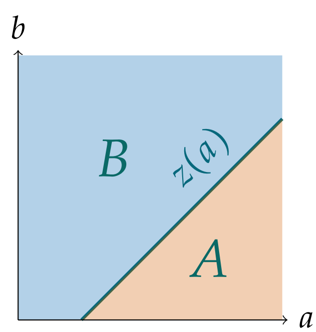

Figure 3a: Sorting pattern implementable with ordeals alone.

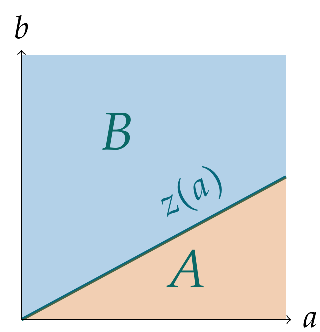

Figure 3b: Sorting pattern whose implementation requires damages.

The fact that achieving certain sorting patterns requires using damages breaks the intuition underlying Proposition 1. I now present an example where damages are required in the optimal mechanism.

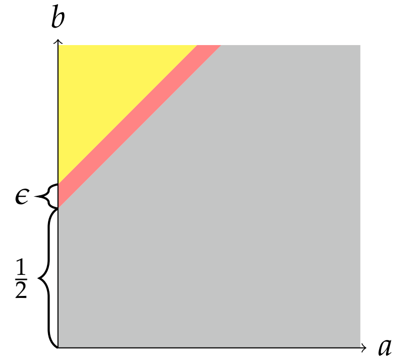

Figure 4: Distribution F from Example 1.

Example 1. *Consider the density:*

$$f(a,b) = \begin{cases} \epsilon \dfrac{2}{\left(\frac{1}{2}-\epsilon\right)^2}, & \text{if } b-a \geq \frac{1}{2}+\epsilon, \\ k\dfrac{2}{\epsilon-\epsilon^2}, & \text{if } b-a \in \left[\frac{1}{2}, \frac{1}{2}+\epsilon\right), \\ \frac{8}{7}(1-k-\epsilon), & \text{if } b-a < \frac{1}{2}. \end{cases}$$

and supplies given by:

$$\mu_A = 1-k-\epsilon, \quad \mu_B = k+\epsilon.$$

For $\epsilon > 0$ sufficiently small, a mechanism using only ordeals is not optimal.

The distribution from the example is illustrated in Figure 4. The probability masses in all three shaded areas are constant in ϵ: they equal to ϵ in the yellow area, k in the red area, and $1-k-\epsilon$ in the gray one. The supply of good B is chosen to exactly match the total mass of the yellow and red areas, while the supply of good A matches the mass in the gray area.

Consider a mechanism for this distribution which uses only ordeals but not damages. Discarding any of either good's supply would not be helpful (this point later shown formally in the proof of Theorem 1), so we can without loss consider only mechanisms where the whole supply is allocated. Without damages, this can only be achieved by giving out good B for free, and giving good A with an ordeal $c = 1/2$. Indeed, this mechanism induces a pattern of sorting illustrated in Figure 5a, with types shaded in orange getting good B and types shaded in blue getting good A. As discussed above, the boundary splitting the two regions is angled at 45 degrees.

Now, notice that for agents in the strip between the solid and dashed lines in Figure 5a, the surplus from getting good B over getting good A is at most ϵ. This is because most of their surplus is consumed by the ordeal $c = 1/2$. Consider then the case where $\epsilon \to 0$. As ϵ falls, this surplus goes to zero for the whole mass k of agents in the aforementioned strip. Similarly, the mass of agents above the dashed line, equal to ϵ, also tends to zero. Consequently, total welfare then tends to that which would have resulted from all agents getting good A for free.

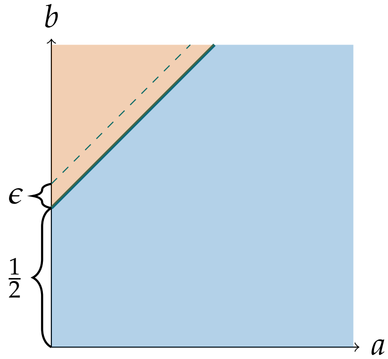

Figure 5a: The mechanism in Example 1 that gives good B with an ordeal.

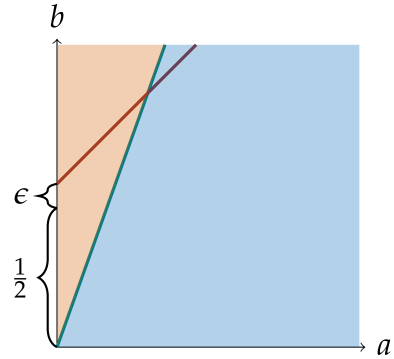

Figure 5b: The mechanism in Example 1 that damages good B.

Consider by contrast the mechanism which uses no ordeals but damages good B to the point where the resulting boundary satisfies both goods' supply constraints (Figure 5b). Here too, total welfare of agents above the red line becomes negligible as $\epsilon \to 0$. Note, however, that agents who are below the red line and far to the left of the green line now benefit substantially from getting good B over getting good A for free. These agents have a strong *relative* preference for B over A, and thus strongly prefer even damaged B to undamaged A. Since the mass of such agents does not go to zero as ϵ decreases, this mechanism creates substantial gains from allocating good B over A even in the limit case. This is in contrast to the ordeal-only mechanism, where all the gains from allocating B over A are "eaten away" by the ordeal used to screen high-b types, thus making the damage-based mechanism superior. Intuitively, this occurs because most of the consumers of good B under the ordeal-based mechanism have values very close to the "market clearing ordeal".

4.3 When are damages not optimal?

While damages can be useful in multidimensional settings, this is not the case when the following condition on the joint value distribution holds:

Assumption 1. *The value distribution F has a continuous density f and its inverse conditional anti-hazard rates*

$$\frac{F_{A|B}(a|b)}{f_{A|B}(a|b)}, \quad \frac{F_{B|A}(b|a)}{f_{B|A}(b|a)},$$

are strictly increasing in one of a and b and non-decreasing in the other one.

Since the argument behind this result uses optimal control methods, it also requires the following technical restriction on admissible mechanisms:

Assumption 2. *The designer is restricted to quality rules $x : [0,1]^2 \to [0,1]$ that are piece-wise continuously differentiable in each dimension of the type.*[8]

Under these assumptions, the designer's optimal mechanism is given by the following result:

Theorem 1. *The optimal mechanism offers only two options: it allocates undamaged goods A and B with ordeals c_A and c_B, respectively. These ordeals are chosen so that the whole supply of both goods is allocated.*

In fact, this optimal mechanism has a simple interpretation: the designer posts a pair of "market-clearing" ordeals which act as prices, replicating the competitive equilibrium allocation of goods (see Section 5 for a discussion). While it is easy to see that this mechanism is optimal when transfers are welfare-neutral, the theorem says the same is true when the "revenue" of the mechanism is completely wasted.

4.4 Proof of Theorem 1

In this section I outline the core steps and intuitions behind the proof of the result, as well as explain the role of both of assumptions. Before we begin the argument, let us note that we can assume without loss that the optimal mechanism features a boundary with $\overline{b} = 1$, that is, one that reaches the ceiling of the unit square before it reaches its right wall. This is because, by the definition of the boundary, one of $\overline{a}, \overline{b}$ equals to 1 and the setting is symmetric with respect to both goods. Consequently, in what follows we restrict attention to boundaries with this feature.

4.4.1 Reformulating the objective. Let us first observe that we can use the boundary z and the A-indirect utility U_A to rewrite total welfare in a concise way. To that end, we define an *extended boundary* $\hat{z}$ which is constructed from z as follows:

$$\hat{z}(a) = \begin{cases} 0, & \text{if } a < \underline{a}, \\ z, & \text{if } a \in [\underline{a}, \overline{a}], \\ 1, & \text{if } a > \overline{a}. \end{cases}$$

That is, $\hat{z}$ equals to z on the latter's domain, takes value zero below it and takes value 1 above it (Figure 6). The following result is then shown in the appendix:

Lemma 1. *Consider a mechanism with a boundary z and A-indirect utility U_A. Total welfare under this mechanism is then given by:*

$$U_A(1) - \int_0^1 U_A'(a) \cdot F(a, \hat{z}(a)) da. \tag{W*}$$

[8] Note Assumption 2 allows for allocation rules that are discontinuous at finitely many points.

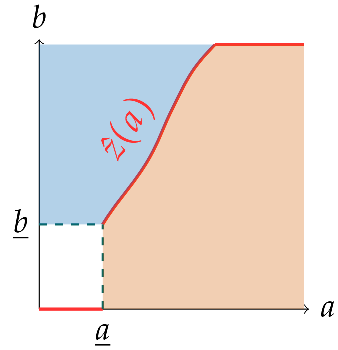

Figure 6: Extended boundary $\hat{z}$.

Intuitively, the lemma shows how the welfare of all agents in the mechanism can be written using only the welfare of agents choosing good A and the shape of the boundary. To see why this is the case, note that the utilities of agents getting good A depend only on their values for good A and not on those for good B, and vice versa. Moreover, recall that the agents on the boundary are indifferent between the two goods. It therefore follows that types on lying on the same L-shaped curves in Figure 7 all have the same utility. Indeed, this observation is captured in the equation defining the boundary:

$$U_A(a) = U_B(z(a)) \quad \text{for all} \quad a \in [\underline{a}, \overline{a}]. \tag{I}$$

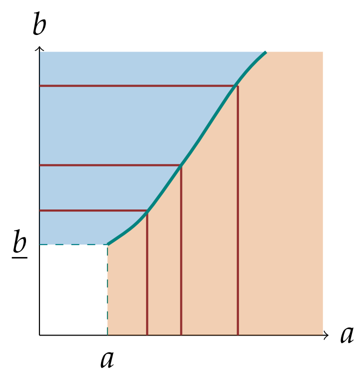

Figure 7: Agents on the same L-shaped curves have equal utilities.

We can therefore calculate welfare by integrating over types who choose A while also taking into account the B-taking types on the same L-shaped curves. Such a calculation yields the expression (W*). It is worth noting that this representation of total welfare does not rely on the anti-hazard rate and piece-wise continuous differentiability assumptions made for Theorem 1.

This form of the objective also bears resemblance to a Myersonian virtual value which would appear in a single-dimensional setting with a welfarist objective (see e.g. Condorelli (2012)). Moreover, by the envelope theorem, we can think of $U_A'(a)$ as the quality allocated to agents who receive good A and value it at a. Unfortunately, however, the setting does not yield itself to a Myersonian solution method. The reason for this is two-fold. First, the supply constraint in my model is not on total quality, but on the *mass of agents who receive good A*; this constraint depends not on the total assignment of U'_A, but on the probability masses on both sides of the boundary z. Second, the shape of the "virtual value" itself is endogenous to the choice of boundary. Indeed, it is optimizing over the shape of the boundary which poses the greatest difficulty in proving the result.[9]

I solve the problem using a novel approach, which proceeds in three steps. By Lemma 1, the pair (z, U_A) is a sufficient statistic for total welfare. In the first step, I therefore characterize the feasible (z, U_A) pairs. Then, in the second step, I fix a boundary z and find the welfare-maximizing U_A that is compatible with it. Finally, I optimize over the space of feasible pairs (z, U_A), where U_A is the optimal A-indirect utility associated with z. I use optimal control tools to show that the optimal boundary z^* has to be linear with a slope of 1, which in turn is optimally implemented without damages.

Characterizing feasible (z, U_A). We say (z, U_A) is *feasible* if there exists a mechanism (c, x, y) with A, B-indirect utilities U_A, U_B such that:

$$U_A(a) = U_B(z(a)) \quad \text{for all} \quad a \in [\underline{a}, \overline{a}]. \tag{I}$$

Lemma 2. *The pair (z, U_A) is feasible if and only if:*

1. *U'_A and U'_A/z' are non-decreasing,*
2. *The boundary z has finite, strictly positive one-sided derivatives at every $a \in (\underline{a}, \overline{a})$, and a finite, strictly positive left derivative at $\overline{a}$.*
3. *The supply constraint (S′) holds:*

$$\int_{\underline{a}}^{1} \int_{0}^{z(\min[a,\overline{a}])} f(a, v) dv\, da \leq \mu_A, \quad \int_{\underline{b}}^{1} \int_{0}^{z^{-1}(\min[b,\overline{b}])} f(v, b) dv\, db \leq \mu_B. \tag{S′}$$

Let us first discuss 1. Note (I) means that U_A and the boundary z uniquely pins down the B-indirect utility U_B. Moreover, differentiating (I) gives:

$$U'_A(a) = U'_B(z(a)) \cdot z'(a) \quad \Rightarrow \quad U'_B(z(a)) = \frac{U'_A(a)}{z'(a)}. \tag{DI}$$

[9] Indeed, the difference between my approach and the Myersonian one is suggested by Assumption 1, which is imposed on the (conditional) *anti*-hazard rate, and not on the hazard rate, as it would be in a standard Myersonian setting with a welfarist objective. The reason for why the anti-hazard rate appears in my condition, explained in the course of the proof, is logically distinct from the one for the presence of the hazard rate in the Myersonian problem.

We know, however, that U_B must be convex and, since z is increasing, this implies U'_A/z' has to be increasing. Thus, 1. boils down to a monotonicity conditions on U'_A and U'_B. 2. then uses (DI) to establish that any boundary implemented by U_A and U_B has to have certain regularity properties. Finally, 3. changes the way we express supply constraints. Rather than look at good allocations $y(a,b)$ directly, it takes advantage of the fact that types who get good A (B) are below (above) the boundary. It then measures the masses of agents getting either good by integrating over agents below and above z (Figure 8).

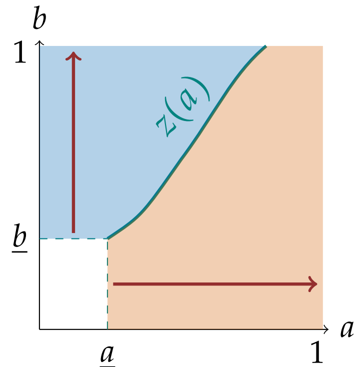

Figure 8: The supply condition (S′) ensures that the probability masses below the boundary (orange) and above it (blue) are at most μ_A and μ_B, respectively. The red arrows mark the directions of integration in the left-hand-sides in (S′).

4.4.2 Role of piece-wise differentiability. The results shown so far did not rely on Assumption 2, which restricts the designer to piece-wise continuously differentiable allocation rules $x(a,b)$.[10] It is, however, necessary for the results that follow. The role of this assumption is to guarantee the following regularity condition on the boundary z:

Fact 1. *Under Assumption 2, every implementable boundary is piece-wise twice continuously differentiable with strictly positive one-sided derivatives on $(\underline{a}, \overline{a}]$.*

This, in turn, lets us optimize over the space of boundaries using optimal control techniques.

Optimal U_A for a fixed boundary z. Let us now fix a boundary z and find the A-indirect utility U_A that maximizes total welfare (W*) subject to (z, U_A) being feasible.

Lemma 3. *Fix a boundary z. Then (z, U_A), with U_A defined by (1), maximizes total welfare (W*)*

[10]In the appendix, I provide a version of Lemma 2 under this additional restriction.

among all feasible pairs $(z, \check{U}_A)$.

$$U'_A(a) = \begin{cases} 0, & a \in (0, \underline{a}), \\ m(a) \cdot c, & a \in (\underline{a}, \overline{a}), \\ 1, & a \in (\underline{a}, 1). \end{cases} \tag{1}$$

where

$$m(a) = \exp\left(\int_{\underline{a}}^{a} \max\left[0, \tfrac{z''(s)}{z'(s)}\right] ds \right) \prod_{\substack{z'^{+}(t) > z'^{-}(t), \\ t < a}} \frac{z'^{+}(t)}{z'^{-}(t)}, \qquad c = \frac{1}{\max[m(\overline{a}),\ m(\overline{a})/z'^{-}(\overline{a})]}.$$

I now explain the intuition behind this lemma. First, note that U'_A is constant on all intervals where z is concave and is proportional to z' on all intervals where z is convex. Recall also that

$$U'_B(z(a)) = \frac{U'_A(a)}{z'(a)}. \tag{DI}$$

Thus, we can equivalently say that $U'_A(a)$ is constant on concave intervals of z while $U'_B(z(a))$ is constant on its convex intervals (Figure 9).

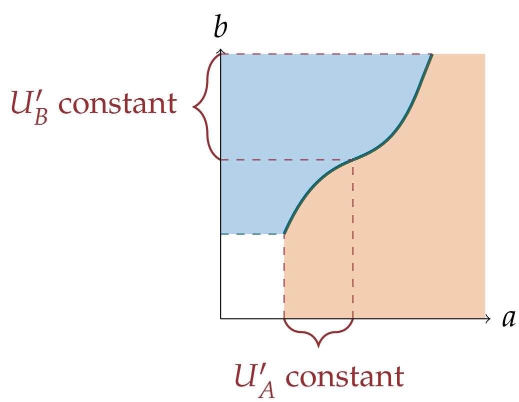

Figure 9: $U'_A(a)$ and $U'_B\left(z(a)\right)$ are constant where the boundary z is concave and convex, respectively.

Now, recall that by Lemma 2, $U'_A(a)$ and $U'_B\left(z(a)\right)$ must be increasing. The above observation therefore means that at least one of these monotonicity constraints must always bind. Intuitively, were neither constraint to bind on some interval, we could increase $U'_A(a)$ and $U'_B\left(z(a)\right)$ point-wise in a (DI)-preserving manner until one of them started binding. This would create point-wise higher utility profiles and thus produce superior welfare (W*).

To understand the economic intuition behind this logic, note that by the envelope theorem, $U'_A(a)$ corresponds to the quality x of the goods received by agents who get good A and value it at a. Consider then some type (a', b') on the boundary and suppose that the boundary is convex on some interval $[a'', a']$. Since $U'_B(z(a))$ is always non-decreasing, (DI) implies that $U'_A(a)$ has to be strictly increasing on this interval below a', and so that agents with $a < a'$ get a damaged version of good A. This means that, in general, making a boundary curve up or down somewhere requires damaging the good on at least one side of the boundary there. Note, however, that altering the shape of the boundary could, in principle, be beneficial despite this damage, as it might entice agents to choose between the A- and B-goods in a more socially efficient way. Nevertheless, more separation cannot be better *conditional on implementing the same boundary*. Consequently, a fixed boundary is implemented most efficiently by a mechanism that separates qualities as little as possible while still satisfying (DI) and both monotonicity constraints.

4.4.3 Showing the optimal boundary is linear. Having pinned down the optimal U_A for a given z, we can turn to searching for the best boundary with its optimal A-indirect utility. Let z^* be the optimal boundary and $\underline{a}^*, \underline{b}^*$ be the lowest participating values associated with it.

Proposition 3. *The optimal boundary* $z^* : [\underline{a}^*, \overline{a}^*] \to [\underline{b}^*, \overline{b}^*]$ *is linear.*

I prove this proposition by considering optimal control problems of choosing a boundary on a part of the boundary where it is concave/convex. I show that no boundary with strictly convex/concave parts or kinks can satisfy the necessary optimality conditions, and thus that the optimal boundary has to be linear. While the formal argument is relegated to the appendix, I provide an informal sketch capturing the core intuition behind it. This heuristic argument shows why the boundary cannot have strictly convex (and symmetrically, concave) segments, and why conditional inverse anti-hazard rate monotonicity is sufficient for this conclusion.

Consider some interval of $[\underline{a}, \overline{a}]$ on which the boundary is concave and, for simplicity, assume it consists of multiple small, linear pieces (Figure 10a). We will consider the perturbations to these linear pieces on this part of the boundary. Notice, however, that such perturbation have to respect the supply constraint (S′), and thus must preserve the probability mass underneath and above the boundary. Still, we can construct perturbations preserving the supply constraint by perturbing one piece of the boundary upwards and another one downwards in a ratio that leaves the probability masses unchanged (Figure 10b). First-order optimality conditions then tell us that when perturbing one such piece, we can capture the effect of this perturbation on the supply constraint by a Lagrange multiplier μ.

Now, recall that by Lemma 1, total welfare is given by (W*):

$$U_A(1) - \int_0^1 U'_A(a) \cdot F(a, \hat{z}(a)) da. \tag{W'}$$

Moreover, Lemma 3 tells us that U'_A is equal to some constant $c > 0$ on the region where the boundary is concave. Thus, we can write the effect of the boundary in the region $[\tilde{a}, \tilde{a} + \delta]$ as

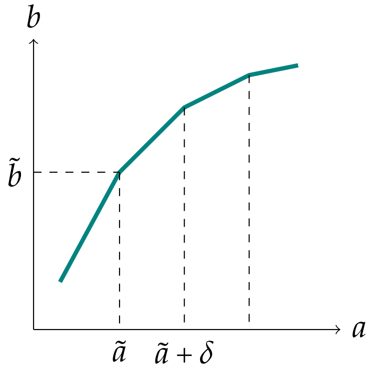

Figure 10a

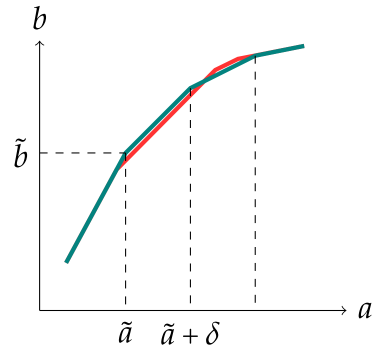

Figure 10b

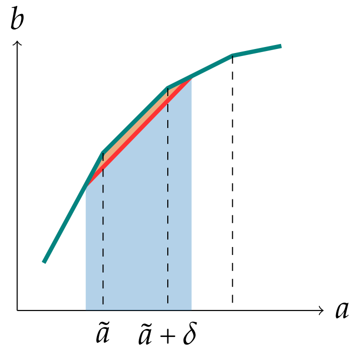

Figure 10c

follows, up to scaling:

$$V(\tilde{b}) := -\int_{\tilde{a}}^{\tilde{a}+\delta} F\left(a, \tilde{b} + (a-\tilde{a}) \cdot s\right) da.$$

Now, consider the effect of a small downward perturbation to the height of the boundary on this interval, as illustrated in Figure 10c. The first-order effect of this perturbation is given by:

$$-\frac{d}{d\tilde{b}} V(s) \;=\; \frac{d}{d\tilde{b}} \int_{\tilde{a}}^{\tilde{a}+\delta} F\left(a, \tilde{b} + (a-\tilde{a}) \cdot s\right) da \;-\; \mu \cdot \frac{d}{d\tilde{b}} \int_{\tilde{a}}^{\tilde{a}+\delta} \int_{0}^{\tilde{b}+(a-\tilde{a})\cdot s} f(a,b) db da.$$

The latter term captures the effect of the perturbation on the probability mass under the boundary, which is valued according to the aforementioned multiplier μ on the supply constraint (S). The first-order condition for this perturbation is then:

$$0 \;=\; \int_{\mathcal{K}} f(a,b)\, \mathrm{d}(a,b) \;-\; \mu \int_{\mathcal{L}} f(a,b)\, \mathrm{d}(a,b)$$

where $\mathcal{K}$ and $\mathcal{L}$ are, respectively, the limits of the blue region in Figure 10c, and of the orange region in Figure 10c, as the perturbation gets small. Dividing through by the latter integral then reduces the condition to:

$$0 \;=\; \frac{\int_{\mathcal{K}} f(a,b)\, \mathrm{d}(a,b)}{\int_{\mathcal{L}} f(a,b)\, \mathrm{d}(a,b)} \;-\; \mu.$$

Now, notice that when the length δ of the perturbed interval becomes small, we can use the following approximation:

$$\frac{\int_{\mathcal{K}} f(a,b)\, \mathrm{d}(a,b)}{\int_{\mathcal{L}} f(a,b)\, \mathrm{d}(a,b)} \;\approx\; \frac{\int_0^{z(\tilde{a})} \tilde{a} \cdot f(\tilde{a},t) dt}{\tilde{a} \cdot f(\tilde{a},t)} \;=\; \frac{F_{A|B}(\tilde{a}|z(\tilde{a}))}{f_{A|B}(\tilde{a}|z(\tilde{a}))}.$$

Thus, when the boundary is strictly concave on some region, a profitable perturbation analogous to that in Figure 10b does not exist only if at every point in that interval we have:

$$0 = \frac{F_{A|B}(\tilde{a}|z(\tilde{a}))}{f_{A|B}(a|z(a))} - \mu.$$

However, this cannot be the case by Assumption 1. Recall the inverse anti-hazard rate is strictly increasing in one of the variables, non-decreasing in the other one, and z is strictly increasing in a. Thus, if we were indifferent about perturbing the boundary slightly up or down at some level of a, we would strictly prefer to perturb it upwards for any higher a.

The proof formalizes this reasoning in the continuous case using optimal control methods. The intuition presented here explains why the boundary does not have strictly convex and concave intervals. The argument extends this reasoning to show the boundary cannot have kinks.[11]

4.4.4 Showing the optimal linear boundary has a unit slope. It therefore remains to find the optimal slope of the linear boundary, as well as the optimal quantities of both goods to allocate. The following lemma solves the former problem:

Lemma 4. *The optimal boundary z^* is linear with a slope of* 1.

I give an intuition for this result below. Suppose the boundary z was initially sloped at $s > 1$ and reached the ceiling of the unit square before it reached its right wall (Figure 11a).

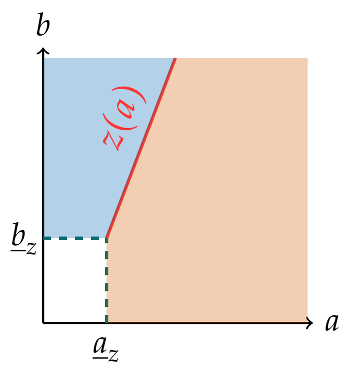

Figure 11a: s-sloped boundary z.

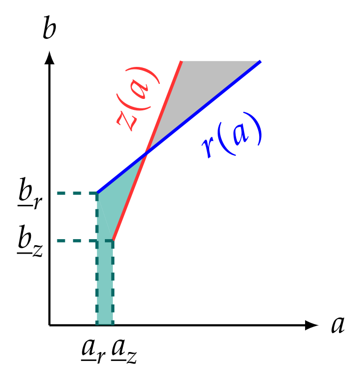

Figure 11b: Regions $\overline{\mathcal{D}}$ (green) and $\underline{\mathcal{D}}$ (grey).

By an argument similar to that in Subsection 4.2, a mechanism with such a boundary can be implemented by offering two options: good A with no damage and an ordeal of $\underline{a}_z$, and good B with quality s at an ordeal of $\underline{b}_z \cdot s$. Indeed, under such a mechanism, agents with values $(a, b) < (\underline{a}_z, \underline{b}_z)$ would not participate, while the others would sort between the two goods according to an s-sloped line. I now explain why such a boundary with a slope exceeding 1 cannot be

[11] Note that the presented intuition never required the inverse anti-hazard rate be strictly *increasing*, but just that it be strictly *monotonic*. In fact, the argument for why strictly convex and concave intervals cannot be optimal survives in the case where inverse anti-hazard rate is strictly decreasing. The latter part of the argument, which shows why the boundary cannot be piece-wise linear with kinks, requires it to be increasing.

optimal under Assumption 1. To that end, consider a slightly less steep boundary r such that the same amount of each good is allocated under it as it was under z (Figure 11b). A few features of this boundary are apparent. First, it crosses z once, and from above. Second, its lowest participating value for A is lower: $\underline{a}_r < \underline{a}_z$. Finally, consider the green and gray regions in Figure 11b; I will refer to them as $\underline{\mathcal{D}}, \overline{\mathcal{D}}$:

$$\underline{\mathcal{D}} = \left\{(a,b): 0 < a < a^*,\ \hat{z}(a) < b < \hat{r}(a)\right\} \quad \overline{\mathcal{D}} = \left\{(a,b): a^* < a < 1,\ \hat{r}(a) < b < \hat{z}(a)\right\}.$$

These denote the types who received good B under z and receive good A under r, and vice versa. Since the masses of agents who receive each good are equal for both boundaries, the masses of agents in these two regions must be equal as well:

$$\int_{\overline{\mathcal{D}}} f(a,b)\,\mathrm{d}(a,b) \;=\; \int_{\underline{\mathcal{D}}} f(a,b)\,\mathrm{d}(a,b). \tag{2}$$

Now, let us consider the difference Δ between the value of the objective (W*) for boundaries r and z. As explained previously, both mechanisms' U'_A correspond to the qualities of good A, and thus is uniformly equal to 1 as good A is undamaged in both mechanisms. The difference can then be written as:

$$\begin{aligned}
\Delta &= (1-\underline{a}_r) - \int_0^1 F\left(a,\hat{r}(a)\right)\,da - \left((1-\underline{a}_r) - \int_0^1 F\left(a,\hat{z}(a)\right)\,da\right) \\
&= (\underline{a}_z - \underline{a}_r) - \left(\int_0^1 F\left(a,\hat{r}(a)\right) - F\left(a,\hat{r}(a)\right)\,da\right) \\
&= (\underline{a}_z - \underline{a}_r) - \left(\int_{\overline{\mathcal{D}}} \frac{F_{A|B}(a \mid b)}{f_{A|B}(a \mid b)} \cdot f(a,b)\,\mathrm{d}(a,b) \;-\; \int_{\underline{\mathcal{D}}} \frac{F_{A|B}(a \mid b)}{f_{A|B}(a \mid b)} \cdot f(a,b)\,\mathrm{d}(a,b)\right).
\end{aligned}$$

We can therefore decompose the difference in welfare between the boundaries into two effects: $\underline{a}_z - \underline{a}_r$, which corresponds to the change in good A's ordeal, and the difference in integrals which corresponds to the change in the total values of the goods to their recipients, taking into account the change in their qualities and the change in recipient composition.

Assumption 1 lets us sign these two effects separately. Indeed, as mentioned, $\underline{a}_r < \underline{a}_z$, so the former effect is positive. To see how to sign the latter effect, recall that Assumption 1 sates that

$$\frac{F_{A|B}(a \mid b)}{f_{A|B}(a \mid b)}, \tag{3}$$

is increasing in both a and b. Note also that the region $\overline{\mathcal{D}}$ lies to the north-east of $\overline{\mathcal{D}}$, and thus the values of (3) are uniformly higher in it. Moreover, $\underline{\mathcal{D}}$ and $\overline{\mathcal{D}}$ contain equal probability masses by (2), and so it follows that the former integral must be higher.

For this reason, the optimal boundary has a slope of 1. Lemma 3 then tells us the mechanism implementing it does not use damages. The intuition for this follows that in Subsection 4.2: a boundary with a slope of 1 can be implemented by having agents choose between both undamaged goods with thier corresponding ordeals.

4.4.5 Discarding goods is never optimal. Finally, we need to show that the optimal mechanism allocates the whole supply of both goods. As we have already shown, this mechanism uses only ordeals to enforce the supply constraints; we can then interpret these ordeals as prices. Then, intuitively, if the designer were to discard some of one good's supply, she could do better by simply lowering its associated ordeal and letting demand for it increase. As a result the designer selects the (unique) pair of ordeals at which the markets for both goods clear.

Lemma 5. *The optimal mechanism allocates the whole supply of both goods.*

5 Partially wasteful ordeals

I now consider the case where every unit of the ordeal generates a benefit $\gamma \in [0,1]$ for the designer; I interpret this partially wasteful screening device as a monetary payment and let γ be the designer's value for the collected revenue. If this revenue could be costlessly rebated to participants, it would be natural to set $\gamma = 1$. This would in turn mean that the transfers in the mechanism are welfare-neutral, effectively reducing the designer's objective to maximizing allocative efficiency. In this case, we recover the standard result about the optimality of market mechanisms. Since the allocation of the mechanism in Theorem 1 replicated the competitive equilibrium allocation, we get the following result:

Proposition 4. *The mechanism described in Theorem 1 maximizes allocative efficiency* (E)*:*

$$\int \mathbb{1}_{y(a,b)=A} \cdot x(a,b) \cdot a + \mathbb{1}_{y(a,b)=B} \cdot x(a,b) \cdot b \; \mathrm{d}F(a,b), \tag{E}$$

subject to the supply constraint (S)*.*

In practice, however, we might expect the designer to consider transfers wasteful. For government programs, this could be the case when rebating revenue to participants is possible but costly due to bureaucratic inefficiency, or because distributing cash lacks the screening benefits of in-kind transfers.[12] In such cases, a redistributive designer might not want to use the program to collect revenue, as its participants are significantly poorer than the average taxpayer.[13]

[12] When the designer hands out a free (or subsidized) inferior good, only relatively poor agents will want to participate as wealthier ones can afford higher-quality alternatives. Thus, the subsidy is automatically targeted to those who need it most (Besley and Coate, 1991). As soon as the designer hands out cash, such targeting disappears as money is desired by everyone, regardless of wealth.

[13] While the model does not explicitly account for wealth differences or heterogeneous welfare weights among agents, this can be viewed as an approximation of a scenario where such differences exist but are relatively small

These concerns can be modelled by considering payments a *partially* wasteful instrument, with a weight $\gamma \in [0,1]$ on revenue in the designer's objective:

$$\int U[a,b,(c,x,y)(a,b)]\,\mathrm{d}F(a,b) \;+\; \gamma \cdot \int c(a,b)\,\mathrm{d}F(a,b). \tag{WR}$$

As may be intuitive, including partial revenue considerations in the designer's objective does not break Theorem 1:

Corollary 1. *Theorem 1 continues to hold when the designer maximizes* (WR) *for any* $\gamma \in [0,1]$.

The corollary follows because (WR) is a linear combination of the objectives (W) and (WR), each of which is maximized by the ordeal-only mechanism (by Theorem 1 and Proposition 4, respectively). The same mechanism must therefore also maximize their combination.

6 Heterogeneous ordeal costs

Certain ordeals are more costly to some agents than to others. For instance, poorer agents who benefit most from social programs may find monetary payments more burdensome; conversely, such agents might be more willing to engage in hassles such as queueing. I extend the model to accommodate such heterogeneous ordeal costs: suppose agents' types are three-dimensional, (a,b,r), where r represents an agent's unit cost for the ordeal. Utilities are then given by:

$$\begin{aligned} &x \cdot a - r \cdot c \;\text{ if }\; y = A, \\ &x \cdot b - r \cdot c \;\text{ if }\; y = B. \end{aligned}$$

I assume r is distributed on $[\underline{r}, \overline{r}]$ with $\underline{r} > 0$. Types (a,b,r) are then distributed according to F with full support on $[0,1]^2 \times [\underline{r}, \overline{r}]$.

Fortunately, the methods from the baseline model carry over to this enriched setup. Following Dworczak et al. (2021), we observe that all agents with types (a,b,r) and $(a/r, b/r, 1)$ are behaviorally equivalent. I can then adopt an approach similar to theirs: I renormalize all agents' ordeal cost to one, and identify agents with transformed values

$$(\tilde{a}, \tilde{b}) = \left(\frac{a}{r}, \frac{b}{r}\right),$$

and a welfare weight $\lambda(\tilde{a},\tilde{b})$ equal to the expectation of r among agents whose types were renormalized to $(\tilde{a},\tilde{b})$:

$$\lambda(\tilde{a},\tilde{b}) = \mathbb{E}\left[r \;\middle|\; \frac{a}{r} = \tilde{a},\; \frac{b}{r} = \tilde{b}\right].$$

between participants compared to the gap between participants and the average taxpayer. This is especially likely when the designer allocates inferior goods, such as public housing in undesirable areas. The designer's welfare-weighted objective can then be approximated by a constant weight on all participants, which is distinct from that on revenue, representing her welfare weight for the average taxpayer.

We will denote by G of transformed types implied by F; note that G has full support on $[0, 1/\underline{r}]^2$. We assume that G has a piece-wise Lipschitz continuous density g. The designer's objective then reduces to maximizing total *weighted* welfare:

$$\int \lambda(\tilde{a}, \tilde{b}) \cdot U[\tilde{a}, \tilde{b}, (c, x, y)(\tilde{a}, \tilde{b})] \; \mathrm{d}G(\tilde{a}, \tilde{b}). \tag{WW}$$

We then get the following result, which is a direct analog of Theorem 1:

Corollary 2. *Let G have a continuous density g and*

$$\frac{\int_0^a \lambda(v, b) \cdot g(v, b) dv}{g(a, b)}, \quad \frac{\int_0^a \lambda(v, b) \cdot g(v, b) dv}{g(a, b)},$$

be strictly increasing in one of a and b and non-decreasing in the other one. Then the optimal mechanism offers only two options: it allocates undamaged goods A and B with ordeals c_A and c_B. These ordeals are chosen so that the whole supply of both goods is allocated.

The conditions of the corollary are analogous to those of Theorem 1, except the density of types is multiplied by the welfare weights the planner attributes to them. While the proof is almost entirely analogous, I outline the key modifications in the appendix.

7 Conclusion

The contribution of this paper is the distinction between two kinds of wasteful screening instruments: ordeals, which are separable from agents' values for the allocated good, and damages, which are more burdensome for agents whose values for the goods are higher. I study their screening properties when they are used together and show that screening with damages is redundant when agents differ only with respect to the value for one good. This is no longer the case, however, when heterogeneity in values is multidimensional. Still, even in this case, ordeal-only mechanisms are optimal under certain regularity conditions.

My analysis provides insights into the optimal design of public housing allocation mechanisms. It recognizes that households choosing developments based on wait-times amounts to screening with a damage instrument. While wait-times are often not an explicit design choice but equilibrium objects, they continue to exhibit screening properties described by the model. Moreover, they can be influenced by rebalancing the subsidies offered for different kinds of units. My results provide an argument for such a reform: by raising the subsidy for less-desired units and increasing the price of popular ones, the designer can bring the lengths of their wait-lists closer together, and in doing so diminish the screening role of the damage instrument in the program. Instead, participants would be incentivized to choose between options based on monetary costs: a (partially) wasteful ordeal instrument.

References

AKBARPOUR, M., P. DWORCZAK, AND F. YANG (2023): "Comparison of Screening Devices," in *Proceedings of the 24th ACM Conference on Economics and Computation*, 60–60.

ALATAS, V., R. PURNAMASARI, M. WAI-POI, A. BANERJEE, B. A. OLKEN, AND R. HANNA (2016): "Self-targeting: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 124, 371–427.

ARNOSTI, N. AND P. SHI (2020): "Design of lotteries and wait-lists for affordable housing allocation," *Management Science*, 66, 2291–2307.

BARZEL, Y. (1974): "A theory of rationing by waiting," *The Journal of Law and Economics*, 17, 73–95.

BESLEY, T. AND S. COATE (1991): "Public Provision of Private Goods and the Redistribution of Income," *The American Economic Review*, 81, 979–984.

——— (1992): "Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs," *The American Economic Review*, 82, 249–261.

BLOCH, F. AND D. CANTALA (2017): "Dynamic Assignment of Objects to Queuing Agents," *American Economic Journal: Microeconomics*, 9, 88–122.

CONDORELLI, D. (2012): "What money can't buy: Efficient mechanism design with costly signals," *Games and Economic Behavior*, 75, 613–624.

DENECKERE, R. J. AND P. R. MCAFEE (1996): "Damaged goods," *Journal of Economics & Management Strategy*, 5, 149–174.

DESHPANDE, M. AND Y. LI (2019): "Who is screened out? Application costs and the targeting of disability programs," *American Economic Journal: Economic Policy*, 11, 213–248.

DUPAS, P., V. HOFFMANN, M. KREMER, AND A. P. ZWANE (2016): "Targeting health subsidies through a nonprice mechanism: A randomized controlled trial in Kenya," *Science*, 353, 889–895.

DWORCZAK, P., S. D. KOMINERS, AND M. AKBARPOUR (2021): "Redistribution Through Markets," *Econometrica*, 89, 1665–1698.

GUL, F. AND E. STACCHETTI (1999): "Walrasian equilibrium with gross substitutes," *Journal of Economic theory*, 87, 95–124.

HARTLINE, J. D. AND T. ROUGHGARDEN (2008): "Optimal mechanism design and money burning," in *Proceedings of the fortieth annual ACM symposium on Theory of computing*, 75–84.

KELSO JR, A. S. AND V. P. CRAWFORD (1982): "Job matching, coalition formation, and gross substitutes," *Econometrica: Journal of the Econometric Society*, 1483–1504.

KLEVEN, H. J. AND W. KOPCZUK (2011): "Transfer program complexity and the take-up of social benefits," *American Economic Journal: Economic Policy*, 3, 54–90.

LESHNO, J. D. (2022): "Dynamic matching in overloaded waiting lists," *American Economic Review*, 112, 3876–3910.

NEUSTADT, L. W. (1976): *Optimization: A Theory of Necessary Conditions*, Princeton University Press.

NICHOLS, A. L. AND R. J. ZECKHAUSER (1982): "Targeting Transfers through Restrictions on Recipients," *The American Economic Review*, 72, 372–377.

NICHOLS, D., E. SMOLENSKY, AND T. N. TIDEMAN (1971): "Discrimination by waiting time in merit goods," *The American Economic Review*, 61, 312–323.

SEIERSTAD, A. AND K. SYDSAETER (1986): *Optimal control theory with economic applications*, Elsevier North-Holland, Inc.

VAN DEN BERG, V. AND E. T. VERHOEF (2011): "Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay," *Journal of Public Economics*, 95, 983–992.

VAN DIJK, W. (2019): "The socio-economic consequences of housing assistance," *University of Chicago Kenneth C. Griffin Department of Economics job market paper, 0–46 i–xi*, 36.

VAN OMMEREN, J. N. AND A. J. VAN DER VLIST (2016): "Households' willingness to pay for public housing," *Journal of Urban Economics*, 92, 91–105.

VICKREY, W. (1973): *Pricing, metering, and efficiently using urban transportation facilities*, 476.

WALDINGER, D. (2021): "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," *American Economic Review*, 111, 2660–2696.

WORLD HEALTH ORGANIZATION (2023): "High-value referrals: learning from challenges and opportunities of the COVID-19 pandemic: concept paper," *High-value referrals: learning from challenges and opportunities of the COVID-19 pandemic: concept paper*.

YANG, F. (2021): "Costly multidimensional screening," *arXiv preprint arXiv:2109.00487*.

A Waitlists

In this section, I consider a simple environment where goods and agents arrive dynamically. I consider a patient designer who optimizes steady-state welfare by selecting menus of wait-times, ordeals, and allocation probabilities for both kinds of good. I show that the designer's problem has the structure of the baseline model.

The designer distributes two types of goods, A and B with agents' values for them given by (a, b). Goods and agents arrive continuously over time. At every instant $\tau \in \mathbb{R}$, flow masses $\mu_A, \mu_B > 0$ of goods A and B arrive, with $\mu_A + \mu_B \leq 1$. Concurrently, a unit flow mass of agents arrives, with types (a, b) distributed according to a joint distribution F.

There are two separate waitlists, one for each good. For each waitlist, the designer chooses a menu of three objects: ordeals $c \in \mathbb{R}_+$, wait-times $t \in \mathbb{R}_+$, and the probabilities $p \in [0, 1]$ with which the agent gets the good at the end of the wait. I assume without loss that the ordeal is completed when the agent joins the waitlist. Thus, when a type-(a, b) agent participates and selects the option (c, t, p) from good y's menu, her utility is:

$$p \cdot e^{-\rho t} \cdot a - c \;\; \text{if} \;\; y = A,$$
$$p \cdot e^{-\rho t} \cdot b - c \;\; \text{if} \;\; y = B,$$

where $\rho > 0$ is the common discount rate. The agent can also choose not to participate and receive utility 0.

All agents of the same type choose the same good and menu option, regardless of when they arrive. Moreover, I assume the designer cannot ban the agent from re-entering the mechanism if she ends up not receiving the good at the end of her wait. Note that if this happens, the agent faces a problem identical to the one she faced upon arrival, and thus will again choose the same good and option from its menu.

I restrict attention to mechanisms that admit a steady state of the system. Note that such a steady state exists if and only if the mass of agents choosing each good in every period is at most the arrival rate of that good:

$$\int \mathbb{1}_{(a,b) \text{ chooses } A} \, \mathrm{d}F(a, b) \leq \mu_A, \quad \int \mathbb{1}_{(a,b) \text{ chooses } B} \, \mathrm{d}F(a, b) \leq \mu_B. \tag{SS}$$

Crucially, the constraint is unaffected by the fact that the designer can offer goods probabilistically: since agents can always re-enter, all types selecting a menu option which gives them the good with a positive probability will get it at some point almost surely. Thus, for balance to be maintained, every agent choosing a given good must have an associated unit of that good arriving in the system in the representative period, regardless of her allocation probability.

Moreover, note that agents' utilities depend only on *expected discounting*, and not the probability of allocation p and wait-time t separately. Since wait-times do not figure in the steady state constraint (SS), we can without loss restrict the designer to letting agents choose between pairs of ordeals c and expected discounting $x \in \mathbb{R}_+$. Moreover, by the revelation principle, we can think of the designer as choosing steady state allocation rules for ordeals, expected discounting, and goods, $(c,x,y):[0,1]^2 \to \mathbb{R}_+^2 \times \{A,B,\varnothing\}$, subject to (IC), (IR) and the steady state constraint (SS′):

$$\text{for all } (a,b),(a',b') \in [0,1]^2, \quad U[a,b,(c,x,y)(a,b)] \geq U[a,b,(c,x,y)(a',b')], \tag{IC}$$

$$\text{for all } (a,b) \in [0,1]^2, \quad U[a,b,(c,x,y)(a,b)] \geq 0. \tag{IR}$$

$$\int \mathbb{1}_{y(a,b)=A} \, \mathrm{d}F(a,b) \leq \mu_A, \quad \int \mathbb{1}_{y(a,b)=B} \, \mathrm{d}F(a,b) \leq \mu_B. \tag{SS′}$$

This makes the problem identical to that in the baseline model. Note that despite allowing the designer to randomize allocations within each waitlist, the free re-entry restriction renders this tool unhelpful. This highlights that the distinction between waitlists and lotteries is contingent on the designer's ability to track agents' identity and exclude those who have already participated (see Arnosti and Shi (2020) for a related discussion). This natural restriction justifies not allowing the designer to allocate goods probabilistically in the baseline model. Note, however, that considering only deterministic mechanisms is still restrictive in that it does not allow the designer to randomize between different kinds of goods.

B Omitted proofs

First, recall that U_A, U_B are convex and strictly increasing on $[\underline{a}, \overline{a}]$ and $[\underline{b}, \overline{b}]$, respectively. As such, they are differentiable except at countably many points and are uniquely pinned down by their derivatives, wherever these exist, and $U_A(\underline{a}) = U_B(\underline{b}) = 0$. Moreover, wherever they fail to be differentiable, we have:

$$U_A'^{-}(t') = \lim_{t \uparrow t',\, t \in D} U_A'(t), \quad U_A'^{+}(t') = \lim_{t \downarrow t',\, t \in D} U_A'(t),$$

where $D \subset [\underline{a}, \overline{a}]$ is the set of points where U_A' exists. Analogous limits hold for U_B.

B.1 Proof of Proposition 2

Suppose some type $(a,b) < (\underline{a}, \underline{b})$ could weakly benefit from requesting either good. Then some type $(a+\epsilon, b+\epsilon) < (\underline{a}, \underline{b})$, for $\epsilon > 0$ small enough, would strictly benefit from it, so one of $U_A(a+\epsilon)$ and $U_B(b+\epsilon)$ would have to be strictly above zero. Since U_A, U_B are increasing, this contradicts the definition of $(\underline{a}, \underline{b})$. Thus, $y(a,b) = \varnothing$ for all $(a,b) < (\underline{a}, \underline{b})$.

Analogously, all types for whom $a > \underline{a}$ or $b > \underline{b}$ strictly benefit from choosing one of the goods. Moreover, a positive mass of types gets either good, and thus $\underline{a}, \underline{b} < 1$.

Now, suppose w.l.o.g. that $U_B(1) \geq U_A(1)$. We will now construct the boundary z and prove it has the above properties. For $a \in [\underline{a}, 1]$, define:

$$z(a) := \inf \left\{ b : U_B(b) \geq U_A(a) \right\}.$$

Note that, by construction, $z(\underline{a}) = \underline{b}$. Also, since $U_B(1) \geq U_A(1)$, it must be that $z(1) := \overline{b} \leq 1$.

Since U_A and U_B are both continuous and strictly increasing on $[\underline{a}, 1]$ and $[\underline{b}, 1]$, $z(a)$ is also continuous and satisfies:

$$U_A(a) = U_B(z(a)) \quad \Longrightarrow \quad z(a) = U_B^{-1}\left(U_A(a)\right).$$

Note $z(a)$ is a composition of two strictly increasing functions, so it is also strictly increasing.

By construction, any type $(a, z(a)) > (\underline{a}, \underline{b})$ is indifferent between her best options for both goods. Then, by a single-crossing argument, any type $(a', z(a))$ with $a' > a$ strictly prefers to get good A. Analogously, any type (a, b') with $b' > z(a)$ strictly prefers to get good B.

B.2 Proof of Lemma 1

Proposition 1 lets us rewrite total welfare (W) in terms of A, B-indirect utilities U_A, U_B and their associated boundary z:

$$\int_{\underline{a}}^{1} \int_{0}^{z(\min[a,\bar{a}])} f(a,v)dv \cdot U_A(a)\, da \; + \; \int_{\underline{b}}^{1} \int_{0}^{z^{-1}(\min[b,\bar{b}])} f(v,b)dv \cdot U_B(b)\, db. \tag{W'}$$

Since $\bar{b} = 1$, the above becomes:

$$\int_{\underline{a}}^{\bar{a}} \int_{0}^{z(a)} f(a,v)dv \cdot U_A(a)\, da \; + \; \int_{z(\underline{a})}^{z(\bar{a})} \int_{0}^{z^{-1}(b)} f(v,b)dv \cdot U_B(b)\, db + \int_{\bar{a}}^{1} \int_{0}^{1} f(a,v)dv \cdot U_A(a)\, da.$$

A change of variables yields:

$$\int_{\underline{a}}^{\bar{a}} \underbrace{\left(\int_{0}^{z(a)} f(a,v)dv + \int_{0}^{a} f(v,z(a))dv \cdot z'(a) \right)}_{=\frac{d}{da}F(a,z(a))} U_A(a)\, da \; + \; \int_{\bar{a}}^{1} \underbrace{\int_{0}^{1} f(a,v)dv}_{=\frac{d}{da}F(a,1)}\, U_A(a)\, da.$$

Denote the former integral by K and the latter one by L. Integrating them by parts gives:

$$K = U_A(\bar{a}) \cdot F(\bar{a}, z(\bar{a})) - U_A(\underline{a}) \cdot F(\underline{a}, z(\underline{a})) - \int_{\underline{a}}^{\bar{a}} U_A'(a) \cdot F(a, z(a))\, da.$$

Recall that $U_A(\underline{a}) = 0$. Thus, we get:

$$K = U_A(\bar{a}) \cdot F(\bar{a}, 1) - \int_{\underline{a}}^{\bar{a}} U_A'(a) \cdot F\left(a, z(a)\right)\, da.$$

Integrating L by parts gives:

$$L = U_A(1) - U_A(\bar{a}) \cdot F(\bar{a}, 1) - \int_{\bar{a}}^{1} U_A'(a) \cdot F(a, 1)\, da.$$

Summing K and L, we get:

$$U_A(1) - \int_{\underline{a}}^{\overline{a}} U_A'(a) \cdot F(a, z(a))\, da - \int_{\overline{a}}^{1} U_A'(a) \cdot F(a,1)\, da = U_A(1) - \int_0^1 U_A'(a) \cdot F(a, \hat{z}(a)) da.$$

B.3 Proof of Lemma 2

($\Leftarrow$). Fix any feasible mechanism (c, x, y). I first show 1. Since U_A and U_B are convex and strictly positive above $\underline{a}, \underline{b}$, respectively, U_A' and U_B' are non-decreasing and strictly positive above $\underline{a}$ and $\underline{b}$. Moreover, Proposition 2 gives:

$$U_A(a) = U_B(z(a)) \quad \text{for all} \quad a \in [\underline{a}, \overline{a}]. \tag{4}$$

Differentiating gives the following wherever z' exists:

$$U_A'(a) = U_B'(z(a)) \cdot z'(a) \quad \Rightarrow \quad U_B'(z(a)) = \frac{U_A'(a)}{z'(a)}.$$

Since z is strictly increasing on $[\underline{a}, \overline{a}]$ and U_B' is non-decreasing, U_A'/z' is non-decreasing and strictly positive above $\underline{a}$. I now show 2. Recall that by Proposition 2, z is pinned down by:

$$z = U_B^{-1} \circ U_A.$$

Fix any $a \in (\underline{a}, \overline{a}]$ and note that:

$$\begin{aligned}
z'^{+}(a) &= \lim_{\delta \downarrow 0} \frac{z(a+\delta) - z(a)}{\delta} \\
&= \lim_{\delta \downarrow 0} \frac{U_B^{-1}(U_A(a+\delta)) - U_B^{-1}(U_A(a))}{U_A(a+\delta) - U_A(a)} \cdot \frac{U_A(a+\delta) - U_A(a)}{\delta} \\
&= \left(U_B^{-1}\right)'^{+} (U_A(a)) \cdot U_A'^{+}(a) \\
&= \frac{U_A'^{+}(a)}{(U_B'^{+} \circ U_B^{-1} \circ U_A)(a)} = \frac{U_A'^{+}(a)}{(U_B'^{+} \circ z)(a)} > 0,
\end{aligned}$$

where the last inequality follows because $z = U_B^{-1} \circ U_A$ by (4). The argument for left-derivatives is analogous. Let us now show 3 holds. Recall $U_A(a) > U_B(b)$ for almost all agents for whom $y(a,b) = A$, so by Proposition 2:

$$\int_{\underline{a}}^{1} \int_0^{z(\min[a,\overline{a}])} f(a,v) dv\, da = \int \mathbb{1}_{y(a,b)=A}\, \mathrm{d}F(a,b) \leq \mu_A,$$

where the inequality holds by (S). An analogous expression holds for B.

($\Rightarrow$). Fix any (z, U_A) satisfying 1. – 3.; I will construct a mechanism (c, x, y) that corresponds to them. Choose some continuous $U_B : [0,1] \to [0,1]$ satisfying:

$$U_B'\big(z(a)\big) = \begin{cases} 0, & a < \underline{a}, \\ \dfrac{U_A'(a)}{z'(a)}, & a \in \big(\underline{a}, \overline{a}\big). \end{cases} \tag{5}$$

Note 1. guarantees that U_B' is non-decreasing.

I will show that U_A, U_B are the A,B-indirect utilities for the following mechanism, and that the mechanism is feasible:

$$c(a,b) = \begin{cases} a \cdot U_A'^{-}(a) - U_A(a), & \text{if } \tilde{U}_A(a) \geq \tilde{U}_B(b), \\ b \cdot U_B'^{-}(b) - U_B(b), & \text{if } \tilde{U}_B(b) > \tilde{U}_A(a), \end{cases}$$

$$x(a,b) = \begin{cases} U_A'^{-}(a), & \text{if } U_A(a) \geq U_B(b), \\ U_B'^{-}(b), & \text{if } U_B(b) > U_A(a), \end{cases}$$

$$y(a,b) = \begin{cases} \varnothing, & \text{if } (a,b) \leq (\underline{a}, \underline{b}), \\ A, & \text{if } a > \underline{a} \text{ and } U_A(a) \geq U_B(b), \\ B, & \text{if } b > \underline{b} \text{ and } U_B(b) > U_A(a). \end{cases}$$

Since U_A, U_B are strictly increasing and convex, a standard argument verifies that, under this ordeal rule, no (a,b) wants to misreport to (a',b') for which $y(a',b') = y(a,b)$. That is, conditional on choosing the good she was assigned, (a,b) prefers her assigned quality and ordeal option. Then U_A, U_B are indeed the A,B-indirect utilities for this mechanism because:

$$a \cdot U_A'^{-}(a) - \big(a \cdot U_A'^{-}(a) - U_A(a)\big) = U_A(a), \qquad b \cdot U_B'^{-}(b) - \big(b \cdot U_B'^{-}(b) - U_B(b)\big) = U_B(b).$$

Verifying that (IC) holds for (c,x,y) thus only requires checking that no (a,b) wants to misreport to (a',b') for which $y(a',b') \neq y(a,b)$. But since $U_A(a)$ and $U_B(b)$ are the best utilities (a,b) can get from either good, this is true by the construction of $y(a,b)$. Note also that (IR) must hold as both A,B-indirect utilities are positive everywhere.

I now verify that the mechanism implements boundary z. (5) tells us that wherever z' exists, we have:

$$z'(a) \cdot U_B'(z(a)) = U_A'(a) \quad \Longrightarrow \quad \frac{d}{da} U_B(z(a)) = \frac{d}{da} U_A(a). \tag{6}$$

Moreover, z is increasing, so z' exists a.e. Since U_A, U_B are absolutely continuous and $U_B(z(\underline{a})) = U_A(\underline{a}) = 0$, we can integrate both sides of (6) to get:

$$U_B(z(a)) = U_A(a).$$

It therefore remains to check the supply condition (S). But by Proposition 2:

$$\int \mathbb{1}_{y(a,b)=A}\, \mathrm{d}F(a,b) = \int_{\underline{a}}^{1} \int_{0}^{z(\min[a,\overline{a}])} f(a,v)\, dv\, da \leq \mu_A,$$

$$\int \mathbb{1}_{y(a,b)=B}\, \mathrm{d}F(a,b) = \int_{\underline{b}}^{1} \int_{0}^{z^{-1}(\min[b,\overline{b}])} f(v,b)\, dv\, db \leq \mu_B,$$

where the inequalities hold by 3.

B.4 Lemma 2 under Assumption 2 and the proof of Fact 1

The following fact refines the characterization of feasible (z, U_A) under Assumption 2. Fact 1 is a corollary of this result.

Lemma 6. *Let $z : [\underline{a}, \overline{a}] \to [\underline{b}, \overline{b}]$ be a boundary and U_A be an A-indirect utility. Then, under Assumption 2, the pair (z, U_A) is feasible if and only if:*

(a) U_A' and U_A'/z' are non-decreasing and strictly positive above $\underline{a}$,

(b) The boundary z has finite, strictly positive one-sided derivatives at every $a \in (\underline{a}, \overline{a})$, and a finite, strictly positive left derivative at $\overline{a}$.

(c) The supply constraint (S′) holds.

(d) $U_A', U_A'/z'$ are piece-wise continuously differentiable and z is piece-wise twice continuously differentiable,

Proof. The $(\Leftarrow)$ direction follows because under (a) – (d), Lemma 2 applies. Let us show the $(\Rightarrow)$ direction. (a), (b) and (c) follow from Lemma 2, so it remains to show (d). By Proposition 2, types $(a, 0)$ where $a > \underline{a}$ choose good A and types $(0, b)$ where $b > \underline{b}$ choose good B. Then, by the envelope theorem, the following holds wherever $x(a, 0)$, $x(0, b)$ are continuous in a, b, respectively:

$$U_A'(a) = x(a,0) \;\text{ for }\; a > \underline{a}, \quad U_B'(b) = x(0,b) \;\text{ for }\; b > \underline{b}. \tag{7}$$

Since x is piece-wise differentiable, this implies that U_A' and U_B' are piece-wise continuously differentiable and U_A, U_B are piece-wise differentiable. Moreover, recall Proposition 2 tells us:

$$U_A = U_B \circ z \quad \Rightarrow \quad U_B^{-1} \circ U_A = z, \tag{8}$$

and thus z is also piece-wise continuously differentiable. Consider then any interval on which z' exists. Differentiating the first equation in (8) then tells us that:

$$\frac{U_A'}{U_B' \circ z} = z'. \tag{9}$$

Then, since U'_A, U'_B and z are all piece-wise continuously differentiable, so is z'. We have thus shown z is piece-wise twice continuously differentiable and U'_A is piece-wise continuously differentiable. The fact that U'_A/z' is piece-wise continuously differentiable then follows from U'_A and z' being strictly positive and piece-wise continuously differentiable on $[\underline{a}, \overline{a}]$. □

B.5 Proof of Lemma 3

First, note that $U'_A(b) = 1$ for $a \in (\underline{a}, \overline{a})$ and that:

$$\max\left[U'^{-}_A(\overline{a}),\ U'^{-}_A(\overline{a})/z'^{-}(\overline{a})\right] = 1. \tag{10}$$

Otherwise, we could replace U_A with $\tilde{U}_A$ defined by:

$$\tilde{U}'_A(a) = \begin{cases} \frac{U'_A(a)}{\max[U'^{-}_A(\overline{a}),\ U'^{-}_A(\overline{a})/z'^{-}(\overline{a})]}, & a \in (0, \underline{a}), \\ 1, & a \in (\underline{a}, \overline{a}). \end{cases}$$

Indeed, note $\tilde{U}_A$ satisfies all the conditions of Lemma 6 and thus $(z, \tilde{U}_A)$ is feasible.

Feasibility. I now check the conditions of Lemma 6 to verify that the triple (z, U_A) defined in the lemma is feasible whenever the boundary z can be implemented. I first show $U'_A, U'_A/z'$ are non-decreasing; note it suffices to check their monotonicity on $[\underline{a}, \overline{a}]$.

Note that U'_A inherits monotonicity from m. Let us then check U'_A/z', or equivalently $\log\left(U'_A/z'\right)$. Consider first smooth intervals of z; there:

$$\begin{aligned} \left(\log(U'_A/z')\right)' &= \log(m)' - \log(z')' \\ &= \max\left[0, \frac{z''}{z'}\right] - \frac{z''}{z'} \geq 0. \end{aligned}$$

Moreover, $U'^{-}_A/z' < U'^{+}_A/z'$ wherever $z'^{-} > z'^{+}$. It remains to verify points where $z'^{-} < z'^{+}$. There:

$$\frac{U'^{+}_A(t)}{z'^{+}(t)} = \frac{m(t)}{z'^{+}(t)} \cdot \frac{z'^{+}(t)}{z'^{-}(t)} = \frac{m(t)}{z'^{-}(t)} = \frac{U'^{-}_A(t)}{z'^{-}(t)}.$$

Hence (a) holds. Furthermore, U'_A and U'_A/z' are piece-wise continuously differentiable as m is continuously differentiable except at the finitely many jump points of z'. The remaining requirements concern only the boundary z.

Optimality. In the remainder of the proof we consider any other feasible $(z, \check{U}_A)$ and show that (z, U_A) give higher welfare. By the above, we can restrict attention to $(z, \check{U}_A)$ such that $\check{U}'_A(a) = 1$ for $a > \overline{a}$ and:

$$\max\left[\check{U}'^{-}_A(\overline{a}),\ \check{U}'^{-}_A(\overline{a})/z'^{-}(\overline{a})\right] = 1. \tag{11}$$

It suffices to show that U'_A lies point-wise above $\check{U}'_A$. Recall that U'_A and $\check{U}'_A$ coincide on $(\bar{a}, 1)$. Moreover, (10) and (11) give:

$$U'^{-}_A(\bar{a}) = \check{U}'^{-}_A(\bar{a}) = \begin{cases} 1/z'^{-}(\bar{a}), & \text{if } z'^{-}(\bar{a}) \geq 1, \\ 1, & \text{if } z'^{-}(\bar{a}) < 1. \end{cases},$$

Since their left-limits $\bar{a}$ coincide, it suffices to show that the following ratio is non-increasing:

$$U'_A \;/\; \check{U}'_A.$$

First, consider intervals on which $\check{U}'_A$ is differentiable and z is smooth. There:

$$\left(\log\left(U'_A/\check{U}'_A\right)\right)' = \left(\log(U'_A)\right)' - \left(\log(\check{U}'_A)\right)'.$$

It suffices to show this derivative is negative. Let us derive the two log terms. Since $\check{U}'_A/z'_A$ is non-increasing, we have:

$$\left(\check{U}'_A/z'\right)' = \frac{\check{U}''_A \cdot z' - \check{U}'_A \cdot z''(a)}{(z')^2} \geq 0 \quad \Rightarrow \quad \left(\log \check{U}'_A\right)' = \frac{\check{U}''_A}{\check{U}'_A} \geq \frac{z''}{z'}.$$

Moreover, we also know U'_A is non-decreasing, so:

$$\left(\log \check{U}'_A\right)' \geq 0.$$

Combining these bounds yields:

$$\left(\log \check{U}'_A\right)' \geq \max\left[0, \frac{z''}{z'}\right]$$

The latter log term is given by:

$$\left(\log U'_A\right)' = \frac{m'}{m} = \max\left[0, \frac{z''}{z'}\right].$$

Thus:

$$\left(\log(U'_A)\right)' - \left(\log(\check{U}'_A)\right)' \leq \max\left[0, \frac{z''}{z'}\right] - \max\left[0, \frac{z''}{z'}\right] = 0.$$

Let us then consider points where z' jumps upwards or $\check{U}'^{-}_A$ jumps downwards. Since $\check{U}'_A/z'$ is non-decreasing, we have:

$$\left(\frac{\check{U}'_A}{z'}\right)^{+} \geq \left(\frac{\check{U}'_A}{z'}\right)^{-} \quad \Rightarrow \quad \frac{\check{U}'^{+}_A}{\check{U}'^{-}_A} \geq \frac{z'^{+}}{z'^{-}}.$$

Moreover,

$$U_A'^{+} = c \cdot m^{+} = c \cdot m^{-} \cdot \frac{z'^{+}}{z'^{-}} = U_A'^{-} \cdot \frac{z'^{+}}{z'^{-}} \quad \Rightarrow \quad \frac{U_A'^{+}}{U_A'^{-}} = \frac{z'^{+}}{z'^{-}}.$$

Thus:

$$\left(\frac{U_A'}{\check{U}_A'}\right)^{-} \geq \left(\frac{U_A'}{\check{U}_A'}\right)^{+}. \tag{12}$$

Finally, consider points where z' jumps downwards. U_A' is continuous at such points, and since $\check{U}_A'$ is increasing, it is either continuous or jumps upwards there. In either case, we get (12).

B.6 Proof of Proposition 3

The optimal boundary is piece-wise linear. I begin by showing that z^* has to solve the following optimal control problem on every closed interval where it is concave and twice continuously differentiable:

Problem 1. *Choose the control $u : [\underline{v}, \overline{v}] \to \mathbb{R}_-$ and state variables $z, y, q : [\underline{v}, \overline{v}] \to \mathbb{R}$ to maximize:*

$$-\int_{\underline{v}}^{\overline{v}} F\left(a, z(a)\right)\, da, \tag{13}$$

subject to the following laws of motion:

$$z'(v) = y(v), \quad y'(v) = u(v), \quad q'(v) = \int_0^{z(v)} f(v,b)db,$$

and the following end-point constraints:

$$z(\underline{v}) = z^*(\underline{v}), \quad z(\overline{v}) = z^*(\overline{v}), \tag{14}$$

$$y(\underline{v}) = z_+^{*\prime}(\underline{v}), \quad y(\overline{v}) = z_-^{*\prime}(\overline{v}), \tag{15}$$

$$q(\underline{v}) = 0, \quad q(\overline{v}) = \int_{\underline{v}}^{\overline{v}} \left(\int_0^{z^*(v)} f(v,b)db\right) dv. \tag{16}$$

The states z and y correspond to the boundary and its derivative, the control u corresponds to its second derivative, and q is introduced to capture the supply constraint.

Lemma 7. *Let $[\underline{v}, \overline{v}]$ with $\underline{v} > \underline{a}$ be such that z^* is twice continuously differentiable with $z^{*\prime\prime} < 0$ on it. Then z^* has to solve Problem 1 on $[\underline{v}, \overline{v}]$.*

Proof. First, note that by Fact 1 z^* is absolutely continuous on $[\underline{v}, \overline{v}]$. It is also concave and twice continuously differentiable on this interval, so it is admissible in Problem 1. By Lemma 3, the optimal U_A for z^* satisfies the following on $[\underline{v}, \overline{v}]$:

$$U_A'(a) = k,$$

for some $k > 0$. Now, take any $z : [\underline{v}, \overline{v}] \to \mathbb{R}$ that is admissible in Problem 1 and consider:

$$\tilde{z}(v) = \begin{cases} z(v) & \text{if } v \in [\underline{v}, \overline{v}], \\ z^*(v) & \text{elsewhere.} \end{cases}$$

Note that, by Lemma 6, the pair $(\tilde{z}, U_A)$ is also implementable. Thus, if z^* is optimal globally, it must be optimal among all $\tilde{z}$ implemented this way. In what follows I show that this is only true if z^* solves Problem 1. Indeed, by Lemma 1, total welfare is:

$$U_A(1) - \int_0^1 U_A'(a) \cdot F\left(a, \hat{z}(a)\right) da = const - k \cdot \int_{\underline{v}}^{\overline{v}} F\left(a, z(a)\right) da.$$

□

I now show $z^{*\prime\prime} = 0$ wherever it exists. Suppose $z^{*\prime\prime} < 0$ somewhere (the case of $z^{*\prime\prime} > 0$ is symmetric). By Fact 1, $z^{*\prime\prime}$ is piece-wise continuous, so there must exist an interval $[\underline{v}, \overline{v}]$, with $\underline{v} > \overline{a}$, on which $z^{*\prime\prime}$ exists and $z^{*\prime\prime} < 0$.

Consider Problem 1 for that interval. As shown, z^* restricted to $[\underline{v}, \overline{v}]$ must be the optimal z for that problem. Let $(z^*, y^*, q^*, u^*, \xi, \phi, \mu)$ be optimal the collection of states, controls and costates associated with z^*. The Hamiltonian for this problem is:

$$\mathcal{H} = -F\left(a, z^*(a)\right) + \mu(a) \cdot \left(\int_0^{z^*(a)} f(a,b)db\right) + \xi(a) \cdot y(a) + \phi(a) \cdot u(a), \tag{17}$$

where $\mu(a)$ is the costate on q, ξ is the costate on z and ϕ is the costate on y. By the Maximum Principle, we then have:

$$\mu'(a) = 0.$$

Since $\mu(a)$ is constant, I will simply write it as μ. Moreover, we have:

$$\xi'(a) = -\left(-\int_0^a f\left(v, z^*(a)\right) dv + \mu \cdot f\left(a, z^*(a)\right)\right) = \left(\int_0^a f\left(v, z^*(a)\right) dv - \mu \cdot f\left(a, z^*(a)\right)\right), \tag{18}$$

and:

$$\phi'(a) = -\xi(a), \tag{19}$$

giving:

$$\phi''(a) = -\xi'(a) = -\left(\int_0^a f\left(v, z^*(a)\right) dv - \mu \cdot f\left(a, z^*(a)\right)\right). \tag{20}$$

The Maximum principle further tells us that controls $u^*(v) < 0$ must maximize the Hamiltonian everywhere in $(\underline{v}, \overline{v})$. However, the Hamiltonian depends on the control linearly and so the optimal control can be interior only if $\phi(v) = 0$ on $(\underline{v}, \overline{v})$. In particular, this means that $\phi''(a)$ has to be zero in that region. This gives:

$$0 = -\int_0^a f\left(v, z^*(a)\right) dv + \mu \cdot f\left(a, z^*(a)\right) \quad \Rightarrow \quad \frac{\int_0^a f\left(v, z^*(a)\right) dv}{f\left(a, z^*(a)\right)} = \frac{F_{A|B}\left(a|z^*(a)\right)}{f_{A|B}\left(a|z^*(a)\right)} = \mu. \tag{21}$$

Note, however, that $z^*(a)$ is strictly increasing in a and, by Assumption 1, the inverse conditional anti-hazard rate is strictly increasing in one of a and z^*, and non-decreasing in the other. Thus, (21) cannot hold on an interval and so $z^{*\prime\prime}(a) = 0$ wherever z^* is twice-differentiable. Since z^* was piece-wise twice continuously differentiable, it follows that z^* is piece-wise linear.

The optimal boundary is linear. We know the optimal boundary z^* is piece-wise linear. Pick $\overline{v}$ such that $[\underline{a}, \underline{v}]$ is the largest interval starting with $\underline{a}$ on which z^* is convex or concave. Assume w.l.o.g. that z^* is concave on it and consider the following problem:

Problem 2. *Choose the control $u : [\underline{a}, \overline{v}] \to \mathbb{R}_-$, state variables $z, y, q : [\underline{a}, \overline{v}] \to \mathbb{R}$, a number of jumps $n \in \mathbb{N}$, jump locations and jump sizes, $a_i \in [\underline{a}, \overline{v}]$ and $v_i \in \mathbb{R}_-$ for $i \in \{1, \ldots, n\}$ to maximize:*

$$-\int_{\underline{a}}^{\overline{v}} F\left(a, z(a)\right) da. \tag{22}$$

subject to the following laws of motion:

$$z'(v) = y(v), \quad y'(v) = u(v), \quad q'(v) = \left(\int_0^{z(v)} f(v,b)db\right),$$

the jump function for every i:

$$v_i = y_+(a_i) - y_-(a_i),$$

and the following end-point constraints:

$$z(\underline{a}) = z^*(\underline{a}), \quad z(\overline{v}) = z^*(\overline{v}), \tag{23}$$

$$y(\underline{a}) \text{ free}, \quad y(\overline{v}) = z^{*\prime}_-(\overline{v}), \tag{24}$$

$$q(\underline{a}) = 0, \quad q(\overline{v}) = \int_{\underline{a}}^{\overline{v}} \left(\int_0^{z^*(v)} f(v,b)db\right) dv. \tag{25}$$

We then get the following lemma whose proof is analogous to that of Lemma 7:

Lemma 8. *Let $[\underline{a}, \overline{v}]$ be a concave interval of z^*. Then the optimal boundary z^* on this interval has to solve Problem 2.*

Since z^* is piece-wise linear, the largest initial concave interval either covers all of $[\underline{a}, \overline{a}]$, or consists of at least two linear pieces. In what follows, I show that the solution to Problem 2 cannot have jumps, and thus that the latter cannot happen. This in turn proves that z^* is linear.

Let us now analyze the necessary conditions for z^* restricted to $[\underline{a}, \overline{v}]$ to solve Problem 2. The Hamiltonian and costate equations for this problem are the same as for Problem 1, and given by (17), (18) and (19), and μ, the costate for q, is also constant. However, the initial value of y is now free, so its costate at the beginning of the interval is zero (see Neustadt (1976), p. 234):

$$\phi(\underline{a}) = 0. \tag{26}$$

Now, by the Maximum Principle with jumps (see Seierstad and Sydsaeter (1986), Theorem 7, p. 196-197) we know that:

1. $\phi(\cdot)$ is continuous and differentiable except possibly at jump points,
2. $\phi(a^*) = 0$ when a^* is a jump point,
3. At all a where there is no jump, $\phi(a) \geq 0$.

Fact 2. *The costate ϕ is twice continuously differentiable on $(\underline{a}, \overline{v})$.*

Proof. For a other than jump points this follows because:

$$\phi'(a) = -\xi(a) = -\xi(\underline{a}) - \int_{\underline{a}}^{a} \xi'(t)dt = -\xi(\underline{a}) - \int_{\underline{a}}^{a} \left(\int_0^t f\left(v, z^*(t)\right) dv - \mu \cdot f\left(t, z^*(t)\right) \right) dt. \quad (27)$$

Now, let $a^* \in (\underline{a}, \overline{v})$ be a jump point. Then (27) holds on some open neighborhoods to the left and right of a^*. We know that $\phi'(a)$ is differentiable there, with:

$$\phi''(a) = -\xi'(a) = -\left(\int_0^a f\left(v, z^*(a)\right) dv - \mu \cdot f\left(a, z^*(a)\right) \right),$$

which, just like $\phi'(a)$, inherits continuity from f and F. Moreover, we see that:

$$\lim_{a \downarrow a^*} \phi'(a) = \lim_{a \uparrow a^*} \phi'(a), \quad \lim_{a \downarrow a^*} \phi''(a) = \lim_{a \uparrow a^*} \phi''(a),$$

where these limits are finite. Thus, $\phi'(a^*), \phi''(a^*)$ also exist and equal to these limits, and so ϕ is indeed twice continuously differentiable on $(\underline{a}, \overline{v})$. $\square$

Point 2 above tells us that if there is an interior jump at $a^* \in (\underline{a}, \overline{v})$, we must have $\phi(a^*) = 0$ there. Point 3 tells is that $\phi(a) \geq 0$ outside of jump points. Since, by Fact 2, ϕ is twice continuously differentiable around a^*, we must therefore have $\phi'(a^*) = 0$ and $\phi''(a^*) \geq 0$ there. I will show this cannot happen, and thus that z^* cannot have jumps. Note that:

$$\phi''(a) \; = \; f\left(a, z^*(a)\right) \left(\mu - \frac{\int_0^a f\left(v, z^*(a)\right) dv}{f\left(a, z^*(a)\right)} \right) \; = \; f\left(a, z^*(a)\right) \left(\mu - \frac{F_{A|B}\left(a|z^*(a)\right)}{f_{A|B}\left(a|z^*(a)\right)} \right).$$

Recall that $z^*(a)$ is strictly increasing in a, so by Assumption 1, the inverse conditional anti-hazard rate is strictly increasing in a. Thus, $\phi''(a)$ is either strictly negative everywhere on $(\underline{a}, \overline{v})$ or positive until some $\tilde{a} \in (\underline{a}, \overline{v})$ and then negative forever after. In the former case, $\phi''(a) < 0$ for all $a \in (\underline{a}, \overline{v})$, and so $\phi''(a^*) \geq 0$ can never hold for an interior a^*. Let us then consider the latter case in which there exists some $\tilde{a} \in (\underline{a}, \overline{v})$ such that:

$$\phi''(a) \begin{cases} > 0, & \text{if } a < \tilde{a}, \\ = 0, & \text{if } a = \tilde{a}, \\ < 0, & \text{if } a > \tilde{a}. \end{cases}$$

Now, I show that for every $a \in (\underline{a}, \tilde{a}]$ we have $\phi'(a) > 0$. For suppose $\phi'(a) \leq 0$ for some $a \in (\underline{v}, \tilde{a}]$. Then, since $\phi''(a) > 0$ on $(\underline{a}, \tilde{a})$, it must be that $\phi'(a) < 0$ everywhere before a. Then, however:

$$\phi(a) = \underbrace{\phi(\underline{a})}_{=0} + \int_{\underline{a}}^{a} \underbrace{\phi'(t)}_{<0} \, dt < 0,$$

which cannot be as $\phi(a) \geq 0$ on the whole interval by the Maximum Principle with jumps. Thus, $\phi'(a) > 0$ for $a \in (\underline{a}, \tilde{a}]$ and $\phi''(a) < 0$ for all $a \in (\tilde{a}, \overline{v})$. Therefore there is no $a^* \in (\underline{a}, \overline{a})$ for which $\phi'(a^*) = 0$ and $\phi''(a^*) \geq 0$.

B.7 Proof of Lemma 4

Fix any linear boundary $\hat{z}$. Throughout the proof we show that if the slope of $\hat{z}$ is not 1, it can be improved by changing the slope in the direction of 1 while preserving the total amounts of both goods allocated. To that end, let $z_s : [\underline{a}_s, \underline{b}_s] \to [\overline{a}_s, \overline{b}_s]$ denote a linear boundary with slope $s > 0$ under which the same amounts of A and B are allocated as under $\hat{z}$. Thus:

$$z_s(a) = \underline{b}_s + s \cdot (a - \underline{a}_s) \quad \text{for } a \in [\underline{a}_s, \overline{a}_s].$$

See Figure 12a. Note $\underline{a}_s$ and $\underline{b}_s$ are uniquely pinned down by the requirement that probability masses above and below z_s match those above and below $\hat{z}$. Note that for every $s_1, s_2 > 0$ such that $s_1 > s_2$, the boundaries z_{s_1} and z_{s_2} cross exactly once, with z_{s_2} crossing from above (Figure 12b). This in turn implies that $\underline{a}_s, \overline{b}_s$ increase in s and $\overline{a}_s, \underline{b}_s$ decrease in s. Moreover, the regularity of the density f guarantees that all four end-points change continuously in s.

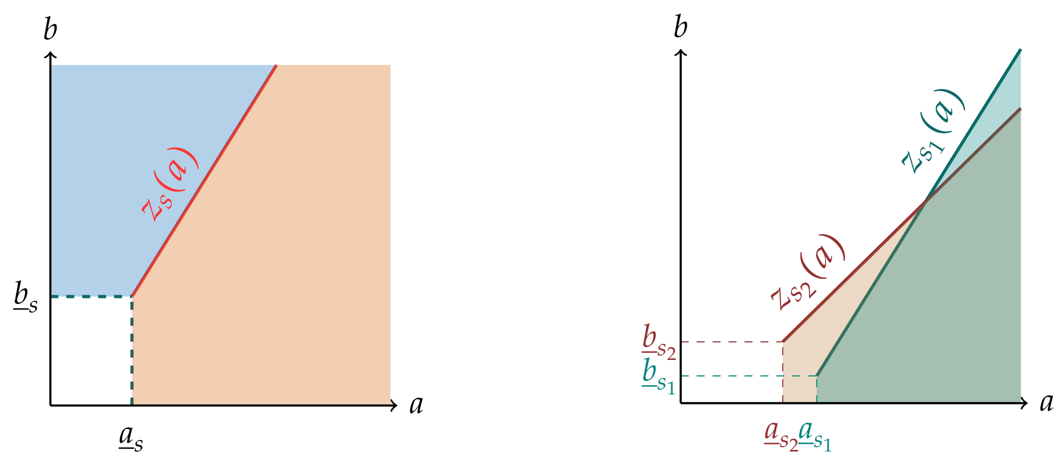

Figure 12a: s-sloped boundary z_s.

Figure 12b: z_{s_2} crosses z_{s_1} once, and from above.

We will take $W[z_s]$ to mean welfare under (z_s, U_A), where U_A is optimal for z_s. It thus suffices to show that $W[z_s]$ decreases in s when $s > 1$ and increases in s when $s < 1$. We show this using the following auxiliary fact:

Fact 3. *The following expression is strictly increasing in s for s > 1:*

$$\int_0^1 F\left(a, \hat{z}_s(a)\right)\, da.$$

Proof. First, note that

$$\begin{aligned}
F\big(a, \hat{z}_{s_1}(a)\big) - F\big(a, \hat{z}_{s_2}(a)\big) &= \int_0^{\hat{z}_{s_1}(a)} \int_0^a f(v,b)\, dv\, db \;-\; \int_0^{\hat{z}_{s_2}(a)} \int_0^a f(v,b)\, dv\, db \\
&= \int_{\hat{z}_{s_2}(a)}^{\hat{z}_{s_1}(a)} \int_0^a f(v,b)\, dv\, db.
\end{aligned}$$

Now, consider the difference:

$$\begin{aligned}
\int_0^1 F\left(a, \hat{z}_{s_1}(a)\right)\, da - \int_0^1 F\left(a, \hat{z}_{s_2}(a)\right)\, da &= \int_0^1 F\left(a, \hat{z}_{s_1}(a)\right) - F\left(a, \hat{z}_{s_2}(a)\right)\, da \\
&= \int_0^1 \left(\int_{\hat{z}_{s_2}(a)}^{\hat{z}_{s_1}(a)} \int_0^a f(v,b) dv\, db\right)\, da.
\end{aligned}$$

Since $s_1 > s_2$, $\hat{z}_{s_2}$ crosses $\hat{z}_{s_1}$ only once, and from above. Let $a^* \in (0,1)$ be their crossing point and define:

$$\overline{\mathcal{D}} = \left\{(a,b)\colon\ a^* < a < 1,\ \hat{z}_{s_1}(a) < b < \hat{z}_{s_2}\right\}, \quad \underline{\mathcal{D}} = \left\{(a,b)\colon\ 0 < a < a^*,\ \hat{z}_{s_2}(a) < b < \hat{z}_{s_1}(a)\right\}.$$

We can then write the difference as:

$$\begin{aligned}
&\int_0^1 F\left(a, \hat{z}_{s_1}(a)\right)\, da - \int_0^1 F\left(a, \hat{z}_{s_2}(a)\right)\, da \\
&= \int_{a^*}^1 \left(\int_{\hat{z}_{s_2}(a)}^{\hat{z}_{s_1}(a)} \int_0^a f(v,b)\, dv\, db\right) da \;+\; \int_0^{a^*} \left(\int_{\hat{z}_{s_2}(a)}^{\hat{z}_{s_1}(a)} \int_0^a f(v,b)\, dv\, db\right) da \\
&= \int_{a^*}^1 \left(\int_{\hat{z}_{s_2}(a)}^{\hat{z}_{s_1}(a)} \int_0^a f(v,b)\, dv\, db\right) da \;-\; \int_0^{a^*} \left(\int_{\hat{z}_{s_1}(a)}^{\hat{z}_{s_2}(a)} \int_0^a f(v,b)\, dv\, db\right) da \\
&= \int_{\overline{\mathcal{D}}} \left(\int_0^a f(v,b)\, dv\right) \mathrm{d}(a,b) \;-\; \int_{\underline{\mathcal{D}}} \left(\int_0^a f(v,b)\, dv\right) \mathrm{d}(a,b) \\
&= \int_{\overline{\mathcal{D}}} \frac{\int_0^a f(v,b)\, dv}{f(a,b)} \cdot f(a,b)\, \mathrm{d}(a,b) \;-\; \int_{\underline{\mathcal{D}}} \frac{\int_0^a f(v,b)\, dv}{f(a,b)} \cdot f(a,b)\, \mathrm{d}(a,b) \\
&= \int_{\overline{\mathcal{D}}} \frac{F_{A|B}(a \mid b)}{f_{A|B}(a \mid b)} \cdot f(a,b)\, \mathrm{d}(a,b) \;-\; \int_{\underline{\mathcal{D}}} \frac{F_{A|B}(a \mid b)}{f_{A|B}(a \mid b)} \cdot f(a,b)\, \mathrm{d}(a,b).
\end{aligned}$$

Now, recall that the inverse conditional anti-hazard rate is strictly increasing in a and non-decreasing in b. Let $b^* := \hat{z}_{s_1}(a^*) = \hat{z}_{s_2}(a^*)$ and notice that, by single-crossing:

$$(a,b) > (a^*,b^*) \quad \text{for all } (a,b) \in \overline{\mathcal{D}}, \qquad (a,b) < (a^*,b^*) \quad \text{for all } (a,b) \in \underline{\mathcal{D}}.$$

Thus, we can bound the difference from below as follows:

$$\begin{aligned}
&\int_0^1 F\left(a,\hat{z}_{s_1}(a)\right)\, da - \int_0^1 F\left(a,\hat{z}_{s_2}(a)\right)\, da \\
&> \int_{\overline{\mathcal{D}}} \frac{F_{A|B}(a^* \mid b^*)}{f_{A|B}(a^* \mid b^*)} f(a,b)\,\mathrm{d}(a,b) \;-\; \int_{\underline{\mathcal{D}}} \frac{F_{A|B}(a^* \mid b^*)}{f_{A|B}(a^* \mid b^*)} f(a,b)\,\mathrm{d}(a,b) \\
&= \frac{F_{A|B}(a^* \mid b^*)}{f_{A|B}(a^* \mid b^*)} \left(\int_{\overline{\mathcal{D}}} f(a,b)\,\mathrm{d}(a,b) - \int_{\underline{\mathcal{D}}} f(a,b)\,\mathrm{d}(a,b) \right) = 0.
\end{aligned}$$

The difference is zero because the probability masses underneath both extended boundaries were equal. □

Recall that by Lemma 1, total welfare equals to:

$$U_A(1) - \int_0^1 U_A'(a) \cdot F\left(a,\hat{z}(a)\right) da. \tag{28}$$

Now, consider two cases.

Case 1: $s > 1$. In this case, the optimal U_A for z_s satisfies:

$$U_A'(a) = \begin{cases} 0, & \text{if } a < \underline{a}_s, \\ 1, & \text{if } a > \underline{a}_s. \end{cases}$$

Thus, total welfare is:

$$W[z_s] = 1 - \underline{a}_s - \int_0^1 F\left(a,\hat{z}_s(a)\right)\, da.$$

Now, since $\overline{a}_s, \overline{b}_s$ and $\overline{a}_s$ change continuously in s and $\overline{a}_{s_1} < 1$, there exists $s_2 \in (1, s_1)$ such that $\overline{a}_{s_2} < 1$. Then also $\overline{b}_{s_2} = 1$ and thus we can apply the formula derived above to both z_{s_1} and z_{s_2}:

$$W[z_{s_1}] = 1 - \underline{a}_{s_1} - \int_0^1 F\left(a,\hat{z}_{s_1}(a)\right)\, da, \quad W[z_{s_2}] = 1 - \underline{a}_{s_2} - \int_0^1 F\left(a,\hat{z}_{s_2}(a)\right)\, da.$$

It thus suffices to show that $W[z_{s_1}] < W[z_{s_2}]$. However, $s_1 > s_2$, so $\underline{a}_{s_1} > \underline{a}_{s_2}$. Moreover, Fact 3 tells us that

$$\int_0^1 F\left(a,\hat{z}_{s_1}(a)\right)\, da > \int_0^1 F\left(a,\hat{z}_{s_2}(a)\right)\, da,$$

which completes the proof.

Case 2: $s < 1$. Since $U_A'(a) = 1$ for $a > \bar{a}_s$, we can write total welfare as:

$$W[z_s] = \int_{\underline{a}_s}^{\bar{a}_s} U_A'(a) \cdot \left[1 - F\left(a, z(a)\right)\right] da + \int_{\bar{a}_s}^{1} \left[1 - F(a, 1)\right] da$$

Recall also that for $a \in (\underline{a}_s, \bar{a}_s)$, we have:

$$U_A'(a) = U_B'(z(a)) \cdot z'(a), \quad U_A'(a) = z'(a) = s, \quad U_B'(z(a)) = 1.$$

Thus:

$$W[z_s] = \int_{\underline{a}_s}^{\bar{a}_s} U_B'(z(a)) \cdot z'(a) \cdot \left[1 - F\left(a, z(a)\right)\right] da + \int_{\bar{a}_s}^{1} \left[1 - F(a, 1)\right] da.$$

Changing variables in the first integral gives:

$$W[z_s] = \int_{\underline{b}_s}^{\bar{b}_s} U_B'(b) \left[1 - F\left(z^{-1}(b), b\right)\right] db + \int_{\bar{a}_s}^{1} \left[1 - F(a, 1)\right] da.$$

Now, by Lemma 3 we have the following when $s < 1$:

$$U_A'(a) = \begin{cases} 0, & \text{if } a < \underline{a}_s, \\ s, & \text{if } a \in (\underline{a}_s, \bar{a}_s), \\ 1, & \text{if } a > \underline{a}_s. \end{cases}$$

Moreover, recall that for $a \in (\underline{a}_s, \bar{a}_s)$:

$$U_B'(z(a)) = \frac{U_A'(a)}{z'(a)} = \frac{s}{s} = 1.$$

We therefore have $U_B'(b) = 1$ on $(\underline{b}_s, \bar{b}_s)$. We can then write:

$$\begin{aligned} W[z_s] &= \int_{\underline{b}_s}^{\bar{b}_s} \left[1 - F\left(z^{-1}(b), b\right)\right] db + \int_{\bar{a}_s}^{1} \left[1 - F(a, 1)\right] da \\ &= \bar{b}_s - \underline{b}_s - \int_0^1 F\left(\hat{z^{-1}}(b), b\right) db + \int_{\bar{a}_s}^{1} \left[1 - F(a, 1)\right] da, \end{aligned}$$

Now, $\bar{b}_s$ increases in s and $\underline{b}_s, \bar{a}_s$ decrease in s. Finally, note that an increase in the slope of z_s leads to a decrease in the slope of z_s^{-1}. Thus, the second term decreases in s by a result symmetric to Fact 3. Thus, $W[z_s]$ increases in s when $s < 1$.

B.8 Proof of Lemma 5

By Lemma 4, we can restrict attention to mechanisms with a boundary of slope 1. Lemma 3 tells us that, under the optimal U_A, such mechanisms offer only two options: good A with $x = 1$ at an ordeal c_A and good B with $x = 1$ at an ordeal c_B. Now, suppose one of the supply constraints

(S) is slack for such a mechanism; assume without loss this is the case for good A.

Since the mechanism allocates a strictly positive amount of good A but its supply constraint is slack, the ordeal for good A has to be interior: $c_A \in (0,1)$. Now, consider an alternative mechanism with ordeals $c_A - \epsilon$ and c_B, respectively. This mechanism improves the utilities of all agents, and strictly so for the positive mass of agents who chose good A under the original mechanism. It therefore suffices to show the alternative mechanism is feasible for $\epsilon > 0$ sufficiently small. Note that the mass of agents who take A under the new mechanism is:

$$\int \mathbb{1}\left[a - (c_A - \epsilon) > \max[0,\ b - c_B]\right]\, dF(a,b),$$

since the set of indifferent agents is zero-mass. Moreover:

$$\lim_{\epsilon \to 0} \int \mathbb{1}\left[a - (c_A - \epsilon) > \max[0,\ b - c_B]\right]\, dF(a,b) = \int \mathbb{1}\left[a - c_A > \max[0,\ b - c_B]\right]\, dF(a,b),$$

which is the mass of agents who got good A under the original mechanism. Since the supply constraint (S) for good A was slack, it remains slack for the alternative one when ϵ is sufficiently small. Similarly, reducing the ordeal for good A can only relax the supply constraint for good B, and thus (S) is satisfied for ϵ small enough.

B.9 Proof of Proposition 4

Maximizing (E) is equivalent to choosing $q_A, q_B : [0,1]^2 \to [0,1]$ to maximize:

$$\int q_A(a,b) \cdot a + q_B(a,b) \cdot b \ \mathrm{d}F(a,b),$$

subject to:

$$\int q_A(a,b)\,\mathrm{d}F(a,b) \le \mu_A, \quad \int q_B(a,b)\,\mathrm{d}F(a,b) \le \mu_B, \tag{29}$$

$$q_A(a,b) + q_B(a,b) \le 1 \ \text{ for every } (a,b) \in [0,1]^2. \tag{30}$$

Since $\mu_A + \mu_B \le 1$ and a unit mass of types has positive values for both goods, both supply constraints (29) will hold with equality. The objective and constraints are linear so the solution exists and must also maximize:

$$\int q_A(a,b) \cdot (a - \eta_A) + q_B(a,b) \cdot (b - \eta_B) \ \mathrm{d}F(a,b), \tag{31}$$

subject to (30) for some multipliers $\eta_A, \eta_B \ge 0$. Note also that $\eta_A, \eta_B < 1$. Otherwise, the maximizer of (31) would not allocate one of the goods at all, and we know that supply constraints must hold with equality. Now, notice that q_A, q_B maximize (31) if and only if they satisfy the following almost everywhere:

$$q_A(a,b) = \begin{cases} 1, & \text{if } a - \eta_A > \max\{0,\ b - \eta_B\}, \\ 0, & \text{otherwise,} \end{cases} \quad , \quad q_B(a,b) = \begin{cases} 1, & \text{if } b - \eta_B > \max\{0,\ a - \eta_A\}, \\ 0, & \text{otherwise.} \end{cases}$$

Such an allocation is implemented by a mechanism with no damages and two posted prices equal to η_A, η_B. Finally, since individual demands satisfy the gross substitutes condition, these prices are unique (Kelso Jr and Crawford, 1982; Gul and Stacchetti, 1999). Since Theorem 1 offered each good with an ordeal and allocated the whole supply of both goods, the two ordeals must therefore have been equal to the unique market-clearing prices.

B.10 Proof of Corollary 2

The argument is analogous to the proof of Theorem 1. First, note that the steps related to implementation, i.e. Proposition 2 and Lemma 2, transfer without any modification as they do not rely on the planner's objective. The proof of Lemma 3 also carries through, as it shows that the optimal implementation guarantees the highest utility profile *point-wise*, and thus the argument is unaffected by the presence of welfare weights. Lemma 5 remains valid for the same reason.

Lemma 1, however, requires modification. As before, Proposition 1 lets us rewrite total welfare (WW) in terms of A, B-indirect utilities U_A, U_B and their associated boundary z as follows:

$$\int_{\underline{a}}^{1} \int_{0}^{z(\min[a,\bar{a}])} \lambda(a,v) \cdot g(a,v) dv \cdot U_A(a)\, da + \int_{\underline{b}}^{1} \int_{0}^{z^{-1}(\min[b,\bar{b}])} \lambda(a,v) \cdot g(v,b) dv \cdot U_B(b)\, db. \tag{WW'}$$

A sequence of steps analogous to those in the proof of Lemma 1 then yields the following expression for the objective:

$$U_A(1) - \int_0^1 U'_A(a) \cdot \tilde{G}(a, \hat{z}(a)) da, \tag{WW*}$$

where:

$$\tilde{G}(a,b) := \frac{\int_{[0,a]\times[0,b]} \lambda(v,w) \cdot g(v,w)\ \mathrm{d}(v,w)}{\int_{[0,1]^2} \lambda(v,w) \cdot g(v,w)\ \mathrm{d}(v,w)}.$$

A similar modification is required in the proof of Proposition 3. Note that while the objective is then written in terms of $\tilde{G}$, the area constraint is still phrased in terms of the unmodified density g. Following the steps of the derivation then yields the analog of (20):

$$\phi''(a) = -\left(\int_0^a \lambda(v, z^*(a)) \cdot g\left(v, z^*(a)\right) dv - \mu \cdot g\left(a, z^*(a)\right) \right). \tag{32}$$

The rest of the original argument carries through exactly, only with the expression

$$\frac{\int_0^a f\left(v, z^*(a)\right) dv}{f\left(a, z^*(a)\right)} = \frac{F_{A|B}\left(a|z^*(a)\right)}{f_{A|B}\left(a|z^*(a)\right)},$$

now replaced with

$$\frac{\int_0^a \lambda(v, z^*(a)) \cdot g\left(v, z^*(a)\right) dv}{g\left(a, z^*(a)\right)}.$$

A similar modification must be made to Lemma 4; the following expression in the original proof

$$\int_{\overline{\mathcal{D}}} \left(\int_0^a g(v,b)\, dv \right) \mathrm{d}(a,b) \; - \; \int_{\underline{\mathcal{D}}} \left(\int_0^a g(v,b)\, dv \right) \mathrm{d}(a,b),$$

is now replaced with

$$\int_{\overline{\mathcal{D}}} \left(\int_0^a \lambda(v,b) \cdot g(v,b)\, dv \right) \mathrm{d}(a,b), \; - \; \int_{\underline{\mathcal{D}}} \left(\int_0^a \lambda(v,b) \cdot g(v,b)\, dv \right) \mathrm{d}(a,b),$$

which can in turn be written as:

$$\int_{\overline{\mathcal{D}}} \frac{\int_0^a \lambda(v,b) \cdot g(v,b)\, dv}{g(a,b)} \cdot g(a,b)\, \mathrm{d}(a,b) \; - \; \int_{\underline{\mathcal{D}}} \frac{\int_0^a \lambda(v,b) \cdot g(v,b)\, dv}{g(a,b)} \cdot g(a,b)\, \mathrm{d}(a,b).$$

The rest of the argument mirrors the original one.