CAN WITHHOLDING OR DAMAGE IMPROVE WELFARE IN BILATERAL TRADE MECHANISMS?

MIHAI MANEA* AND ERIC MASKIN†

Abstract. We study the welfare consequences of allowing the mechanism designer to withhold or damage resources in the optimal mechanism for bilateral trade with independent private values. We show that withholding monetary transfers or withholding the good from both traders is never optimal. Similarly, damaging the good for the buyer cannot improve welfare. By contrast, damaging the good for the seller may improve welfare in the optimal mechanism. However, such welfare improvements are feasible only if the damage hurts seller types with lower initial valuations more severely.

1. Introduction

A major theme of the literature on mechanism design with asymmetric information is that to achieve an optimal allocation of resources, it may be necessary to *deliberately* introduce distortions that would be inefficient in a complete information setting. That is, it is optimal to violate *first-best* efficiency in order to relax incentive and individual rationality constraints.

For example, consider a seller who has one unit of an indivisible good to which he attaches no value facing a buyer whose valuation for the good is equally likely to be 3 or 7. In the mechanism that maximizes expected revenue, the seller sets a price of 7, and the buyer acquires the good when his valuation is 7, but not when it is 3. This outcome is inefficient because either buyer type derives greater value from the good than the seller. The inefficiency arises because by not trading with the low-value buyer, the seller is able to demand a high price from the high-value buyer. If the valuation of the high-type buyer were reduced from 7 to 5, then the revenue-maximizing mechanism prescribes that the seller trade with both buyer types at price 3. The reduction in the value of the high-type buyer not only increases the probability of an efficient allocation under the revenue-maximizing mechanism from 1/2 to 1, but also increases total expected welfare (the expected sum of buyer's and seller's payoffs) from 7/2 to 4.

Going a step further, Laffont and Maskin (1979) and Myerson and Satterthwaite (1983) exhibit models in which, because of asymmetric information, there is *no* mechanism that

*Stony Brook University; mihai.manea@stonybrook.edu
†Harvard University; emaskin@fas.harvard.edu
Date: October 2, 2020.

is first-best efficient. For an illustration, consider the following instance of the Myerson-Satterthwaite model of bilateral trade with incomplete information. There is a seller with an indivisible good who has a valuation that is either $s_1 \in [0, 1)$ or 3 with equal probability, and a buyer with possible valuations 1 and 4 also with equal probability, where buyer and seller types are independently distributed. First-best efficiency dictates that the good be sold with probability 1 for the pairs of buyer and seller types $(1, s_1), (4, s_1), (4, 3)$ and with probability 0 for the pair $(1, 3)$. However, there is no mechanism consistent with the buyer's and the seller's incentive and individual rationality constraints that delivers these outcomes if $s_1 \in (0, 1)$. The *second-best* welfare optimal mechanism entails that for the pair of types $(1, s_1)$, the good is traded with probability $2/(2 + s_1)$. In this example, reducing the value s_1 of the low-type seller increases the probability that the good is allocated efficiently in the optimal mechanism, and also enhances total welfare if $s_1 \leq \sqrt{6} - 2$.

The examples above suggest that reducing valuations and, more generally, creating inefficiencies at the ex ante stage of a mechanism may improve ex post allocative efficiency and social welfare. In this paper, we explore the possibility of such welfare improvements in the Myerson-Satterthwaite bilateral trade setting when the mechanism designer has access to two types of inefficient actions: withholding and damage.

Under *withholding*, we imagine that when the good is transferred by the seller, it does not necessarily reach the buyer: the mechanism designer can withhold or destroy the good with some probability. Similarly, the mechanism designer has the option to withhold part of the monetary transfer from the buyer to the seller (i.e., run an ex post budget surplus).

The general principle behind creating inefficiencies is that they may enhance a mechanism's ability to discriminate between types. Since our model (as much of the mechanism design literature) assumes quasilinear preferences, there is no difference between types in their marginal utility for money. Hence, there should be no welfare gain from withholding money, and this intuition is borne out by our Proposition 1.

More interestingly, Proposition 2 establishes that withholding the good never improves welfare either. To understand the issues involved in establishing this result, note that if a mechanism withholds the good for some pair of buyer and seller types, the natural path to a welfare improvement requires increasing the probability with which the good is allocated to either of these types. However, this perturbation may result in a violation of the monotonicity of each trader's probability of being allocated the good with respect to his type, which is a necessary condition for implementation. In this situation, we show how the allocation can be further perturbed via a sequence of changes in trading probabilities that has a neutral effect on the overall probability that each buyer and seller type affected by the changes is allocated the good until we eventually reach a type for which the monotonicity condition is not binding. With suitable adjustments in monetary transfers, the perturbed allocation is implementable and improves welfare over the original one.

When the good is deliberately *damaged*, it is reduced in value for one or more buyer or seller types. This may entail physical damage that is intentional or exogenous. Regulations that restrict the use of the good or prohibit bundling may also reduce a trader's valuation and have idiosyncratic effects on types (whereas property taxes reduce valuations uniformly across types). For a concrete example, suppose that the good being sold is a private airplane that has a navigational device attuned to the seller's geographic region. The buyer lives in a different region, and removing the device will not affect him at all. However, removing it will impact a seller with low valuation more because higher-valuation seller types are intensely invested in flying and have other equipment that could substitute for the missing device. Similarly, a broker handling an estate or an art or wine collection can decide which items should be included for sale as a package, and dictate what the seller should do with the excluded items prior to the sale.

Proposition 3 shows that damaging the good for the buyer is never socially optimal. The intuition is that reducing buyer value entails less scope for allocating the good to the buyer and thus decreases potential gains from trade.

In Proposition 4, by contrast, we demonstrate that damaging the good for the seller may enhance social welfare for a rich set of value distributions, thus expanding the conclusion of our second opening example. The idea is that a seller type whose valuation has been reduced finds it less attractive to pretend to be a higher-type seller. The slack created in the underlying incentive constraint can be leveraged to increase the probability of trade for seller types with reduced valuations. This is true even if the seller is given the option to withdraw from the mechanism and consume the *undamaged* good. In the context of the examples above, Proposition 4 implies that it may be valuable to employ a mechanism designer who is more "active" than usually considered.

However, Proposition 4 relies on the mechanism designer being able to damage the good to a greater extent for seller types who value it less. In some situations, the mechanism designer may be constrained to damage the good in a fashion that affects higher-valuation seller types more (e.g., proportionately). Then, Proposition 5 reverses the conclusion of Proposition 4 (under the maintained assumption that the seller has the right to consume the undamaged good): damage that hurts seller types with higher valuations more severely cannot improve welfare. The proof of Proposition 5 shows that even though reduced valuations for lower seller types decrease the information rents commanded by the seller as highlighted in Proposition

4, the transfers the designer saves on lower types by damaging the good are not sufficient to compensate higher seller types to participate in the mechanism and risk being allocated the damaged good.

There are a few contributions related to our paper. Cramton, Gibbons, and Klemperer (1987) show that the tension between incomplete information and allocative efficiency highlighted by the impossibility result of Myerson and Satterthwaite (1983) can be remediated by dispersing the initial property rights among agents. Matsuo (1989) and Kos and Manea (2009) dispense with another key ingredient for the Myerson-Satterthwaite impossibility theorem—the assumption that agents have continuous value distributions—and characterize the discrete value distributions for which an ex post efficient allocation is implementable. Makowski and Mezzetti (1993) point out that the impossibility theorem is not robust to the introduction of a second buyer. Postlewaite (1979) shows that agents can manipulate the competitive allocation in their favor by withholding or destroying endowments in an exchange economy. Condorelli and Szentes (2019) analyze a game between a buyer who can stochastically reduce his value for a good and a seller who sets the price after observing the value distribution selected by the buyer. In their game, the buyer strategically introduces incomplete information and inefficiency in order to appropriate some of the surplus resulting from monopoly pricing.

The rest of the paper is organized as follows. In the next section, we set up our model. Sections 3 and 4 cover the results for withholding monetary transfers and the good, respectively. In Sections 5 and 6, we present the results related to damaging the good for the buyer and for the seller, respectively. Proofs are relegated to the Appendix.

2. Framework

We consider a discrete-value version of the bilateral trade problem with incomplete information modeled by Myerson and Satterthwaite (1983), in which a *buyer* is interested in acquiring an indivisible *good* that a *seller* owns. The *valuations* of the buyer and of the seller for the good are independently distributed random variables with *probability mass functions* p_b and p_s, respectively, which have corresponding *supports* $V_b = \{b_1, b_2, \ldots, b_m\}$ and $V_s = \{s_1, s_2, \ldots, s_n\}$, where $0 \leq b_1 < b_2 < \ldots < b_m$ and $0 \leq s_1 < s_2 < \ldots < s_n$. Each player is privately informed about his own value for the good (his *type*) and believes that the other player's value is a random variable drawn from the specified distribution.

The two players are risk-neutral and have additively separable utility functions for money and the value of the good. The mechanism designer needs to specify a game that determines the probability of allocating the good to each of the two players and the monetary transfers sent or received by the players. The objective of the mechanism designer is to maximize expected total welfare. By the revelation principle, we can focus the analysis on direct mechanisms without loss of generality. In a *direct mechanism*, the players simultaneously report their values, and the outcome is determined by four functions (x_b, x_s, t_b, t_s) with $x_b, x_s : V_b \times V_s \to [0, 1]$ and $t_b, t_s : V_b \times V_s \to \mathbb{R}$. If the buyer reports value b_i and the seller reports value s_j, then $x_b(b_i, s_j)$ and $x_s(b_i, s_j)$ are the probabilities with which the good is *allocated* to the buyer and to the seller, respectively, and $t_b(b_i, s_j)$ is the *monetary transfer sent* by the buyer, whereas $t_s(b_i, s_j)$ is the *monetary transfer received* by the seller.

This formulation departs from the Myerson-Satterthwaite model by allowing the mechanism designer to withhold both money and the good: $t_b(b_i, s_j) - t_s(b_i, s_j)$ is the amount of money withheld, and $1 - x_b(b_i, s_j) - x_s(b_i, s_j)$ is the probability that the good is withheld by the mechanism designer for a profile of reports (b_i, s_j). The mechanism designer does not bring any physical resources to the market, and *feasibility* requires that $t_b(b_i, s_j) - t_s(b_i, s_j) \geq 0$ and $x_b(b_i, s_j) + x_s(b_i, s_j) \leq 1$.

Recall that a direct mechanism is *incentive compatible* if honest reporting of values forms a Bayesian Nash equilibrium of the game between the buyer and the seller induced by the mechanism. A mechanism is *individually rational* if each player type obtains non-negative expected utility gains from participating in the mechanism. We say that a mechanism (x_b, x_s, t_b, t_s) is *implementable* if it is feasible, individually rational, and incentive compatible. An allocation (x_b, x_s) is *implementable* if there exist transfer functions (t_b, t_s) such that (x_b, x_s, t_b, t_s) is an implementable mechanism. The designer seeks to develop an implementable mechanism that maximizes *total welfare*—the sum of expected utility for the buyer and the seller. Employing the notation

$$\bar{x}_b(b_i) = \sum_{j=1}^{n} p_s(s_j) x_b(b_i, s_j); \qquad \bar{t}_b(b_i) = \sum_{j=1}^{n} p_s(s_j) t_b(b_i, s_j)$$

$$\bar{x}_s(s_j) = \sum_{i=1}^{m} p_b(b_i) x_s(b_i, s_j); \qquad \bar{t}_s(s_j) = \sum_{i=1}^{m} p_b(b_i) t_s(b_i, s_j)$$

for the expected outcomes relevant for a buyer of type b_i and a seller of type s_j, the *optimal implementable mechanism* (x_b, x_s, t_b, t_s) solves the following linear program:[1]

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{m} p_b(b_i)(\bar{x}_b(b_i)b_i - \bar{t}_b(b_i)) + \sum_{j=1}^{n} p_s(s_j)(\bar{x}_s(s_j)s_j + \bar{t}_s(s_j)) \\
\text{s.t.} \quad & IR_{b_i} : \bar{x}_b(b_i)b_i - \bar{t}_b(b_i) \geq 0, \forall i = \overline{1,m} \\
& IR_{s_j} : \bar{t}_s(s_j) - (1 - \bar{x}_s(s_j))s_j \geq 0, \forall j = \overline{1,n} \\
& IC_{b_i \to b_k} : \bar{x}_b(b_i)b_i - \bar{t}_b(b_i) \geq \bar{x}_b(b_k)b_i - \bar{t}_b(b_k), \forall i, k = \overline{1,m} \\
& IC_{s_j \to s_k} : \bar{t}_s(s_j) - (1 - \bar{x}_s(s_j))s_j \geq \bar{t}_s(s_k) - (1 - \bar{x}_s(s_k))s_j, \forall j, k = \overline{1,n} \\
& FT_{b_i,s_j} : t_b(b_i, s_j) - t_s(b_i, s_j) \geq 0, \forall i = \overline{1,m}, j = \overline{1,n} \\
& FX_{b_i,s_j} : x_b(b_i, s_j) + x_s(b_i, s_j) \leq 1, \forall i = \overline{1,m}, j = \overline{1,n}.
\end{aligned}
$$

Myerson and Satterthwaite (1983) assume that the mechanism designer cannot withhold either money or the good, which corresponds to imposing the feasibility constraints FT_{b_i,s_j} and FX_{b_i,s_j} with equality for all i and j. This assumption restricts the set of mechanisms the designer can implement. The next two sections investigate whether the designer can implement mechanisms that achieve greater total welfare when withholding is feasible.

3. Withholding Monetary Transfers

We first show that withholding monetary transfers cannot enhance welfare in the optimal mechanism.

Proposition 1. *It is never optimal to withhold monetary transfers.*

The proof of this result is presented in the Appendix along with all other proofs. The argument is straightforward. If an implementable mechanism involves money withholding for a pair (b_i, s_j), we can increase the transfer to seller type s_j to match the transfer from b_i when the buyer reports type b_i, and reduce the transfers s_j receives from all buyer types uniformly and credit them to the buyer, so that s_j enjoys the same expected outcomes following the perturbation. The perturbed mechanism preserves individual rationality and incentive compatibility, and increases the utility of the buyer without affecting the utility of the seller.

[1]The notation $\overline{a,b}$ stands for the sequence of integers in the interval $[a, b]$.

4. Withholding the Good

The previous section shows that in solving the linear program for the optimal mechanism, we can assume that the FT_{b_i,s_j} constraints hold with equality, i.e., $t_b(b_i, s_j) = t_s(b_i, s_j)$ for all (b_i, s_j). In what follows, we restrict attention to mechanisms with this property, and use a single function t to describe monetary transfers; notation for a mechanism simplifies to (x_b, x_s, t) with the understanding that $t_b(b_i, s_j) = t_s(b_i, s_j) = t(b_i, s_j)$.

The next result shows that mechanisms that withhold the good are not optimal.

Proposition 2. *It is never optimal to withhold the good with positive probability.*

The following lemmata collect standard insights from mechanism design about binding "local" incentive constraints and monotonicity properties of the allocation needed for the proof.

Lemma 1. *A mechanism (x_b, x_s, t) is implementable if and only if it satisfies IR_{b_1}, IR_{s_n}, $IC_{b_i \to b_{i+1}}$, $IC_{b_{i+1} \to b_i}$, $IC_{s_j \to s_{j+1}}$ and $IC_{s_{j+1} \to s_j}$ for $i = \overline{1, m-1}$ and $j = \overline{1, n-1}$. If (x_b, x_s, t) is implementable, then $\bar{x}_b(b_{i+1}) \geq \bar{x}_b(b_i)$ and $\bar{x}_s(s_{j+1}) \geq \bar{x}_s(s_j)$ for $i = \overline{1, m-1}$ and $j = \overline{1, n-1}$. If $\bar{x}_b(b_{i+1}) > \bar{x}_b(b_i)$ and $IC_{b_{i+1} \to b_i}$ holds with equality, then $IC_{b_i \to b_{i+1}}$ is satisfied with strict inequality. Similarly, if $\bar{x}_s(s_{j+1}) > \bar{x}_s(s_j)$ and $IC_{s_j \to s_{j+1}}$ holds with equality, then $IC_{s_{j+1} \to s_j}$ is satisfied with strict inequality.*

Lemma 2. *For every implementable mechanism (x_b, x_s, t), there exists a transfer specification t' such that (x_b, x_s, t') is implementable, and IR_{b_1}, $IC_{b_i \to b_{i-1}}$ and $IC_{s_j \to s_{j+1}}$ hold with equality under (x_b, x_s, t') for $i = \overline{2, m}$ and $j = \overline{1, n-1}$.*

To sketch the proof of Proposition 2, consider an implementable mechanism (x_b, x_s, t'') that does not fully allocate the good for some profile of reports (b_i, s_j). Lemma 2 implies the existence of a transfer function t that implements the same allocation (x_b, x_s) such that (x_b, x_s, t) satisfies the constraints listed in the lemma with equality. To increase the welfare generated by the mechanism, we contemplate a perturbation of the allocation whereby the seller of type s_j keeps the good with higher probability: $x'_s(b_i, s_j) = x_s(b_i, s_j) + \varepsilon$. If $\bar{x}_s(s_j) < \bar{x}_s(s_{j+1})$ and ε is sufficiently small, this perturbation preserves the monotonicity of $\bar{x}_s$, which per Lemma 1 is a necessary condition for the implementation of the resulting allocation. To provide incentives supporting the new allocation x', we modify transfers to type s_j uniformly so that s_j receives the same utility in the perturbed mechanism: $t'(b_k, s_j) = t(b_k, s_j) - p_b(b_i)\varepsilon s_j$ for all k.

The main constraint from Lemma 1 we need to check in order to prove that (x_b, x'_s, t') is an implementable mechanism is $IC_{s_{j+1} \to s_j}$. Indeed, the perturbation reduces transfers for a type s_j by $p_b(b_i)\varepsilon s_j$ to account for an increase of $p_b(b_i)\varepsilon$ in the probability of this type being allocated the good. This marginal change is also attractive to type s_{j+1}, for whom the marginal increase in the probability of being allocated the good is worth $p_b(b_i)\varepsilon s_{j+1}$, which is greater than the implicit cost of $p_b(b_i)\varepsilon s_j$. However, the condition $\bar{x}_s(s_j) < \bar{x}_s(s_{j+1})$ and the assumption that (x_b, x_s, t) satisfies $IC_{s_j \to s_{j+1}}$ with equality, along with Lemma 1, imply that (x_b, x_s, t) satisfies $IC_{s_{j+1} \to s_j}$ with strict inequality. By continuity, (x_b, x'_s, t') must also satisfy $IC_{s_{j+1} \to s_j}$ for small ε.

The case $\bar{x}_s(s_j) = \bar{x}_s(s_{j+1})$ requires a more substantial perturbation since the monotonicity condition necessary for implementability from Lemma 1 cannot be maintained if we increase $\bar{x}_s(s_j)$ without altering $\bar{x}_s(s_{j+1})$. However, in seeking a welfare improvement over the mechanism (x_b, x_s, t), we still need to take advantage of the slack in the constraint FT_{b_i,s_j} by increasing the probability that type s_j keeps the good to $x'_s(b_i, s_j) = x_s(b_i, s_j) + \varepsilon$ for some small ε. For this change to have a neutral effect on $\bar{x}_s(s_j)$, we decrease $x_s(b_{i'}, s_j)$ and increase $x_b(b_{i'}, s_j)$ by $\varepsilon p_b(b_i)/p_b(b_{i'})$ for some other buyer type $b_{i'}$. We can then neutralize the effect of this perturbation on type $b_{i'}$ by decreasing the probability that $b_{i'}$ receives the good from some other seller type $s_{j'}$. Finally, the increase in the probability that $s_{j'}$ keeps the good can be implemented without a further change in the allocation of $s_{j'}$ using the idea from the first step of the proof if j' can be selected such that $\bar{x}_s(s_{j'}) < \bar{x}_s(s_{j'+1})$.

Formally, we perturb the allocation as follows:

$$x'_s(b_i, s_j) = x_s(b_i, s_j) + \varepsilon \leq 1 - x_b(b_i, s_j)$$

$$x'_s(b_{i'}, s_j) = x_s(b_{i'}, s_j) - \frac{p_b(b_i)}{p_b(b_{i'})}\varepsilon > 0 \ \& \ x'_b(b_{i'}, s_j) = x_b(b_{i'}, s_j) + \frac{p_b(b_i)}{p_b(b_{i'})}\varepsilon < 1$$

$$x'_b(b_{i'}, s_{j'}) = x_b(b_{i'}, s_{j'}) - \frac{p_b(b_i)p_s(s_j)}{p_b(b_{i'})p_s(s_{j'})}\varepsilon > 0 \ \& \ x'_s(b_{i'}, s_{j'}) = x_s(b_{i'}, s_{j'}) + \frac{p_b(b_i)p_s(s_j)}{p_b(b_{i'})p_s(s_{j'})}\varepsilon < 1.$$

In the proof, we choose $s_{j'}$ to be the highest seller type such that $\bar{x}_s(s_{j'}) = \bar{x}_s(s_j)$, and identify a buyer type $b_{i'}$ such that these perturbations satisfy the right-hand side inequalities and generate allocation probabilities in $[0, 1]$. We then adjust the transfers such that $t'(b_k, s_{j'}) = t(b_k, s_{j'}) - s_{j'}\varepsilon p_b(b_i)p_s(s_j)/p_s(s_{j'})$ for $k = \overline{1, m}$ as in the first step of the proof in order to implement the same expected payoff for $s_{j'}$ under (x'_b, x'_s, t') and preserve the structure of incentives for the seller. Using Lemma 1, it can be checked that the perturbed mechanism (x'_b, x'_s, t') is implementable. The values of $(\bar{x}_b, \bar{x}_s)$ and $(\bar{x}'_b, \bar{x}'_s)$ differ only for type $s_{j'}$, who keeps the good with $\varepsilon p_b(b_i)p_s(s_j)/p_s(s_{j'})$ greater probability under x'_s than under x_s. Hence, the mechanism (x'_b, x'_s, t') boosts the welfare of (x_b, x_s, t) by $p_b(b_i)p_s(s_j)\varepsilon s_{j'}$, proving that (x_b, x_s, t) is not an optimal mechanism.

5. Damaging the Good for the Buyer

In this section and the next, we examine whether damaging the good for the buyer or for the seller can enhance welfare. Specifically, we assume that the mechanism designer can reduce the value of either the buyer or the seller for the good by shifting the corresponding value distribution down in the sense of first-order stochastic dominance. The shift does not have to be uniform across player types. If the good has multiple components (or is offered as a bundle) and there is heterogeneity in how player types value each component, then the mechanism designer may "damage" the good by prohibiting the sale or regulating the use of certain components with varying effects on different types. Uniform value reduction can also be accomplished by imposing a property tax.

Since withholding money or the good cannot improve welfare, we revert to the Myerson-Satterthwaite formulation of the optimal mechanism in which the constraints FT_{b_i,s_j} and FX_{b_i,s_j} are imposed with equality. Following the notation from the previous section, the *transfer* function $t : V_b \times V_s \to \mathbb{R}$ specifies a payment from the buyer to the seller. Notation for allocations is *updated* to reflect the fact that withholding the good is ruled out: the *allocation* is now determined by a single function $x : V_b \times V_s \to [0, 1]$ that describes the probability of trade between the buyer and the seller. In the *mechanism* (x, t), when the buyer reports type b_i and the seller reports type s_j, the two players trade the good with probability $x(b_i, s_j)$ in exchange for a monetary transfer $t(b_i, s_j)$.

Let

$$\bar{x}_b(b_i) = \sum_{j=1}^{n} p_s(s_j)x(b_i, s_j) \text{ and } \bar{t}_b(b_i) = \sum_{j=1}^{n} p_s(s_j)t(b_i, s_j)$$

denote the probability with which the buyer of type b_i *receives* the good and the expected monetary transfer type b_i *sends* to the seller, respectively, and let

$$\bar{x}_s(s_j) = \sum_{i=1}^{m} p_b(b_i)x(b_i, s_j) \text{ and } \bar{t}_s(s_j) = \sum_{i=1}^{m} p_b(b_i)t(b_i, s_j)$$

denote the probability with which the seller of type s_j *sells* the good and the expected monetary transfer type s_j *receives* from the buyer, respectively (note the *change in notation*: unlike in the previous sections, $\bar{x}_s$ here denotes trading probabilities for the seller). The optimal implementable mechanism (x,t) then solves the following linear program:

$$\begin{aligned}
\max \quad & \sum_{i=1}^{m} p_b(b_i)\bar{x}_b(b_i)b_i + \sum_{j=1}^{n} p_s(s_j)(1-\bar{x}_s(s_j))s_j \\
\text{s.t.} \quad & IR_{b_i}: \bar{x}_b(b_i)b_i - \bar{t}_b(b_i) \geq 0, \forall i = \overline{1,m} \\
& IR_{s_j}: \bar{t}_s(s_j) - \bar{x}_s(s_j)s_j \geq 0, \forall j = \overline{1,n} \\
& IC_{b_i \to b_k}: \bar{x}_b(b_i)b_i - \bar{t}_b(b_i) \geq \bar{x}_b(b_k)b_i - \bar{t}_b(b_k), \forall i,k = \overline{1,m} \\
& IC_{s_j \to s_k}: \bar{t}_s(s_j) - \bar{x}_s(s_j)s_j \geq \bar{t}_s(s_k) - \bar{x}_s(s_k)s_j, \forall j,k = \overline{1,n}.
\end{aligned}$$

We establish that damaging the good for the buyer can never improve welfare.

Proposition 3. *If p_b and p'_b are two value distributions for the buyer such that p_b first-order stochastically dominates p'_b, then for any seller value distribution p_s, no implementable mechanism for the pair of value distributions (p'_b, p_s) achieves greater total welfare than the optimal implementable mechanism for (p_b, p_s).*

The intuition for this result is that reducing buyer value entails less scope for allocating the good to the buyer and thus decreases potential gains from trade. For the proof, it is helpful to approach this intuition from the opposite angle: if the buyer value distribution shifts up, it becomes easier to provide incentives for trade with buyer types whose values increase. It is sufficient to prove the result for cases in which $p'_b(b_i) = p_b(b_i) + \varepsilon$ and $p'_b(b_{i+1}) = p_b(b_{i+1}) - \varepsilon$ with $\varepsilon > 0$ for some i, and $p'_b(b_k) = p_b(b_k)$ for $k \neq i, i+1$. We show that any mechanism (x', t') that is implementable when players' values are distributed according to (p'_b, p_s) and satisfies the constraint $IC_{b_{i+1} \to b_i}$ from Lemma 2 with equality can be perturbed to construct an implementable mechanism (x,t) for the value distributions (p_b, p_s) that differs from (x', t') only at the following profiles of types:

$$\begin{aligned}
x(b_{i+1}, s_j) &= \frac{\varepsilon}{p_b(b_{i+1})} x'(b_i, s_j) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} x'(b_{i+1}, s_j) \\
t(b_{i+1}, s_j) &= \frac{\varepsilon}{p_b(b_{i+1})} t'(b_i, s_j) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} t'(b_{i+1}, s_j), \forall j = \overline{1,n}.
\end{aligned}$$

The perturbed mechanism (x,t) channels $\varepsilon \bar{x}'_b(b_i)$ volume of trade from the ε-measure of buyer types whose valuations switch from b_i to b_{i+1} in the shift from p'_b to p_b without affecting the expected pattern of trade for other types, and consequently generates $\varepsilon \bar{x}'_b(b_i)(b_{i+1} - b_i)$ more surplus than (x', t').

6. Damaging the Good for the Seller

We finally consider situations in which the mechanism designer can damage the good for the seller. It is reasonable to assume that the seller has the option to keep the undamaged good, so the mechanism designer is constrained to provide each seller type at least the utility the seller would derive from consuming the good prior to the damage. To express this participation constraint for the seller, we focus on *type-by-type* damage whereby the value of each seller type j is reduced from s_j to s'_j. Thus, the mechanism designer can shift down the initial distribution of seller values p_s with support $s_1 < s_2 < \ldots < s_n$ to a distribution p'_s with support $s'_1 < s'_2 < \ldots < s'_n$ such that $s'_j \leq s_j$ and $p_s(s_j) = p'_s(s'_j)$ for $j = \overline{1,n}$. Then, in the notation of the previous section, the *damage participation constraint* for seller type j in an implementable mechanism (x, t) following the value reduction from p_s to p'_s requires that

$$\bar{t}_s(s'_j) + (1 - \bar{x}_s(s'_j))s'_j \geq s_j.$$

Note that the standard $IR_{s'_j}$ constraint for the pair of distributions (p_b, p'_s) is weaker, with the original value s_j being replaced by the reduced value s'_j on the right-hand side of the inequality above.

In contrast to damaging the good for the buyer, we discover that damaging the good for the seller may enhance welfare for many value distributions.

Proposition 4. *Let x be an implementable allocation for a pair of value distributions (p_b, p_s). Suppose that there exist $i \leq m - 1$ and $j \leq n - 1$ such that $\bar{x}_b(b_1) = \bar{x}_b(b_i) < \bar{x}_b(b_{i+1})$ and $\bar{x}_s(s_1) = \bar{x}_s(s_j) > \bar{x}_s(s_{j+1}) > 0$. If $b_i > s_1$ and $x(b_i, s_1) < 1$, then for any sufficiently small $\varepsilon > 0$, damaging the good to reduce the valuations of seller types $1, \ldots, j$ by ε (without affecting other types), enables the mechanism designer to implement an allocation x' that obeys the seller's damage participation constraint and coincides with x for every pair of corresponding buyer and seller types except that*

$$x'(b_i, s_1 - \varepsilon) = x(b_i, s_1) + \varepsilon \frac{(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1})) \sum_{l=1}^{j} p_s(s_l)}{p_s(s_1)(b_{i+1} - b_i) \sum_{k=i+1}^{m} p_b(b_k)}.$$

It is possible that the allocation x' generates greater total welfare given the reduced seller values than x for the original value distributions even in cases in which x is the optimal allocation for (p_b, p_s).[2]

[2]We can replace the hypothesis $x(b_i, s_1) < 1$ with the weaker condition $\bar{x}_s(s_1) < 1$, and adjust the conclusion to state that $\bar{x}'$ coincides with $\bar{x}$ for all corresponding buyer and seller types except that $\bar{x}'_b(b_i) = \bar{x}_b(b_i) + \varepsilon' p_s(s_1)$ and $\bar{x}'_s(s_1 - \varepsilon) = \bar{x}_s(s_1) + \varepsilon' p_b(b_i)$ for some $\varepsilon' > 0$. See the Appendix for a proof.

To understand this result, note that any mechanism that implements the allocation x does not distinguish between seller types 1 through j, and damaging the good for these types decreases their incentives to report higher values. Indeed, higher seller types keep the good with higher probability, but keeping the good is less attractive for a low-type seller whose value has been reduced. The slack created in the $IC_{s_{j'} \to s_{j+1}}$ constraints by the ε reduction in valuations for seller types $j' \leq j$ can be used to decrease transfers from the buyer of type $i+1$ and higher to the seller of type j and lower. This perturbation introduces slack in the $IC_{b_{i+1} \to b_i}$ constraint, which can be leveraged to increase the probability of an ex post efficient trade between the seller with reduced valuation $s_1 - \varepsilon$ and the buyer with valuation b_i at marginal terms that would be attractive to buyer type b_{i+1} if not for the decrease in transfers granted to this type under the perturbation.

However, the improvement in allocative efficiency generated by increased trading between types b_i and $s_1 - \varepsilon$ is achieved at the cost of damaging the good and thus reducing the utility of affected seller types in the event of no trade. Hence, there is a trade-off between the additional gains from trade for the pair $(b_i, s_1 - \varepsilon)$,

$$\begin{aligned} (1) \quad & p_b(b_i)p_s(s_1)(b_i - s_1 + \varepsilon)(x'(b_i, s_1 - \varepsilon) - x(b_i, s_1)) \\ & \qquad = \varepsilon \frac{(b_i - s_1 + \varepsilon)(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))p_b(b_i)\sum_{l=1}^{j} p_s(s_l)}{(b_{i+1} - b_i)\sum_{k=i+1}^{m} p_b(b_k)}, \end{aligned}$$

and the expected utility loss suffered by seller types 1 through j in the event that they do not trade the damaged good,

$$(2) \qquad \varepsilon(1 - \bar{x}_s(s_1)) \sum_{l=1}^{j} p_s(s_l).$$

Both the marginal efficiency gain deriving from the perturbed allocation and the marginal loss resulting from the damage are of order ε, and the net effect of the perturbation on total welfare depends on the original allocation x and the value distributions (p_b, p_s).[3]

[3] If the mechanism designer has the option to damage the good in a way that reduces the valuations of seller types $1, \ldots, j$ by ε without affecting other types, and can condition the exercise of this option on the seller's reported type, then the improvement in allocative efficiency can be achieved without actually damaging the good when the seller reports having one of the types $1, \ldots, j$. Indeed, the perturbed allocation x' can be implemented by damaging the good only if the seller reports having one of the types $j+1, \ldots, n$. Since seller types $j+1, \ldots, n$ are not affected by the damage, this form of report-contingent damage enables the mechanism designer to improve allocative efficiency without hurting the seller in the event of no trade, resulting in an unambiguous welfare gain. Note that report-contingent damage relaxes the $IC_{s_{j'} \to s_{j+1}}$ constraints for $j' \leq j$ more than ex ante damage, making it possible to also attain greater allocative efficiency. In the example regarding the sale of a private airplane from the introduction, report-contingent damage would entail removing the navigational device only when the seller claims to have high value for the airplane.

We illustrate this trade-off in an example which shows that either effect (1) or (2) can dominate when x is the optimal implementable allocation for the original value distributions. Suppose that there are two buyer and two seller types ($m = n = 2$). The buyer's valuations are $b_1 = 1$ and $b_2 = 4$ with equal probability, while the seller's valuations are (a parameter) $s_1 \in [0, 1)$ and $s_2 = 3$ with equal probability. We assume that the designer can damage the good for the low-type seller alone, and consider the welfare consequences of reducing s_1. The following lemma provides a method for computing the optimal implementable mechanism in this family of examples.

Lemma 3. *If (x, t) is an optimal implementable mechanism and x is not an ex post efficient allocation, then (x, t) satisfies all the constraints IR_{b_1}, IR_{s_n}, $IC_{b_i \to b_{i-1}}$, and $IC_{s_j \to s_{j+1}}$ with equality for $i = \overline{2, m}$ and $j = \overline{1, n-1}$.*

In the Appendix, we solve the relaxed linear program that determines the optimal implementable mechanism for the assumed value distributions in which the constraints are replaced by equalities derived from the binding constraints IR_{b_1}, IR_{s_2}, $IC_{b_2 \to b_1}$, and $IC_{s_1 \to s_2}$ identified in Lemma 3, and then show that the solution indeed represents the optimal mechanism. We find that the optimal mechanism implements the allocation

$$x(b_1, s_1) = 2/(2 + s_1), x(b_1, s_2) = 0, x(b_2, s_1) = 1, x(b_2, s_2) = 1. \tag{3}$$

The optimal mechanism allocates the good efficiently with higher probability as s_1 decreases to 0, reaching ex post efficiency for $s_1 = 0$.[4] As implied by Proposition 4, making the good less valuable for the low-type seller s_1 reduces his incentive to mimic the high type s_2, who trades less in the optimal mechanism. The resulting slack in the $IC_{s_1 \to s_2}$ constraint can be used to induce type s_1 to trade more with type b_1.

The optimal implementable allocation x identified in (3) generates a total welfare of

$$\frac{24 + 11s_1 + s_1^2}{4(2 + s_1)},$$

which is decreasing for $s_1 \in [0, \sqrt{6} - 2]$ and increasing for $s_1 \in [\sqrt{6} - 2, 1)$ as seen in Figure 1. At $s_1 = \sqrt{6} - 2 \approx 0.449$, the marginal improvement in gains from trade quantified by formula (1) resulting from a small decrease in s_1 exactly offsets the marginal utility loss described by expression (2) that seller type s_1 suffers in the event he retains the damaged

[4]Myerson and Satterthwaite (1983) note that their result about the impossibility of implementing the ex post efficient allocation for continuous value distributions does not extend to the case of discrete distributions. Matsuo (1989) and Kos and Manea (2009) characterize the pairs of buyer and seller discrete value distributions for which the ex post efficient allocation is implementable, as is the case in this example for $s_1 = 0$.

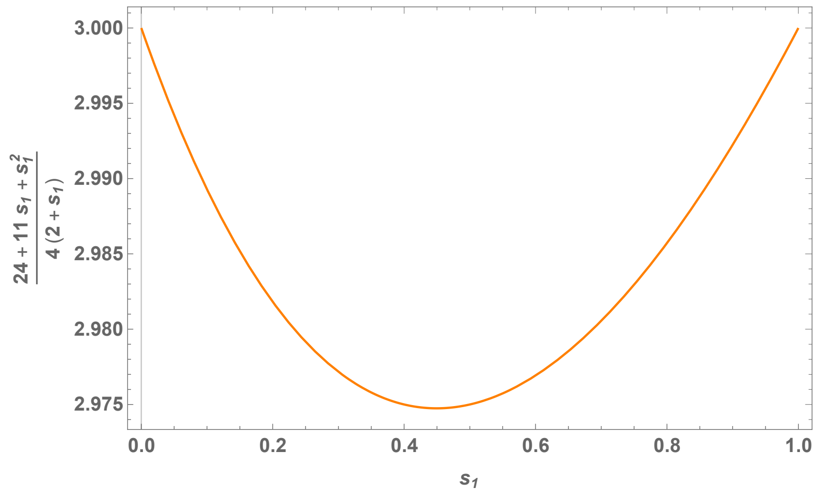

FIGURE 1. Total welfare in the optimal mechanism as a function of s_1

good. For $s_1 < \sqrt{6}-2$, the gain is greater than the loss, while for $s_1 > \sqrt{6}-2$ the opposite is true. For a concrete computation, x achieves total welfare 2.975 for $s_1 = 0.4$, and 3 for $s_1 = 0$. Therefore, reducing the value of the low-type seller from 0.4 to 0 enhances welfare in the optimal mechanism.

The seller's damage participation constraint is satisfied in this example when the mechanism designer damages the good for the low-type seller and implements the optimal allocation with the reduced valuation. Indeed, a seller with any initial valuation $s_1 \in (0, 1)$ would agree to have the good damaged and then participate in the optimal mechanism with a reduced valuation $s'_1 \in [0, s_1)$ because trading in the optimal mechanism with value s'_1 yields an expected payoff of at least $3/2$, which exceeds the seller's value s_1 for the undamaged good. To see this, note that after reducing the value of the low-type seller to s'_1, the individual rationality constraint IR_{s_2} and the incentive compatibility constraint $IC_{s'_1 \to s_2}$ hold with equality under the optimal mechanism. Hence, seller type s_2 trades the good with probability $1/2$ and receives an expected payment of $1/2 \times s_2 = 3/2$, and the low-type seller enjoys a utility of $1/2 \times s'_1 + 3/2 \geq 3/2$.

Proposition 4 presumes that the mechanism designer is able to target the damage of the good at seller types with low valuations. Suppose, however, that the mechanism designer is limited to damage that affects a high-value seller at least as much as a low-value seller. Specifically, damage is restricted to shifting down the seller value distribution p_s with support $s_1 < s_2 < \ldots < s_n$ to distributions p'_s with support $s'_1 < s'_2 < \ldots < s'_n$ such that $0 \leq s_1 - s'_1 \leq s_2 - s'_2 \leq \ldots \leq s_n - s'_n$ and $p'_s(s'_j) = p_s(s_j)$ for $j = \overline{1, n}$. A distribution p'_s with this property is said to be a *monotonic reduction* of p_s.

We show that damage via a monotonic reduction of seller values cannot raise expected welfare.

Proposition 5. *If p'_s is a monotonic reduction of the seller value distribution p_s, then for any buyer value distribution p_b, there is no implementable mechanism for the pair of value distributions (p_b, p'_s) that obeys the seller's damage participation constraint for the value reduction from p_s to p'_s and achieves greater total welfare than the optimal mechanism for (p_b, p_s).*[5]

To develop intuition for this result, it is helpful to revisit the binding constraints for the seller from Lemma 3 and the example discussed after Proposition 4. As suggested by the example, reducing the valuation of a seller type that is not the highest introduces slack in the binding incentive constraint for that type if the same allocation is to be implemented; this decreases the information rent needed to compensate that type relative to the next higher type. However, this potentially positive welfare effect is counterbalanced by the direct loss created by the damage as well as by the increased transfers required to meet the seller's damage participation constraint. In particular, the damage participation constraint for the highest-type seller is more restrictive than the post-damage individual rationality constraint for this type; a higher expected transfer is required for this type to implement the same allocation. We show that when damage entails a monotonic reduction in values, the additional transfer imposed by the highest-type seller's damage participation constraint is greater than the sum of transfers saved on lower types as a result of the diminished information rents dispensed to them after the damage.

For a proof sketch, fix a buyer value distribution p_b, and consider a seller value distribution p'_s with support $s'_1 < s'_2 < \ldots < s'_n$ that is a monotonic reduction of the distribution p_s with support $s_1 < s_2 < \ldots < s_n$. Suppose that (x', t') is an implementable mechanism for the pair of distributions (p_b, p'_s) that satisfies the seller's damage participation constraint for the value reduction from p_s to p'_s. We prove that the transfers prescribed by (x', t') can be perturbed to implement the corresponding type-by-type allocation x with $x(b_i, s_j) = x'(b_i, s'_j)$ for all i and j under the value distributions (p_b, p_s). The perturbed transfer function t for the new

[5]If the mechanism designer could damage the good even when the seller refuses to trade in the mechanism—i.e., has to abide by the individual rationality constraint with respect to the seller's reduced valuation, but not by the more restrictive damage participation constraint—then a modification of the example demonstrating Proposition 4 shows that welfare improvement can be extended to the case of a monotonic reduction. This is possible even when the mechanism designer is subject to an ex ante damage participation constraint requiring that the seller's expected payoff in the post-damage mechanism is not smaller than the seller's expected value for the undamaged good. See the Appendix for a discussion.

mechanism (x,t) is constructed as follows:

$$t(b_i,s_j)=t'(b_i,s'_j)-(1-\bar{x}_s(s_n))(s_n-s'_n)+\sum_{l=j}^{n-1}(\bar{x}_s(s_l)-\bar{x}_s(s_{l+1}))(s_l-s'_l),\forall i=\overline{1,m},j=\overline{1,n}.$$

In this specification, the amount $(1-\bar{x}_s(s_n))(s_n-s'_n)$ represents the designer's cost for reducing the value of the highest type of seller from s_n to s'_n in the event this type keeps the good, which has probability $1-\bar{x}_s(s_n)$. The term $(\bar{x}_s(s_l)-\bar{x}_s(s_{l+1}))(s_l-s'_l)$ captures the designer's savings on the information rent of seller type l relative to type $l+1$ when the value of type l drops from s_l to s'_l, which is weighted by the difference in trading probabilities $\bar{x}_s(s_l)-\bar{x}_s(s_{l+1})$ for the two types.

Individual rationality and incentive compatibility constraints for the seller under (x',t') with the value distributions (p_b,p'_s) and the monotonicity of the value reduction imply the corresponding constraints under (x,t) with the distributions (p_b,p_s). Since changes in transfers are independent of buyer type, (x,t) satisfies the buyer's incentive compatibility constraints for the distributions (p_b,p_s). The monotonicity assumption $0\le s_1-s'_1\le s_2-s'_2\le\ldots\le s_n-s'_n$ guarantees that

$$\sum_{l=1}^{n-1}(\bar{x}_s(s_l)-\bar{x}_s(s_{l+1}))(s_l-s'_l)\le(1-\bar{x}_s(s_n))(s_n-s'_n),$$

which implies that $t(b_i,s_j)\le t'(b_i,s'_j)$ for all i and j. Then, (x,t) satisfies the buyer's individual rationality constraints because (x',t') does.

Appendix: Proofs

Proof of Proposition 1. Let (x_b,x_s,t_b,t_s) be an optimal implementable mechanism. We prove by contradiction that $t_b(b_i,s_j)=t_s(b_i,s_j)$ for all pairs of types (b_i,s_j). Suppose that there exists a pair (b_i,s_j) with $\varepsilon:=t_b(b_i,s_j)-t_s(b_i,s_j)>0$. Consider a perturbation (x_b,x_s,t'_b,t'_s) of the mechanism (x_b,x_s,t_b,t_s) specifying that the transfer to seller type s_j is increased to match the transfer from b_i when the buyer reports type b_i and that the transfers s_j receives from all buyer types are uniformly reduced and credited to the buyer, so that s_j has the same expected utility following the perturbation. The perturbed mechanism (x_b,x_s,t'_b,t'_s) is identical to (x_b,x_s,t_b,t_s) except for the following profiles of reports:

$$\begin{aligned}t'_s(b_i,s_j)&=t_s(b_i,s_j)+\varepsilon-p_b(b_i)\varepsilon\\t'_s(b_k,s_j)&=t_s(b_k,s_j)-p_b(b_i)\varepsilon,\forall k=\overline{1,m},k\neq i\\t'_b(b_k,s_j)&=t_b(b_k,s_j)-p_b(b_i)\varepsilon,\forall k=\overline{1,m}.\end{aligned}$$

The definition is so that $(\bar{x}_b, \bar{x}_s, \bar{t}'_b, \bar{t}'_s)$ is identical to $(\bar{x}_b, \bar{x}_s, \bar{t}_b, \bar{t}_s)$ at all values with the following exception:

$$\bar{t}'_b(b_k) \quad = \quad \bar{t}_b(b_k) - p_b(b_i)p_s(s_j)\varepsilon, \forall k = \overline{1, m}.$$

We can easily compare expected outcomes under the mechanism (x_b, x_s, t'_b, t'_s) to those under (x_b, x_s, t_b, t_s). Each buyer type receives the good with the same probability and expects a uniform decrease in transfers. Each seller type obtains the good with the same probability and receives the same expected transfer. Therefore, (x_b, x_s, t'_b, t'_s) satisfies all buyer and seller rationality and incentive constraints because (x_b, x_s, t_b, t_s) does. We conclude that the mechanism (x_b, x_s, t'_b, t'_s) is implementable and yields an increase in welfare of $p_b(b_i)p_s(s_j)\varepsilon$ over the mechanism (x_b, x_s, t_b, t_s), a contradiction with the optimality of the latter. □

Proof of Lemma 1. We need to show that the set of selected constraints imply all the other ones. $IC_{b_i \to b_{i+1}}$ and $IC_{b_{i+1} \to b_i}$ lead to

$$(4) \qquad b_i(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i)) \leq \bar{t}_b(b_{i+1}) - \bar{t}_b(b_i) \leq b_{i+1}(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i)), \forall i = \overline{1, m-1}.$$

In particular, $b_i(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i)) \leq b_{i+1}(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i))$ and $b_{i+1} \geq b_i$ imply that

$$(5) \qquad \bar{x}_b(b_{i+1}) \geq \bar{x}_b(b_i), \forall i = \overline{1, m-1}.$$

If $IC_{b_{i+1} \to b_i}$ holds with equality, then the second inequality in (4) holds with equality. If additionally $\bar{x}_b(b_{i+1}) > \bar{x}_b(b_i)$, then the leftmost expression in (4) is strictly smaller than the rightmost expression in (4), so the first inequality must be strict, which means that $IC_{b_i \to b_{i+1}}$ is satisfied with strict inequality.

If $i < k \leq m$, then (4) and (5) imply that

$$b_i(\bar{x}_b(b_{g+1}) - \bar{x}_b(b_g)) \leq b_g(\bar{x}_b(b_{g+1}) - \bar{x}_b(b_g)) \leq \bar{t}_b(b_{g+1}) - \bar{t}_b(b_g)$$
$$\leq b_{g+1}(\bar{x}_b(b_{g+1}) - \bar{x}_b(b_g)) \leq b_k(\bar{x}_b(b_{g+1}) - \bar{x}_b(b_g)), \forall g = \overline{i, k-1}.$$

Adding up these inequalities, we obtain

$$b_i(\bar{x}_b(b_k) - \bar{x}_b(b_i)) \leq \bar{t}_b(b_k) - \bar{t}_b(b_i) \leq b_k(\bar{x}_b(b_k) - \bar{x}_b(b_i)),$$

which is equivalent to $IC_{b_i \to b_k}$ and $IC_{b_k \to b_i}$.

Also, $IC_{b_k \to b_1}$ and IR_{b_1} imply that

$$b_k\bar{x}_b(b_k) - \bar{t}_b(b_k) \geq b_k\bar{x}_b(b_1) - \bar{t}_b(b_1) \geq b_1\bar{x}_b(b_1) - \bar{t}_b(b_1) \geq 0,$$

verifying IR_{b_k} for $k = \overline{2, m}$.

The claims regarding the seller are checked similarly. □

Proof of Lemma 2. Among the monetary transfer functions t that implement (x_b, x_s), there exists one, t', that maximizes the expected payoff of seller type s_n. We show that t' has the desired properties. Suppose that $IC_{b_{i+1}\to b_i}$ holds with strict inequality under (x_b, x_s, t'). For $\varepsilon > 0$ sufficiently small, the transfer function t'' defined by

$$\begin{aligned} t''(b_k, s_j) &= t'(b_k, s_j) \text{ for } k \leq i, j = \overline{1, n} \\ t''(b_k, s_j) &= t'(b_k, s_j) + \varepsilon \text{ for } k \geq i+1, j = \overline{1, n} \end{aligned}$$

satisfies the sufficient individual rationality and incentive constraints from Lemma 1, and increases the expected payoff of s_n by $\varepsilon \sum_{k=i+1}^{n} p_b(b_k) > 0$, a contradiction with the definition of t'. The conclusion regarding IR_{b_1} follows from perturbing the transfers as in the formulae above with $i = 0$.

We obtain a similar contradiction if (x_b, x_s, t') satisfies $IC_{s_j\to s_{j+1}}$ with strict inequality by perturbing the transfer function as follows:

$$\begin{aligned} t''(b_i, s_l) &= t'(b_i, s_l) - \frac{\varepsilon}{\sum_{l'=1}^{j} p_s(s_{l'})} \text{ for } l \leq j, i = \overline{1, m} \\ t''(b_i, s_l) &= t'(b_k, s_j) + \frac{\varepsilon}{\sum_{l'=j+1}^{n} p_s(s_{l'})} \text{ for } l \geq j+1, i = \overline{1, m}. \end{aligned}$$

(This perturbation keeps buyer's incentives in place because the expected transfer of each buyer type under t'' is the same as under t.) □

Proof of Proposition 2. We proceed by contradiction. Suppose that (x_b, x_s, t) is an optimal implementable mechanism where the good is withheld with positive probability for some buyer-seller value profiles. We can assume without loss of generality that t is selected such that (x_b, x_s, t) satisfies all the constraints in Lemma 2 with equality. Let (b_i, s_j) be a pair of buyer and seller types where the good is withheld with positive probability, i.e., $x_b(b_i, s_j) + x_s(b_i, s_j) < 1$. Let j' be the highest seller type such that $\bar{x}_s(s_j) = \bar{x}_s(s_{j'})$. By Lemma 1, since $IC_{s_{j'}\to s_{j'+1}}$ holds with equality, and $\bar{x}_s(s_{j'}) < \bar{x}_s(s_{j'+1})$, it must be that $IC_{s_{j'+1}\to s_{j'}}$ holds with strict inequality.

Step 1. We show that $j \neq j'$ and $j \neq n$.

For a proof by contradiction, suppose that $j = j'$ or $j = n$. Consider the mechanism (x'_b, x'_s, t'), which perturbs (x_b, x_s, t) at the following values:

$$\begin{aligned} x'_s(b_i, s_j) &= x_s(b_i, s_j) + \varepsilon \\ t'(b_k, s_j) &= t(b_k, s_j) - s_j p_b(b_i)\varepsilon, \forall k = \overline{1, m}. \end{aligned}$$

For small $\varepsilon > 0$, $x_b'(b_i, s_j) + x_s'(b_i, s_j) \leq 1$. The definition is so that $(\bar{x}_b', \bar{x}_s', \bar{t}_b', \bar{t}_s')$ take the same values as $(\bar{x}_b, \bar{x}_s, \bar{t}_b, \bar{t}_s)$ with the following exceptions:

$$\begin{aligned}\bar{x}_s'(s_j) &= \bar{x}_s(s_j) + p_b(b_i)\varepsilon \\ \bar{t}_s'(s_j) &= \bar{t}_s(s_j) - s_j p_b(b_i)\varepsilon \\ \bar{t}_b'(b_k) &= \bar{t}_b(b_k) - s_j p_b(b_i) p_s(s_j)\varepsilon, \forall k = \overline{1, m}.\end{aligned}$$

We compare expected outcomes under the mechanism (x_b', x_s', t') to those under (x_b, x_s, t). For each type report, the buyer receives the good with the same probability, and the expected transfer is uniformly reduced by a constant. Hence, (x_b', x_s', t') satisfies all buyer rationality and incentive constraints because (x_b, x_s, t) does.

The mechanism (x_b', x_s', t') satisfies all seller constraints that do not involve type s_j because $(\bar{x}_s', \bar{t}_s')$ and $(\bar{x}_s, \bar{t}_s)$ coincide for those types. Moreover, seller type s_j expects the same utility from reporting any given type under (x_b', x_s', t') and (x_b, x_s, t), so (x_b', x_s', t') also satisfies the constraints involving the rationality and incentives of type s_j. We are left to check that (x_b', x_s', t') satisfies the constraints $IC_{s_{j-1} \to s_j}$ and $IC_{s_{j+1} \to s_j}$ from Lemma 1.

$IC_{s_{j-1} \to s_j}$ for (x_b', x_s', t') is equivalent to

$$\begin{aligned}&\bar{t}_s(s_{j-1}) - (1 - \bar{x}_s(s_{j-1}))s_{j-1} \geq (\bar{t}_s(s_j) - s_j p_b(b_i)\varepsilon) - (1 - (\bar{x}_s(s_j) + p_b(b_i)\varepsilon))s_{j-1} \\ \iff\ &\bar{t}_s(s_{j-1}) - (1 - \bar{x}_s(s_{j-1}))s_{j-1} \geq \bar{t}_s(s_j) - (1 - \bar{x}_s(s_j))s_{j-1} - (s_j - s_{j-1})p_b(b_i)\varepsilon,\end{aligned}$$

which follows from $s_j > s_{j-1}$ and the fact that (x_b, x_s, t) satisfies $IC_{s_{j-1} \to s_j}$.

For small $\varepsilon > 0$, (x_b', x_s', t') satisfies $IC_{s_{j+1} \to s_j}$ by continuity since (x_b', x_s', t') satisfies $IC_{s_{j+1} \to s_j}$ with strict inequality (for $j = j'$, this was argued immediately after the definition of j'; for $j = n$, the argument is unnecessary).

Therefore, for small $\varepsilon > 0$, the mechanism (x_b', x_s', t') is implementable and yields an increase in welfare of $s_j p_b(b_i) p_s(s_j)\varepsilon$ over the mechanism (x_b, x_s, t), a contradiction with the optimality of (x_b, x_s, t).

Step 2. It must be that $x_b(b_{i'}, s_{j'}) + x_s(b_{i'}, s_{j'}) = 1$ for all $i' = \overline{1, m}$.

This follows by replacing i with i' everywhere in the argument from Step 1.

Step 3. We have that $x_b(b_i, s_{j'}) = 0$ and $x_s(b_i, s_{j'}) = 1$.

Steps 1 and 2 show that $j < j'$ and $x_b(b_i, s_{j'}) + x_s(b_i, s_{j'}) = 1$. Suppose, for a contradiction, that $x_b(b_i, s_{j'}) > 0$.

Construct the mechanism (x'_b, x'_s, t') identical to (x_b, x_s, t) on $V_b \times V_s$ with the following exceptions:

$$\begin{aligned} x'_b(b_i, s_j) &= x_b(b_i, s_j) + \varepsilon \\ x'_b(b_i, s_{j'}) &= x_b(b_i, s_{j'}) - \frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon \\ x'_s(b_i, s_{j'}) &= x_s(b_i, s_{j'}) + \frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon \\ t'(b_k, s_{j'}) &= t(b_k, s_{j'}) - s_{j'} p_b(b_i)\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon, \forall k = \overline{1, m}. \end{aligned}$$

For small $\varepsilon > 0$, we have $x'_b(b_i, s_j) + x'_s(b_i, s_j) \leq 1$ and $x'_b(b_i, s_{j'}) \geq 0$.

Arguments similar to those used in Step 1 show that the mechanism (x'_b, x'_s, t') is implementable for small $\varepsilon > 0$. We reach the contradiction that (x'_b, x'_s, t') generates $s_{j'} p_s(s_{j'}) p_b(b_i) \frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon = s_{j'} p_b(b_i) p_s(s_j)\varepsilon$ more surplus than (x_b, x_s, t).

Step 4. We reach the final contradiction.

Steps 1 and 3 establish that $j < j'$ and $x_s(b_i, s_{j'}) = 1$. In particular, $x_s(b_i, s_j) \leq x_b(b_i, s_j) + x_s(b_i, s_j) < 1 = x_s(b_i, s_{j'})$. Since $x_s(b_i, s_j) < x_s(b_i, s_{j'})$, $p_b(b_i) > 0$, and $\bar{x}_s(s_j) = \bar{x}_s(s_{j'})$, there exist a buyer type $b_{i'}$ such that $x_s(b_{i'}, s_j) > x_s(b_{i'}, s_{j'})$. It follows that $x_s(b_{i'}, s_j) > 0$, $x_b(b_{i'}, s_j) < 1$, and $x_s(b_{i'}, s_{j'}) < 1$. By Step 2, $x_b(b_{i'}, s_{j'}) + x_s(b_{i'}, s_{j'}) = 1$, and hence $x_b(b_{i'}, s_{j'}) > 0$. We collect the relevant inequalities:

$$(6) \qquad x_s(b_i, s_j) < 1,\ x_s(b_{i'}, s_j) > 0,\ x_b(b_{i'}, s_j) < 1,\ x_b(b_{i'}, s_{j'}) > 0,\ x_s(b_{i'}, s_{j'}) < 1.$$

Construct a mechanism (x'_b, x'_s, t') that perturbs (x_b, x_s, t) for the following types:

$$\begin{aligned} &x'_s(b_i, s_j) = x_s(b_i, s_j) + \varepsilon \\ &x'_s(b_{i'}, s_j) = x_s(b_{i'}, s_j) - \frac{p_b(b_i)}{p_b(b_{i'})}\varepsilon \quad \& \quad x'_b(b_{i'}, s_j) = x_b(b_{i'}, s_j) + \frac{p_b(b_i)}{p_b(b_{i'})}\varepsilon \\ &x'_b(b_{i'}, s_{j'}) = x_b(b_{i'}, s_{j'}) - \frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon \quad \& \quad x'_s(b_{i'}, s_{j'}) = x_s(b_{i'}, s_{j'}) + \frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon \\ &t'(b_k, s_{j'}) = t(b_k, s_{j'}) - s_{j'} p_b(b_{i'})\frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon = t(b_k, s_{j'}) - s_{j'} p_b(b_i)\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon, \forall k = \overline{1, m}. \end{aligned}$$

For small $\varepsilon > 0$, the set of inequalities (6) guarantees that all values of x' belong to the interval of probabilities $[0, 1]$. We can then argue as in Steps 1 and 3 that for small $\varepsilon > 0$, the mechanism (x'_b, x'_s, t') is implementable and improves the welfare of (x_b, x_s, t) by

$$s_{j'} p_b(b_{i'}) p_s(s_{j'}) \frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_j)}{p_s(s_{j'})}\varepsilon = s_{j'} p_s(s_j) p_b(b_i)\varepsilon,$$

a contradiction.

Steps 1 through 4 show that if (x_b, x_s, t) is an optimal implementable mechanism, then $x_b(b_i, s_j) + x_s(b_i, s_j) = 1$. □

Proof of Proposition 3. We can assume that the optimal mechanisms for (p_b, p_s) and (p'_b, p_s) are defined for buyer types in the union of the supports of p_b and p'_b since any implementable mechanism can be extended to a set of probability-zero types by specifying allocations and transfers for each such type as an optimal selection for the type from the allocations and transfers prescribed by the mechanism for existing types along with the option of no trade and zero transfers. It is then sufficient to establish the result for cases in which $p'_b(b_i) = p_b(b_i) + \varepsilon$ and $p'_b(b_{i+1}) = p_b(b_{i+1}) - \varepsilon$ with $\varepsilon > 0$, and $p'_b(b_k) = p_b(b_k)$ for $k \neq i, i+1$.

Let (x', t') be a mechanism that maximizes welfare when players' values are distributed according to (p'_b, p_s) and satisfies all constraints in Lemma 2 with equality (we use only the equality in the constraint $IC_{b_{i+1} \to b_i}$ for this proof). Define the mechanism (x, t) to coincide with (x', t') for all type profiles with the following exceptions:

$$\begin{aligned} x(b_{i+1}, s_j) &= \frac{\varepsilon}{p_b(b_{i+1})} x'(b_i, s_j) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} x'(b_{i+1}, s_j), \forall j = \overline{1, n} \\ t(b_{i+1}, s_j) &= \frac{\varepsilon}{p_b(b_{i+1})} t'(b_i, s_j) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} t'(b_{i+1}, s_j), \forall j = \overline{1, n}. \end{aligned}$$

Denote by $\bar{x}_b, \bar{x}_s, \bar{t}_b, \bar{t}_s$ and $\bar{x}'_b, \bar{x}'_s, \bar{t}'_b, \bar{t}'_s$ the probabilities of trade and expected transfers for each buyer and seller type under the mechanism (x, t) when values are distributed according to (p_b, p_s) and under the mechanism (x', t') when values are distributed according to (p'_b, p_s), respectively. The allocation x is defined so that

$$p_b(b_i)x(b_i, s_j) + p_b(b_{i+1})x(b_{i+1}, s_j) = p'_b(b_i)x'(b_i, s_j) + p'_b(b_{i+1})x'(b_{i+1}, s_j),$$

which implies that $\bar{x}_s(s_j) = \bar{x}'_s(s_j)$ for all j. Similarly, $\bar{t}_s(s_j) = \bar{t}'_s(s_j)$ for all j. Moreover, $(\bar{x}'_b, \bar{x}'_s, \bar{t}'_b, \bar{t}'_s)$ is identical to $(\bar{x}_b, \bar{x}_s, \bar{t}_b, \bar{t}_s)$ at all values except b_{i+1}, for which we have

$$\begin{aligned} \bar{x}_b(b_{i+1}) &= \frac{\varepsilon}{p_b(b_{i+1})} \bar{x}'_b(b_i) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} \bar{x}'_b(b_{i+1}) \\ \bar{t}_b(b_{i+1}) &= \frac{\varepsilon}{p_b(b_{i+1})} \bar{t}'_b(b_i) + \frac{p_b(b_{i+1}) - \varepsilon}{p_b(b_{i+1})} \bar{t}'_b(b_{i+1}). \end{aligned}$$

We show that the mechanism (x, t) satisfies all the IR and IC constraints in Lemma 1 when players' values are distributed according to (p_b, p_s). Since $\bar{x}_b, \bar{t}_b, \bar{x}_s, \bar{t}_s$ coincide with $\bar{x}'_b, \bar{t}'_b, \bar{x}'_s, \bar{t}'_s$ at all values except for b_{i+1}, the only constraints we need to check are those that involve buyer type b_{i+1}. Neither type b_i nor type b_{i+2} have an incentive to mimic type b_{i+1} in the mechanism (x, t) since each of these types is assigned the same allocation probabilities and expected transfers under (x,t) and (x',t'), and prefers his outcome under (x',t') to any other outcome achievable under (x',t') as well as to the convex combination of outcomes for types b_i and b_{i+1} specified by $(x(b_{i+1},\cdot),t(b_{i+1},\cdot))$.

We are left to check that (x,t) satisfies $IC_{b_{i+1}\to b_i}$ and $IC_{b_{i+1}\to b_{i+2}}$. The constraint $IC_{b_{i+1}\to b_i}$ requires that

$$\frac{\varepsilon}{p_b(b_{i+1})}(\bar{x}'_b(b_i)b_{i+1}-\bar{t}'_b(b_i))+\frac{p_b(b_{i+1})-\varepsilon}{p_b(b_{i+1})}(\bar{x}'_b(b_{i+1})b_{i+1}-\bar{t}'_b(b_{i+1}))\geq \bar{x}'_b(b_i)b_{i+1}-\bar{t}'_b(b_i),$$

which is is equivalent to

$$x'_b(b_{i+1})b_{i+1}-\bar{t}'_b(b_{i+1})\geq \bar{x}'_b(b_i)b_{i+1}-\bar{t}'_b(b_i).$$

The latter inequality is a consequence of $IC_{b_{i+1}\to b_i}$ under the mechanism (x',t').

To check that (x,t) satisfies $IC_{b_{i+1}\to b_{i+2}}$, we need to show that

$$\frac{\varepsilon}{p_b(b_{i+1})}(\bar{x}'_b(b_i)b_{i+1}-\bar{t}'_b(b_i))+\frac{p_b(b_{i+1})-\varepsilon}{p_b(b_{i+1})}(\bar{x}'_b(b_{i+1})b_{i+1}-\bar{t}'_b(b_{i+1}))\geq \bar{x}'_b(b_{i+2})b_{i+1}-\bar{t}'_b(b_{i+2}).$$

As (x',t') is assumed to satisfy $IC_{b_{i+1}\to b_i}$ with equality, we have $\bar{x}'_b(b_i)b_{i+1}-\bar{t}'_b(b_i)=\bar{x}'_b(b_{i+1})b_{i+1}-\bar{t}'_b(b_{i+1})$, so the inequality above follows from $IC_{b_{i+1}\to b_{i+2}}$ under (x',t').

Since $p_b,p_s,\bar{x}_b,\bar{t}_b,\bar{x}_s,\bar{t}_s$ coincide with $p'_b,p_s,\bar{x}'_b,\bar{t}'_b,\bar{x}'_s,\bar{t}'_s$ at all values except for b_i and b_{i+1}, the difference in the welfare achieved by (x,t) with distributions (p_b,p_s) and (x',t') with distributions (p'_b,p_s) is given by

$$\begin{aligned}
&p_b(b_i)\bar{x}_b(b_i)b_i+p_b(b_{i+1})\bar{x}_b(b_{i+1})b_{i+1}-p'_b(b_i)\bar{x}'_b(b_i)b_i-p'_b(b_{i+1})\bar{x}'_b(b_{i+1})b_{i+1}\\
&=p_b(b_i)\bar{x}'_b(b_i)b_i+p_b(b_{i+1})\left(\frac{\varepsilon}{p_b(b_{i+1})}\bar{x}'_b(b_i)+\frac{p_b(b_{i+1})-\varepsilon}{p_b(b_{i+1})}\bar{x}'_b(b_{i+1})\right)b_{i+1}\\
&-(p_b(b_i)+\varepsilon)\bar{x}'_b(b_i)b_i-(p_b(b_{i+1})-\varepsilon)\bar{x}'_b(b_{i+1})b_{i+1}\\
&=\varepsilon\bar{x}'_b(b_i)(b_{i+1}-b_i),
\end{aligned}$$

which is non-negative. Therefore, the optimal mechanism for (p_b,p_s) yields at least the same amount of welfare as the optimal mechanism for (p'_b,p_s). □

Proof of Proposition 4. Lemma 2 implies the existence of a transfer function t that implements the allocation x for the original pair of value distributions (p_b,p_s) such that (x,t) satisfies $IC_{b_{i+1}\to b_i}$ with equality. Fix $\varepsilon>0$ and let $s'_l=s_l-\varepsilon$ for $l=\overline{1,j}$. Define

$$\varepsilon'=\varepsilon\frac{(\bar{x}_s(s_j)-\bar{x}_s(s_{j+1}))\sum_{l'=1}^{j}p_s(s_{l'})}{p_s(s_1)(b_{i+1}-b_i)\sum_{k'=i+1}^{m}p_b(b_{k'})}. \tag{7}$$

We seek to implement an allocation x' when the values of seller types $l = \overline{1, j}$ are reduced from s_l to s_l' that differs from x (for corresponding seller types) only in that

$$x'(b_i, s_1') = x(b_i, s_1) + \varepsilon'.$$

For this purpose, we perturb the transfer function t for the following pairs of types:

$$\begin{aligned} t'(b_i, s_1') &= t(b_i, s_1) + \varepsilon' b_i \\ t'(b_k, s_1') &= t(b_k, s_1) - \varepsilon'(b_{i+1} - b_i)\frac{p_s(s_1)}{\sum_{l'=1}^{j} p_s(s_{l'})} - \varepsilon'(b_i - s_1')\frac{p_b(b_i)}{\sum_{k'=i+1}^{m} p_b(b_{k'})}, \forall k = \overline{i+1, m} \\ t'(b_k, s_l') &= t(b_k, s_l) - \varepsilon'(b_{i+1} - b_i)\frac{p_s(s_1)}{\sum_{l'=1}^{j} p_s(s_{l'})}, \forall k = \overline{i+1, m}, l = \overline{2, j}. \end{aligned}$$

Transfers are specified so that buyer type b_i pays exactly his gain from the additional probability of receiving the good from seller type s_1. Similarly, we have that $\bar{t}_s'(s_1') - \bar{t}_s'(s_2') = \bar{t}_s(s_1) - \bar{t}_s(s_2) + \varepsilon' p_b(b_i) s_1'$, so that the new seller type s_1' is compensated relative to type s_2' for the utility lost by trading with buyer type b_i with the extra ε' probability.

To show that (x', t') is implementable following the value reduction, we verify that (x', t') satisfies the sufficient individual rationality and incentive constraints from Lemma 1 as well as the stricter participation constraints required for damaging the good for the seller.

Buyer expectations $(\bar{x}_b', \bar{t}_b')$ under the new mechanism differ from the original ones in the following cases:

$$\begin{aligned} \bar{x}_b'(b_i) &= \bar{x}_b(b_i) + \varepsilon' p_s(s_1) \\ \bar{t}_b'(b_i) &= \bar{t}_b(b_i) + \varepsilon' p_s(s_1) b_i \\ \bar{t}_b'(b_k) &= \bar{t}_b(b_k) - \varepsilon' p_s(s_1)(b_{i+1} - b_i) - \varepsilon'(b_i - s_1')\frac{p_b(b_i) p_s(s_1)}{\sum_{k'=i+1}^{m} p_b(b_{k'})}, \forall k = \overline{i+1, m}. \end{aligned}$$

The mechanism (x', t') satisfies the incentive constraints $IC_{b_k \to b_{k+1}}$ and $IC_{b_{k+1} \to b_k}$ for $k = \overline{i+1, m-1}$ since the involved buyer types receive the same allocation under x and x' and their expected payments are reduced by the same amount when shifting from t to t'. It is clear that (x', t') satisfies the constraints $IC_{b_k \to b_{k+1}}$ and $IC_{b_{k+1} \to b_k}$ for $k = \overline{1, i-2}$ and IR_{b_1}. We next check that (x', t') satisfies the incentive constraints $IC_{b_{i-1} \to b_i}$, $IC_{b_i \to b_{i-1}}$, $IC_{b_i \to b_{i+1}}$ and $IC_{b_{i+1} \to b_i}$.

The mechanism (x', t') satisfies $IC_{b_i \to b_{i-1}}$ because buyer type b_i expects the same utility under (x, t) and (x', t') when he reports type b_i, and the same is true when he reports type b_{i-1}. $IC_{b_{i-1} \to b_i}$ holds under (x', t') since for buyer type b_{i-1}, the marginal loss from misreporting his type to be b_i when we shift from (x, t) to (x', t') entails receiving the good with additional probability $\varepsilon' p_s(s_1)$ for an expected benefit of $\varepsilon' p_s(s_1) b_{i-1}$ at the greater cost of $\varepsilon' p_s(s_1) b_i$.

The buyer's incentives we are left to check for (x', t') are $IC_{b_{i+1}\to b_i}$ and $IC_{b_i \to b_{i+1}}$. Since $IC_{b_{i+1}\to b_i}$ holds with equality under (x, t), we have that

$$\bar{x}_b(b_{i+1})b_{i+1} - \bar{t}_b(b_{i+1}) = \bar{x}_b(b_i)b_{i+1} - \bar{t}_b(b_i). \tag{8}$$

To see that (x', t') satisfies $IC_{b_{i+1}\to b_i}$, we use (8) to infer that

$$\bar{x}'_b(b_{i+1})b_{i+1} - \bar{t}'_b(b_{i+1}) = \bar{x}_b(b_{i+1})b_{i+1} - \bar{t}_b(b_{i+1}) + \varepsilon' p_s(s_1)(b_{i+1} - b_i) + \varepsilon'(b_i - s'_1)\frac{p_b(b_i)p_s(s_1)}{\sum_{k'=i+1}^m p_b(b_{k'})}$$

$$= \bar{x}_b(b_i)b_{i+1} - \bar{t}_b(b_i) + \varepsilon' p_s(s_1)(b_{i+1} - b_i) + \varepsilon'(b_i - s'_1)\frac{p_b(b_i)p_s(s_1)}{\sum_{k'=i+1}^m p_b(b_{k'})}$$

$$= \bar{x}'_b(b_i)b_{i+1} - \bar{t}'_b(b_i) + \varepsilon'(b_i - s'_1)\frac{p_b(b_i)p_s(s_1)}{\sum_{k'=i+1}^m p_b(b_{k'})} > \bar{x}'_b(b_i)b_{i+1} - \bar{t}'_b(b_i).$$

The inequality follows from the hypothesis that $b_i > s_1$, which implies that $b_i > s'_1$.

Also using (8), we obtain that

$$\bar{x}'_b(b_i)b_i - \bar{t}'_b(b_i) = \bar{x}_b(b_i)b_i - \bar{t}_b(b_i) = \bar{x}_b(b_{i+1})b_i - \bar{t}_b(b_{i+1}) + (b_{i+1} - b_i)(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i))$$

$$= \bar{x}'_b(b_{i+1})b_i - \bar{t}'_b(b_{i+1}) - \varepsilon' p_s(s_1)(b_{i+1} - b_i) - \varepsilon'(b_i - s'_1)\frac{p_b(b_i)p_s(s_1)}{\sum_{k'=i+1}^m p_b(b_{k'})} + (b_{i+1} - b_i)(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i)).$$

$IC_{b_i\to b_{i+1}}$ for (x', t') is thus equivalent to

$$\varepsilon' p_s(s_1)(b_{i+1} - b_i) + \varepsilon'(b_i - s'_1)\frac{p_b(b_i)p_s(s_1)}{\sum_{k'=i+1}^m p_b(b_{k'})} \leq (b_{i+1} - b_i)(\bar{x}_b(b_{i+1}) - \bar{x}_b(b_i)),$$

which holds for small $\varepsilon > 0$ when ε' is given by (7) since $\bar{x}_b(b_{i+1}) > \bar{x}_b(b_i)$ implies that the right-hand side of the inequality is positive.

We now turn to verifying the seller constraints for (x', t'). Seller expectations $(\bar{x}'_s, \bar{t}'_s)$ under the perturbed mechanism differ from the corresponding ones in the original mechanism in the following instances:

$$\begin{aligned}
\bar{x}'_s(s'_1) &= \bar{x}_s(s_1) + \varepsilon' p_b(b_i) \\
\bar{t}'_s(s'_1) &= \bar{t}_s(s_1) + \varepsilon' p_b(b_i)s'_1 - \varepsilon' p_s(s_1)(b_{i+1} - b_i)\frac{\sum_{k'=i+1}^m p_b(b_{k'})}{\sum_{l'=1}^j p_s(s_{l'})} \\
\bar{t}'_s(s'_l) &= \bar{t}_s(s_l) - \varepsilon' p_s(s_1)(b_{i+1} - b_i)\frac{\sum_{k'=i+1}^m p_b(b_{k'})}{\sum_{l'=1}^j p_s(s_{l'})}, \forall l = \overline{2, j}.
\end{aligned}$$

We provide a proof for $j \geq 2$ (the case $j = 1$ involves similar arguments). Since $\bar{x}_s(s_1) = \bar{x}_s(s_j)$, it must be that $\bar{x}_s(s_l) = \bar{x}_s(s_1)$ and $\bar{t}_s(s_l) = \bar{t}_s(s_1)$ for $l = \overline{2, j}$. Hence $\bar{x}'_s(s_l) = \bar{x}'_s(s_2)$

and $\bar{t}'_s(s_l) = \bar{t}'_s(s_2)$ for $l = \overline{2, j}$. This implies that (x', t') satisfies $IC_{s'_l \to s'_{l+1}}$ and $IC_{s'_{l+1} \to s'_l}$ for $l = \overline{2, j-1}$. As $(\bar{x}'_s, \bar{t}'_s)$ coincides with $(\bar{x}_s, \bar{t}_s)$ for seller types $l = \overline{j+1, n}$, and (x, t) is implementable, (x', t') satisfies $IC_{s_l \to s_{l+1}}$ and $IC_{s_{l+1} \to s_l}$ for $l = \overline{j+1, n-1}$ and IR_{s_n}. We are left to check the following incentive compatibility constraints from Lemma 1: $IC_{s'_1 \to s'_2}$, $IC_{s'_2 \to s'_1}$, $IC_{s'_j \to s_{j+1}}$ and $IC_{s_{j+1} \to s'_j}$.

The constraints $IC_{s'_1 \to s'_2}$ and $IC_{s'_2 \to s'_1}$ for (x', t') are equivalent to

$$(\bar{x}'_s(s'_1) - \bar{x}'_s(s'_2))s'_1 \le \bar{t}'_s(s'_1) - \bar{t}'_s(s'_2) \le (\bar{x}'_s(s'_1) - \bar{x}'_s(s'_2))s'_2.$$

These inequalities follow from $\bar{x}'_s(s'_1) - \bar{x}'_s(s'_2) = \varepsilon' p_b(b_i)$, $\bar{t}'_s(s'_1) - \bar{t}'_s(s'_2) = \varepsilon' p_b(b_i) s'_1$ and $s'_1 < s'_2$.

To check $IC_{s'_j \to s_{j+1}}$ for the mechanism (x', t'), note that $IC_{s_j \to s_{j+1}}$ for (x, t) implies that

$$\begin{aligned}
&\bar{t}'_s(s'_j) - \bar{t}'_s(s'_{j+1}) - (\bar{x}'_s(s'_j) - \bar{x}'_s(s'_{j+1}))s'_j \\
&= \bar{t}_s(s_j) - \bar{t}_s(s_{j+1}) - (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s_j + \varepsilon(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1})) - \varepsilon' p_s(s_1)(b_{i+1} - b_i)\frac{\sum_{k'=i+1}^m p_b(b_{k'})}{\sum_{l'=1}^j p_s(s_{l'})} \\
&\ge \varepsilon(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1})) - \varepsilon' p_s(s_1)(b_{i+1} - b_i)\frac{\sum_{k'=i+1}^m p_b(b_{k'})}{\sum_{l'=1}^j p_s(s_{l'})} = 0,
\end{aligned}$$

where the last equality follows from the definition of ε' in (7). The slack created in $IC_{s_j \to s_{j+1}}$ by the ε value reduction for seller type j is used here to decrease the transfers from buyer types $i+1$ through m to seller types 1 through j, which makes it possible to introduce the slack in $IC_{b_{i+1} \to b_i}$ necessary for increasing the probability of trade between seller type s_1 and buyer type b_i at terms that would be otherwise attractive to buyer type b_{i+1}.

The mechanism (x', t') satisfies $IC_{s_{j+1} \to s'_j}$ because (x, t) satisfies $IC_{s_{j+1} \to s_j}$ and we have that $\bar{x}'_s(s_{j+1}) = \bar{x}_s(s_{j+1})$, $\bar{t}'_s(s_{j+1}) = \bar{t}_s(s_{j+1})$, $\bar{x}'_s(s'_j) = \bar{x}_s(s_j)$ and $\bar{t}'_s(s'_j) < \bar{t}_s(s_j)$.

We finally check that seller types $l = \overline{1, j}$ have incentives to participate in the damage of the good. Since (x', t') satisfies $IC_{s'_l \to s_{j+1}}$ and $IR_{s_{j+1}}$ (Lemma 1) and $\bar{x}'_s(s_{j+1}) = \bar{x}_s(s_{j+1}) > 0$, we have that

$$\begin{aligned}
\bar{t}'_s(s'_l) - \bar{x}'_s(s'_l)s'_l \ge \bar{t}'_s(s_{j+1}) - \bar{x}'_s(s_{j+1})s'_l &= \bar{t}'_s(s_{j+1}) - \bar{x}'_s(s_{j+1})s_{j+1} + \bar{x}_s(s_{j+1})(s_{j+1} - s'_l) \\
&\ge \bar{x}_s(s_{j+1})(s_{j+1} - s'_l) > \bar{x}_s(s_{j+1})(s_{j+1} - s_j).
\end{aligned}$$

Then,

$$\bar{t}'_s(s'_l) + (1 - \bar{x}_s(s'_l))s'_l - s_l = \bar{t}'_s(s'_l) - \bar{x}_s(s'_l)s'_l + s'_l - s_l > \bar{x}_s(s_{j+1})(s_{j+1} - s_j) - \varepsilon,$$

which is positive if $\varepsilon < \bar{x}_s(s_{j+1})(s_{j+1}-s_j)$. This upper bound on ε is positive since $\bar{x}_s(s_{j+1}) > 0$. $\square$

Proof of footnote 2. Suppose that $x(b_i, s_1) = 1$ and $\bar{x}_s(s_1) < 1$. Then, there exists a buyer type i' such that $x(b_{i'}, s_1) < 1$. Since $\bar{x}_b(b_i) = \bar{x}_b(b_1) \leq \bar{x}_b(b_{i'})$ and $x(b_{i'}, s_1) < 1 = x(b_i, s_1)$, there must exist a seller type j' such that $x(b_{i'}, s_{j'}) > x(b_i, s_{j'})$. In particular, we have $x(b_{i'}, s_{j'}) > 0$ and $x(b_i, s_{j'}) < 1$. Construct a perturbation x^1 of the allocation x that differs from x only in the following cases:

$$\begin{aligned}
x^1(b_i, s_1) &= x(b_i, s_1) - \eta \\
x^1(b_i, s_{j'}) &= x(b_i, s_{j'}) + \frac{p_s(s_1)}{p_s(s_{j'})}\eta \\
x^1(b_{i'}, s_{j'}) &= x(b_{i'}, s_{j'}) - \frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_1)}{p_s(s_{j'})}\eta \\
x^1(b_{i'}, s_1) &= x(b_{i'}, s_1) + \frac{p_b(b_i)}{p_b(b_{i'})}\eta.
\end{aligned}$$

For sufficiently small $\eta > 0$, the inequalities above imply that all the values of x^1 are in $[0, 1]$. Moreover, x^1 generates the same allocation probability as x for each buyer and seller type, and has the property that $x^1(b_i, s_1) < 1$. The conclusion of footnote 2 then follows from applying Proposition 4 to allocation x^1. $\square$

Proof of Lemma 3. We first prove the claim for optimal implementable mechanisms (x, t) with the property that $\bar{x}_b$ is strictly increasing and $\bar{x}_s$ is strictly decreasing. Then, we deal with the more involved case in which these assumptions do not hold.

Case 1. Suppose that (x, t) is an optimal implementable mechanism that is not ex post efficient such that $\bar{x}_b$ is strictly increasing and $\bar{x}_s$ is strictly decreasing. Then, (x, t) must satisfy the constraints from the statement with equality.

Let T be the set of monetary transfer functions among those that implement x and maximize the expected utility of a type s_n seller. As argued in the proof of Lemma 2, every mechanism (x, t') with $t' \in T$ satisfies the constraints IR_{b_1}, $IC_{b_i \to b_{i-1}}$ and $IC_{s_j \to s_{j+1}}$ with equality. We will show that for any $t' \in T$, (x, t') also satisfies IR_{s_n} with equality. However, if IR_{s_n} holds with equality at all (x, t') with $t' \in T$, then the maximum utility of a type s_n seller is zero for all monetary transfer functions that implement x. Then (x, t) also achieves the maximum, hence $t \in T$ and (x, t) satisfies all the required constraints with equality as well.

Let $t' \in T$. The fact that all $IC_{b_i \to b_{i-1}}$ and $IC_{s_j \to s_{j+1}}$ hold with equality, the strict monotonicity of $\bar{x}_b$ and $\bar{x}_s$, and Lemma 1 imply that all $IC_{b_{i-1} \to b_i}$ and $IC_{s_{j+1} \to s_j}$ hold with strict inequality for (x, t'). Suppose that IR_{s_n} does not hold with equality for (x, t'). Consider the monetary transfer function t'' defined by

$$t''(b_i, s_j) = t'(b_i, s_j) - (i + j)\varepsilon.$$

By continuity, for $\varepsilon > 0$ sufficiently small, the mechanism (x, t'') still satisfies all IR_{s_n}, $IC_{b_{i-1} \to b_i}$ and $IC_{s_{j+1} \to s_j}$ strictly, but also satisfies all $IC_{b_i \to b_{i-1}}$, $IC_{s_j \to s_{j+1}}$, and IR_{b_1} strictly.

Hence, (x, t'') satisfies the sufficient constraints for implementability from Lemma 1 with strict inequality. Then, (x', t'') also satisfies the sufficient constraints for any small perturbation x' around x, so (x', t'') is implementable for all x' s.t. $\sum_{i,j} |x'(b_i, s_j) - x(b_i, s_j)|$ is sufficiently small. If x is not ex post efficient, there exists such an x' that generates more welfare, a contradiction with the optimality of x. This concludes the proof for this case as explained in the first paragraph.

Case 2. Suppose that (x, t) is an optimal mechanism that is not ex post efficient and does not satisfy all the constraints listed in the statement of the lemma with equality. We reach a contradiction.

Since x is not ex post efficient, either $\{(i, j)|b_i > s_j, x(b_i, s_j) < 1\}$ or $\{(i, j)|b_i < s_j, x(b_i, s_j) > 0\}$ is nonempty. We assume that the set $\{(i, j)|b_i > s_j, x(b_i, s_j) < 1\}$ is not empty and reach a contradiction. A similar contradiction obtains if we assume that the second set is non-empty.

Let $(i, j) \in \arg\max_{(i,j)}\{i - j|b_i > s_j, x(b_i, s_j) < 1\}$. Let i' (j') be the highest (lowest) type of buyer (seller) with $\bar{x}_b(b_{i'}) = \bar{x}_b(b_i)$ ($\bar{x}_s(s_{j'}) = \bar{x}_s(s_j)$); we have $i' \geq i, j' \leq j$. Using arguments similar to those for Case 1, since $\bar{x}_b(b_{i'}) < \bar{x}_b(b_{i'+1})$ and $\bar{x}_s(s_{j'-1}) > \bar{x}_s(s_{j'})$, one can specify the transfer function t such that IR_{b_1}, IR_{s_n}, $IC_{b_{i'} \to b_{i'+1}}$, $IC_{b_{i'+1} \to b_{i'}}$, $IC_{s_{j'} \to s_{j'-1}}$ and $IC_{s_{j'-1} \to s_{j'}}$ hold strictly under (x, t). We seek a perturbation of the mechanism (x, t) that increases $x(b_i, s_j)$ while preserving the monotonicity of $\bar{x}_b$ and $\bar{x}_s$ as in the proof of Proposition 2.

If $i \neq i'$, then the definition of (i, j) implies that $x(b_i, s_j) < 1 = x(b_{i'}, s_j)$. Since $\bar{x}_b(b_{i'}) = \bar{x}_b(b_i)$, there exists j'' such that $x(b_i, s_{j''}) > x(b_{i'}, s_{j''})$. Analogously, if $j \neq j'$, then there exists i'' such that $x(b_{i''}, s_j) > x(b_{i''}, s_{j'})$. In particular, $x(b_i, s_j) < 1, x(b_i, s_{j''}) > 0, x(b_{i'}, s_{j''}) <$

$1, x(b_{i''}, s_j) > 0, x(b_{i''}, s_{j'}) < 1$. Let x' be a perturbation of x for the following types:

$$\begin{aligned}
x'(b_i, s_j) &= x(b_i, s_j) + \varepsilon \\
x'(b_i, s_{j''}) &= x(b_i, s_{j''}) - \frac{p_s(s_j)}{p_s(s_{j''})}\varepsilon \\
x'(b_{i'}, s_{j''}) &= x(b_{i'}, s_{j''}) + \frac{p_b(b_i)}{p_b(b_{i'})}\frac{p_s(s_j)}{p_s(s_{j''})}\varepsilon \\
x'(b_{i''}, s_j) &= x(b_{i''}, s_j) - \frac{p_b(b_i)}{p_b(b_{i''})}\varepsilon \\
x'(b_{i''}, s_{j'}) &= x(b_{i''}, s_{j'}) + \frac{p_s(s_j)}{p_s(s_{j'})}\frac{p_b(b_i)}{p_b(b_{i''})}\varepsilon.
\end{aligned}$$

If $i = i'$, then we set $i = i' = i''$ and $x'(b_i, s_j) = x(b_i, s_j) + \varepsilon$ hereafter. We proceed analogously for $j = j'$.

The selection of i, j, i', j', i'', j'' guarantees that all values of x' are in $[0, 1]$ for small ε. Note that $\bar{x}'_b$ differs from $\bar{x}_b$ only at $b_{i'}$, and $\bar{x}'_s$ differs from $\bar{x}_s$ only at $s_{j'}$, with the differences computed as follows:

$$\begin{aligned}
\bar{x}'_b(b_{i'}) &= \bar{x}_b(b_{i'}) + \frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon \\
\bar{x}'_s(s_{j'}) &= \bar{x}_s(s_{j'}) + \frac{p_s(s_j)}{p_s(s_{j'})}p_b(b_i)\varepsilon.
\end{aligned}$$

Since IR_{b_1}, IR_{s_n}, $IC_{b_{i'} \to b_{i'+1}}$, $IC_{b_{i'+1} \to b_{i'}}$, $IC_{s_{j'} \to s_{j'-1}}$ and $IC_{s_{j'-1} \to s_{j'}}$ hold strictly under (x, t), the only constraints from Lemma 1 that (x', t) might fail for small $\varepsilon > 0$ are $IC_{b_{i'-1} \to b_{i'}}$ and $IC_{s_{j'} \to s_{j'+1}}$. In order to relax these constraints and make x' an implementable allocation, consider a perturbation t' that coincides with t' with the following exceptions:

$$\begin{aligned}
t'(b_{i'}, s_{j'}) &= t(b_{i'}, s_{j'}) + b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon + s_{j'}\frac{p_s(s_j)}{p_s(s_{j'})}p_b(b_i)\varepsilon \\
t'(b_k, s_{j'}) &= t(b_k, s_{j'}) + s_{j'}\frac{p_s(s_j)}{p_s(s_{j'})}p_b(b_i)\varepsilon, \forall k \neq i' \\
t'(b_{i'}, s_g) &= t(b_{i'}, s_g) + b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon, \forall g \neq j'.
\end{aligned}$$

We have that

$$\begin{aligned}
\bar{t}'_b(b_{i'}) &= \bar{t}_b(b_{i'}) + s_{j'}p_s(s_j)p_b(b_i)\varepsilon + b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon \\
\bar{t}'_s(s_{j'}) &= \bar{t}_s(s_{j'}) + b_{i'}p_b(b_i)p_s(s_j)\varepsilon + s_{j'}\frac{p_s(s_j)}{p_s(s_{j'})}p_b(b_i)\varepsilon \\
\bar{t}'_b(b_k) &= \bar{t}_b(b_k) + s_{j'}p_s(s_j)p_b(b_i)\varepsilon, \forall k \neq i' \\
\bar{t}'_s(s_g) &= \bar{t}_s(s_g) + b_{i'}p_b(b_i)p_s(s_j)\varepsilon, \forall g \neq j'.
\end{aligned}$$

By continuity, for small $\varepsilon > 0$, $IR_{b_1}, IR_{s_n}, IC_{b_{i'}\to b_{i'+1}}, IC_{b_{i'+1}\to b_{i'}}, IC_{s_{j'}\to s_{j'-1}}$ and $IC_{s_{j'-1}\to s_{j'}}$ hold strictly under (x', t'). All constraints that do not involve $b_{i'}$ or $s_{j'}$ also hold because at those values $\bar{x}'_b$ and $\bar{x}'_s$ are identical to $\bar{x}_b$ and $\bar{x}_s$, and $\bar{t}'_b$ and $\bar{t}'_s$ are translations of $\bar{t}_b$ and $\bar{t}_s$, respectively. The constraints for (x', t') from Lemma 1 that we are left to check are $IC_{b_{i'}\to b_{i'-1}}$, $IC_{b_{i'-1}\to b_{i'}}$, $IC_{s_{j'}\to s_{j'+1}}$, and $IC_{s_{j'+1}\to s_{j'}}$. The constraints $IC_{b_{i'}\to b_{i'-1}}$ and $IC_{b_{i'-1}\to b_{i'}}$ for (x', t') are equivalent to the following chain of inequalities

$$b_{i'-1}(\bar{x}'_b(b_{i'}) - \bar{x}'_b(b_{i'-1})) \leq \bar{t}'_b(b_{i'}) - \bar{t}'_b(b_{i'-1}) \leq b_{i'}(\bar{x}'_b(b_{i'}) - \bar{x}'_b(b_{i'-1})),$$

which reduces to

$$\begin{aligned}
b_{i'-1}(\bar{x}_b(b_{i'}) - \bar{x}_b(b_{i'-1})) + b_{i'-1}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon &\leq \bar{t}_b(b_{i'}) - \bar{t}_b(b_{i'-1}) + b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon \\
&\leq b_{i'}(\bar{x}_b(b_{i'}) - \bar{x}_b(b_{i'-1})) + b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon.
\end{aligned}$$

The first inequality follows from $b_{i'-1} < b_{i'}$ and the fact that (x, t) satisfies $IC_{b_{i'-1}\to b_{i'}}$, while the second inequality is equivalent to $IC_{b_{i'}\to b_{i'-1}}$ for (x, t). The constraints $IC_{s_{j'}\to s_{j'+1}}$ and $IC_{s_{j'+1}\to s_{j'}}$ for (x', t') are checked similarly.

Therefore, (x', t') is an implementable mechanism. The welfare difference between (x', t') and (x, t) depends on the differences in $\bar{x}'_b$ and $\bar{x}_b$ at $b_{i'}$, and $\bar{x}'_s$ and $\bar{x}_s$ at $s_{j'}$, and is given by

$$p_b(b_{i'})b_{i'}\frac{p_b(b_i)}{p_b(b_{i'})}p_s(s_j)\varepsilon - p_s(s_{j'})s_{j'}\frac{p_s(s_j)}{p_s(s_{j'})}p_b(b_i)\varepsilon = p_b(b_i)p_s(s_j)(b_{i'} - s_{j'})\varepsilon.$$

As $b_i > s_j$, $i' \geq i$, and $j' \leq j$, we have that $b_{i'} > s_{j'}$, which implies that the expression above is positive. This contradicts the assumed optimality of the mechanism (x, t). □

Computations for Proposition 4 and footnote 5. Here we find the optimal implementable mechanisms in the example used to demonstrate Proposition 4 and in an example supporting footnote 5. Assume that the buyer's valuations are $b_1 = 1$ and $b_2 = 4$ with probability 1/2 each, and that the seller's valuations are $s_1 = \varepsilon_1$ and $s_2 = 3 + \varepsilon_2$ with probability 1/2 each, where $\varepsilon_1, \varepsilon_2$ are parameters in $[0, 1)$.

We first solve for the optimal mechanism that satisfies the binding constraints from Lemma 3 with equality and then argue that the solution we find is an optimal implementable mechanism. The optimal mechanism with relaxed constraints solves the following linear program:

$$\begin{aligned}
\tfrac{1}{4}\max \quad & (1-\varepsilon_1)x(b_1,s_1)+(-2-\varepsilon_2)x(b_1,s_2)+(4-\varepsilon_1)x(b_2,s_1)+(1-\varepsilon_2)x(b_2,s_2)+2(\varepsilon_1+3+\varepsilon_2)\\
\text{s.t.} \quad & x(b_1,s_1)-t(b_1,s_1)+x(b_1,s_2)-t(b_1,s_2)=0\\
& t(b_1,s_2)-(3+\varepsilon_2)x(b_1,s_2)+t(b_2,s_2)-(3+\varepsilon_2)x(b_2,s_2)=0\\
& 4x(b_2,s_1)-t(b_2,s_1)+4x(b_2,s_2)-t(b_2,s_2)=4x(b_1,s_1)-t(b_1,s_1)+4x(b_1,s_2)-t(b_1,s_2)\\
& t(b_1,s_1)-\varepsilon_1x(b_1,s_1)+t(b_2,s_1)-\varepsilon_1x(b_2,s_1)=t(b_1,s_2)-\varepsilon_1x(b_1,s_2)+t(b_2,s_2)-\varepsilon_1x(b_2,s_2)\\
& x(b_1,s_1),x(b_1,s_2),x(b_2,s_1),x(b_2,s_2)\in[0,1].
\end{aligned}$$

The equality constraints, correspond in order, to the binding constraints $IR_{b_1}, IR_{s_2}, IC_{b_2\to b_1}$ and $IC_{s_1\to s_2}$ (multiplied by 2).

Rearranging, the system of binding constraints becomes

$$\begin{aligned}
t(b_1,s_1)+t(b_1,s_2) &= x(b_1,s_1)+x(b_1,s_2)\\
t(b_1,s_2)+t(b_2,s_2) &= (3+\varepsilon_2)(x(b_1,s_2)+x(b_2,s_2))\\
t(b_1,s_1)+t(b_1,s_2)-t(b_2,s_1)-t(b_2,s_2) &= 4(x(b_1,s_1)+x(b_1,s_2)-x(b_2,s_1)-x(b_2,s_2))\\
t(b_1,s_1)-t(b_1,s_2)+t(b_2,s_1)-t(b_2,s_2) &= \varepsilon_1(x(b_1,s_1)-x(b_1,s_2)+x(b_2,s_1)-x(b_2,s_2)),
\end{aligned}$$

or

$$\begin{pmatrix} 1 & 1 & 0 & 0\\ 0 & 1 & 0 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1 \end{pmatrix}\begin{pmatrix} t(b_1,s_1)\\ t(b_1,s_2)\\ t(b_2,s_1)\\ t(b_2,s_2)\end{pmatrix}=\begin{pmatrix} x(b_1,s_1)+x(b_1,s_2)\\ (3+\varepsilon_2)(x(b_1,s_2)+x(b_2,s_2))\\ 4(x(b_1,s_1)+x(b_1,s_2)-x(b_2,s_1)-x(b_2,s_2))\\ \varepsilon_1(x(b_1,s_1)-x(b_1,s_2)+x(b_2,s_1)-x(b_2,s_2))\end{pmatrix}.$$

The linear system above has rank 3, and the vector $(-2,2,1,1)$ belongs to the left null space of the coefficient matrix. The system has a solution if and only if the rank of the extended matrix is 3 as well, which is equivalent to

$$(-2,2,1,1)\begin{pmatrix} x(b_1,s_1)+x(b_1,s_2)\\ (3+\varepsilon_2)(x(b_1,s_2)+x(b_2,s_2))\\ 4(x(b_1,s_1)+x(b_1,s_2)-x(b_2,s_1)-x(b_2,s_2))\\ \varepsilon_1(x(b_1,s_1)-x(b_1,s_2)+x(b_2,s_1)-x(b_2,s_2))\end{pmatrix}=0,$$

or

$$(9)\quad (2+\varepsilon_1)x(b_1,s_1)+(8+2\varepsilon_2-\varepsilon_1)x(b_1,s_2)+(2+2\varepsilon_2-\varepsilon_1)x(b_2,s_2)=(4-\varepsilon_1)x(b_2,s_1).$$

The system has a solution $t(b_1, s_1), t(b_1, s_2), t(b_2, s_1), t(b_2, s_2)$ if and only if (9) is satisfied.

Substituting (9) in the objective function, the optimization problem becomes

$$\begin{aligned} \tfrac{1}{4}\max \quad & 3x(b_1, s_1) + (6 + \varepsilon_2 - \varepsilon_1)x(b_1, s_2) + (3 + \varepsilon_2 - \varepsilon_1)x(b_2, s_2) + 2(\varepsilon_1 + 3 + \varepsilon_2) \\ \text{s.t.} \quad & (2 + \varepsilon_1)x(b_1, s_1) + (8 + 2\varepsilon_2 - \varepsilon_1)x(b_1, s_2) + (2 + 2\varepsilon_2 - \varepsilon_1)x(b_2, s_2) = (4 - \varepsilon_1)x(b_2, s_1) \\ & x(b_1, s_1), x(b_1, s_2), x(b_2, s_1), x(b_2, s_2) \in [0, 1]. \end{aligned}$$

At the optimum, we need $x(b_1, s_2) = 0$. Indeed, if $x(b_1, s_2) > 0$ we can increase the value of the objective and satisfy (9) by decreasing $x(b_1, s_2)$ by a small positive δ, and adding $\delta(8 + 2\varepsilon_2 - \varepsilon_1)/(2 + \varepsilon_1)$ to $x(b_1, s_1)$ or $\delta(8 + 2\varepsilon_2 - \varepsilon_1)/(2 + 2\varepsilon_2 - \varepsilon_1)$ to $x(b_2, s_2)$ (if both $x(b_1, s_1)$ and $x(b_2, s_2)$ equal 1, then constraint (9) is violated when $\varepsilon_1 \geq 0, \varepsilon_2 \geq 0$ with at least one strict inequality). By a similar argument, $x(b_2, s_1) = 1$.

The problem reduces to

$$\begin{aligned} \tfrac{1}{4}\max \quad & 3x(b_1, s_1) + (3 + \varepsilon_2 - \varepsilon_1)x(b_2, s_2) + 2(\varepsilon_1 + 3 + \varepsilon_2) \\ \text{s.t.} \quad & (2 + \varepsilon_1)x(b_1, s_1) + (2 + 2\varepsilon_2 - \varepsilon_1)x(b_2, s_2) = 4 - \varepsilon_1 \\ & x(b_1, s_1), x(b_2, s_2) \in [0, 1]. \end{aligned}$$

We now specialize the analysis to two examples. For the example used to illustrate Proposition 4, we set $\varepsilon_2 = 0$. The unique solution to the linear program above in this case is $x(b_1, s_1) = 2/(2 + \varepsilon_1), x(b_2, s_2) = 1$. It is easily checked that the allocation $x(b_1, s_1) = 2/(2 + \varepsilon_1), x(b_1, s_2) = 0, x(b_2, s_1) = 1, x(b_2, s_2) = 1$ can be implemented with transfers that solve the linear system above. For $\varepsilon_1 = 0$, x is the ex post efficient allocation, so clearly (x, t) is an optimal mechanism.

We next argue that (x, t) is also an optimal implementable mechanism for $\varepsilon_1 \in (0, 1)$. It is sufficient to show that the ex post efficient allocation is not implementable in this case. Then, Lemma 3 implies that any optimal implementable mechanism must satisfy the equality constraints we imposed in the linear program, and hence (x, t) is optimal. For a proof by contradiction, assume that there exists a mechanism (x', t') that implements the ex post efficient allocation $x'(b_1, s_1) = 1, x'(b_1, s_2) = 0, x'(b_2, s_1) = 1, x'(b_2, s_2) = 1$ for $\varepsilon_1 \in (0, 1)$. Since $\bar{x}'_b(b_1) = 1/2$ and $\bar{x}'_b(b_2) = 1$, $IC_{b_2 \to b_1}$ and IR_{b_1} imply that the expected payoff of type b_2 under (x', t') is $4 - \bar{t}'_b(b_2) \geq 4 \times 1/2 - \bar{t}'_b(b_1) = 3/2 + 1 \times 1/2 - \bar{t}'_b(b_1) \geq 3/2$. Similarly, the net expected benefit of type s_1 from trading under (x', t') is $\bar{t}'_s(s_1) - \varepsilon_1 \times 1 \geq \bar{t}'_s(s_2) - \varepsilon_1 \times 1/2 = 3/2 - \varepsilon_1/2 + \bar{t}'_s(s_2) - 3 \times 1/2 \geq 3/2 - \varepsilon_1/2$. Hence, the expected gains from trade under (x', t') should be at least $1/2(3/2 + 3/2 - \varepsilon_1/2)$, which is greater than the gains x' generates, $1/4((1 - \varepsilon_1) + (4 - \varepsilon_1) + (4 - 3))$, a contradiction.

For an example corroborating footnote 5, let $\varepsilon_1 = \varepsilon_2 = \varepsilon$ in the arguments above (a similar conclusion can be reached for cases with $\varepsilon_1 < \varepsilon_2$). The linear program has multiple solutions, one of which is $x(b_1, s_1) = (2-2\varepsilon)/(2+\varepsilon), x(b_1, s_2) = 0, x(b_2, s_1) = 1, x(b_2, s_2) = 1$. Arguments similar to those for the first example show there exists an optimal mechanism that implements x for small ε.

The allocation x achieves a total welfare of $(24+11\varepsilon+4\varepsilon^2)/(4(2+\varepsilon))$, which is decreasing for small $\varepsilon \geq 0$. For instance, the welfare generated by x is 3 for $\varepsilon = 0$ and approximately 2.993 for $\varepsilon = 0.1$. Note that following the damage implied by the decrease in ε from 0.1 to 0, the optimal mechanism satisfies the seller's ex ante damage participation constraint described in footnote 5. Indeed, the seller's expected payoff from participating in the post-damage optimal mechanism is 2.25, which is greater than his expected value 1.6 for the undamaged good. □

Proof of Proposition 5. Fix a buyer value distribution p_b, and consider a seller value distribution p'_s with support $s'_1 < s'_2 < \ldots < s'_n$ that is a monotonic reduction of the distribution p_s with support $s_1 < s_2 < \ldots < s_n$. Let (x', t') be an implementable mechanism for the pair of distributions (p_b, p'_s) that satisfies the seller's damage participation constraint for the value reduction from p_s to p'_s. We construct a mechanism (x, t) for the pair of distributions (p_b, p_s) that implements the same type-by-type allocation x', i.e., $x(b_i, s_j) = x'(b_i, s'_j)$ for all i and j, and perturbs the transfer function t' as follows:

$$t(b_i, s_j) = t'(b_i, s'_j) - (1-\bar{x}_s(s_n))(s_n - s'_n) + \sum_{l=j}^{n-1}(\bar{x}_s(s_l) - \bar{x}_s(s_{l+1}))(s_l - s'_l), \forall i = \overline{1,m}, j = \overline{1,n}.$$

Since x and x' represent the same allocation of the good for any profile of corresponding types, and seller values under p_s are higher than under p'_s, the mechanism (x', t') with value distributions (p_b, p'_s) does not achieve greater welfare than the mechanism (x, t) with value distributions (p_b, p_s).

We prove that (x, t) is an implementable mechanism for the pair of distributions (p_b, p_s). It suffices to verify that (x, t) satisfies the constraints from Lemma 1. The mechanism (x, t) satisfies the incentive constraints for the buyer because the corresponding incentive constraints hold for (x', t'), and $\bar{x}_b$ coincides with $\bar{x}'_b$, while $\bar{t}_b(b_i) - \bar{t}'_b(b_i) = \bar{t}_b(b_k) - \bar{t}'_b(b_k)$ for $i, k = \overline{1,m}$. The mechanism (x, t) satisfies the individual rationality constraint for buyer b_1

since $t(b_1, s_j) \leq t'(b_1, s_j)$ for $j = \overline{1, n}$. This inequality holds because

$$\sum_{l=j}^{n-1} (\bar{x}_s(s_l) - \bar{x}_s(s_{l+1}))(s_l - s'_l) \leq \sum_{l=j}^{n-1} (\bar{x}_s(s_l) - \bar{x}_s(s_{l+1}))(s_n - s'_n)$$
$$= (\bar{x}_s(s_j) - \bar{x}_s(s_n))(s_n - s'_n) \leq (1 - \bar{x}_s(s_n))(s_n - s'_n), \forall j = \overline{1, n}$$

as $s_l - s'_l \leq s_n - s'_n$, $\bar{x}_s(s_l) - \bar{x}_s(s_{l+1}) \geq 0$ for $l = \overline{j, n-1}$, and $\bar{x}_s(s_j) \leq 1$.

We are left to check that (x, t) satisfies the constraints IR_{s_n}, $IC_{s_j \to s_{j+1}}$, and $IC_{s_{j+1} \to s_j}$. The constraint IR_{s_n} holds under (x, t) since the damage participation constraint for seller type n (whose value has been reduced from s_n to s'_n) under the mechanism (x', t') requires that $\bar{t}'_s(s'_n) + (1 - \bar{x}'_s(s'_n))s'_n \geq s_n$, and the definition of (x, t) implies that $\bar{t}'_s(s'_n) + (1 - \bar{x}'_s(s'_n))s'_n = \bar{t}_s(s_n) + (1 - \bar{x}_s(s_n))s_n$.

To show that (x, t) satisfies $IC_{s_j \to s_{j+1}}$ and $IC_{s_{j+1} \to s_j}$, note that the corresponding constraints under (x', t') imply that

$$(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s'_j \leq \bar{t}'_s(s'_j) - \bar{t}'_s(s'_{j+1}) \leq (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s'_{j+1}.$$

Since $\bar{t}_s(s_j) - \bar{t}_s(s_{j+1}) = \bar{t}'_s(s'_j) - \bar{t}'_s(s'_{j+1}) + (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))(s_j - s'_j)$, the chain of inequalities above implies that

$$(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s'_j + (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))(s_j - s'_j)$$
$$\leq \bar{t}_s(s_j) - \bar{t}_s(s_{j+1}) \leq (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s'_{j+1} + (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))(s_j - s'_j),$$

which along with the hypothesis that $s_j - s'_j \leq s_{j+1} - s'_{j+1}$ leads to

$$(\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s_j \leq \bar{t}_s(s_j) - \bar{t}_s(s_{j+1}) \leq (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))(s'_{j+1} + s_j - s'_j)$$
$$\leq (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))(s'_{j+1} + s_{j+1} - s'_{j+1}) = (\bar{x}_s(s_j) - \bar{x}_s(s_{j+1}))s_{j+1}.$$

This proves that (x, t) satisfies $IC_{s_j \to s_{j+1}}$ and $IC_{s_{j+1} \to s_j}$.

We have argued that (x, t) is an implementable mechanism for the pair of value distributions (p_b, p_s) that generates at least the same amount of welfare as (x', t') for the pair of value distributions (p_b, p'_s). Therefore, the optimal implementable mechanism for the value distributions (p_b, p_s) achieves at least the same amount of welfare as the mechanism (x', t') for the value distributions (p_b, p'_s). □

References

1. Condorelli, C. and B. Szentes (2019) Information Design in the Hold-up Problem, forthcoming in *J. Pol. Econ.*

2. Cramton, P., R. Gibbons, and P. Klemperer (1987) Dissolving a Partnership Efficiently, *Econometrica*, **55**, 615-632.
3. Kos, N. and M. Manea (2009) Ex Post Efficient Trade Mechanisms with Discrete Types, mimeo.
4. Laffont, J.-J. and E. Maskin (1979) A Differential Approach to Expected Utility Maximizing Mechanisms, Ch. 16 in *Aggregation and Revelation of Preferences* (ed: Laffont, J.-J.), North-Holland Publishing Company.
5. Makowski, L. and C. Mezzetti (1993) The Possibility of Efficient Mechanisms for Trading an Indivisible Object, *J. Econ. Theory*, **59**, 451-465.
6. Matsuo, T. (1989) On Incentive Compatibility, Individual Rationality, and Ex Post Efficient Mechanisms for Bilateral Trading, *J. Econ. Theory*, **49**, 189-194.
7. Myerson, R.B. and M.A. Satterthwaite (1983) Efficient Mechanisms for Bilateral Trading, *J. Econ. Theory*, **29**, 265-281.
8. Postlewaite, A. (1979) Manipulation via Endowments, *Rev. Econ. Stud.*, **46**, 255-62.