# Optimal Bilateral Trade of Multiple Objects*


Ran Eilat — Tel Aviv University

Ady Paunzer — Tel Aviv University


May 2008


**Abstract**

We consider a private-values buyer-seller problem with multiple objects. Valuations are binary, i.i.d., and such that the problem does not have a trivial solution. We characterize mechanisms that span the Pareto frontier. These have a very simple form: Call the seller "good" if he has a low valuation and "bad" if he has a high valuation. Call the buyer "good" if he has a high valuation and "bad" if he has a low valuation. For each object, if both say "bad" – there is no trade. If both say "good" – they trade. If agent $j$ says "bad" and the other says "good", they trade only if the number of good objects in $j$'s announcement is above a certain threshold (at the threshold itself they trade with probability between 0 and 1). The thresholds depend on the weights given to each agent in the designer's objective function: as she leans more towards one of the agents, his trading threshold weakly decreases, and the rival's weakly increases.



*We thank Eddie Dekel, Matt Jackson, Nadav Levy and semimar participants at Tel Aviv University for helpful comments and discussions.

# 1 Introduction

Bilateral trade with private values is a fundamental problem in mechanism design. Myerson and Satterthwaite (1983) have shown that in any non-trivial buyer-seller problem with continuous valuations, full efficiency cannot be attained. Also with discrete (finite) valuations, for many parameter values the first-best outcome cannot be reached.

Jackson and Sonnenschein (2007) have recently shown that with many objects (and valuations that are independent across objects), the problem disappears in the limit: the first-best outcome can be approached as the number of objects tends to infinity.[1,2]. While their mechanism approaches the first best in the limit, for a finite number of objects it is not the optimal one (not second best).

In this paper we characterize the set of second-best outcomes for bilateral trade of multiple objects. That is, for any pair of weights, we find the second-best mechanism that maximizes the respective weighted sum of the buyer's and seller's ex ante gains from trade. We confine ourselves to the case of binary valuations.

The basic tension in the buyer-seller problem, as mirrored in Myerson and Satterthwaite's proof, is that the information rents that must be paid to agents in order to induce them to reveal their types (incentive compatibility) and agree to participate (individual rationality), are higher than the surplus generated by trade. Myerson and Satterthwaite show that the optimal (second-best) solution involves participation restrictions, i.e., not allowing beneficial trade in some cases. That is done by reducing the trade probability for types who gain less from trade; As a result, the incentives of types who gain more from trade to pretend to be of the former kind are relaxed, and thus they demand lower rents. While restricting participation also reduces the surplus from trade, rents are affected more. The optimal solution involves the minimal restrictions of trade that suffice to equate rents with surplus.[3]

What is the optimal way of restricting participation when there are multiple objects? Trading each object separately, using for each the one-object-optimal mechanism, is not the best solution.

[1] Jackson and Sonnenschein treat a more general collective decision problem; in this context their mechanism works even without monetary transfers between agents.

[2] Athey and Miller (2007) study a different but related question: a sequential trading problem with infinitely many periods. They show that the first-best outcome can be attained (for discount factors above 1/2) with ex-ante budget balance. With ex post budget balance only the one-shot outcome is feasible, and with a bounded, collateralized credit line the first-best outcome can be approached as the discount factor tends to 1.

[3] Specifically, Myerson and Satterthwaite show that the mechanism that maximizes the expected gains from trade restricts trade whenever the difference between the announced valuations is below some threshold.

Rather, one can link the different trade problems, allowing the restrictions on the trade of one object to depend on the (announced) valuations of the other objects. This paper identifies the best way of allocating the trade restrictions, optimally trading off the reduced rents and lost surplus. The resulting mechanisms, that together span the Pareto frontier of agents' utilities, share a very simple structure:

Call the seller "good" if he has a low valuation and "bad" if he has a high valuation. Call the buyer "good" if he has a high valuation and "bad" if he has a low valuation. For each object, if both say "bad" – there is no trade. If both say "good" – they trade. If agent $j$ says "bad" and the other says "good", there is trade only if the total number of good valuations in agent $j$'s announcement is more than $M_j$. (At exactly $M_j$ "good" announcements there can be trade with probability less than one). In other words, whenever an agent announces "bad", he is allowed to trade (in case the other agent announced "good") only if there were sufficiently many other objects for which he announced "good".

The respective thresholds for the two agents, i.e., the numbers $M_j$ of good announcements that are required to get the permission to trade at bad valuations, depend on the weights given to the utilities of the two agents in the objective function of the mechanism designer. The higher an agent's weight, the lower his threshold. At the extreme cases, when the mechanism strongly prefers one agent (approaching the monopoly/monopsony solution), that agent's threshold is 0. This means that the mechanism allows him to trade whatever his announcement. The only trade restrictions come, in such a case, from the disfavored agent (the buyer facing a monopolist or seller facing a monopsonist). If there are only few objects, the disfavored agent might be allowed to trade only his good-valuation objects. But if there are many objects, his trade threshold is less than the number of objects. This means that he sometimes trades his bad objects, and thus obtains positive rents. This case corresponds to mixed bundling in the monopoly literature.

In related and independent work, Fang and Norman (2008) study the problem of provision of multiple public goods when exclusion is allowed. That is, there are many potential users with private information on their willingness to pay, and after they report their types the government decides whether to produce each public good, whether to exclude certain agents from using it, and how much to charge each one. For the case of binary valuations, they characterize the optimal mechanism (with equal weights on all consumers). They show that as the number of agents grows without bounds: (1) Either all or none of the public goods is provided. (2) Each consumer with a low valuation is excluded iff he announced low valuations too many times. Note however that while there are strong mathematical connections between the buyer-seller problem and the public-goods problem (with two agents and without exclusion), Fang and Norman's result regarding the exclusion rule is not parallel to ours. This is because the public-goods analogue to trade in the buyer-seller problem is the provision, rather than exclusion. Yet, many of the fundamental ideas underlying our proofs are similar to those of Norman and Fang's.[4]

The remainder of the paper is organized as follow. In section 2 we present an example that illustrates our main result and the intuition for the construction of optimal mechanisms. Section 3 introduces the model and some preliminary results that allows us to simplify the analysis. In section 4 we derive the optimal mechanisms and present our main result. Section 5 presents some graphs that shed light on our results. The proofs of all the lemmas and propositions are relegated to the appendix.

## 2 Example

Let us consider a simple example that illustrates our main result and sheds some light on the intuition behind it. We start by constructing the optimal trade mechanism for the case of one object; we then analyze a replication of problem to three objects, and identify the optimal mechanism in this case. Importantly, this mechanism is not just a replication of the one-object optimal mechanism, but rather conditions the probability of trade in one object on the announcements regarding the other objects.

Consider first the single-object trading problem. Suppose that the seller's valuation of the object is either 11 or 0, While the buyer valuates it as either 12 or 1. Valuations are independently drawn, with equal probabilities for the high and the low valuation. Agents are risk neutral and each knows only his own valuation. To allow addressing both agents parallely, call the low valuation of the buyer (which is 1) and the high valuation of the seller (11) "bad", and call the high valuation of the buyer (12) and the low valuation of the seller (0) "good". That is, "good" refers to types that are more eager to trade.

The (direct) trading mechanism receives a message from each agent regarding his valuation, and determines the probability of trade and the payments. We look for a mechanism that is incentive compatible, individually rational and budget balanced.

[4] Fang and Norman (2006) deal with the provision and exlusion of two public goods and is only peripherally related to our paper.

The parameters in our example are set such that full efficiency cannot be attained. For that, the probability of trade should have been 1 for the three combinations of valuations in which the buyer values the object more than the seller (i.e., at least one agent is "good"), and with probability 0 when the buyer has the lower valuation (both agents are "bad"). The total expected surplus from trade would then be $(12 + 1 + 1)/4 = 3.5$. But this is insufficient to finance the information rents to agents, needed to ensure their participation and truth telling. To see this, let us compute the minimal expected rent to each agent. The expected payoff to the bad type must be at least 0 for him to participate. The good type can always pretend to be a bad one; in that case he obtains the same payoff as the bad type, plus the difference between their valuations, which is 11, times the probability 0.5 that the bad type trades (recall that the bad type only trades with the good type of the opponent). Thus, for him to reveal his true type, his rent must be at least $0.5 \cdot 11 = 5.5$. Consequently, the agent's expected rent must be at least $(5.5 + 0)/2 = 2.75$, and thus the sum of rents to both agents, 5.5, is more that the surplus of 3.5 generated by the mechanism.

To regain budget balance, we can reduce the probabilities of trade of bad-good encounters. If this probability is reduced to $p < 1$, the rent to good types is proportionally reduced, to $p \cdot 5.5$. The surplus created by the mechanism is also reduced, but in a smaller ratio: it is now $(12 + p \cdot 1 + p \cdot 1)/4 = 3 + 0.5 \cdot p$. By reducing $p$ to 0.6 the surplus exactly suffices to finance the rents, and equals 3.3 (which is less that the first-best surplus of 3.5).

Suppose now that the same buyer and seller wish to trade three objects. The valuations of the different objects are independently drawn from the same distribution. The argument of the single-object case shows again that the first-best outcome is unattainable. By trading each of the objects using the second-best mechanism identified above, one can reach a surplus of $3 \cdot 3.3$. However, we can fare better by allocating the (unavoidable) restrictions on trade in a better way, exploiting the richer structure of the agents' type space

Consider then the following mechanism. Trade in each object takes place according to the first-best plan, unless one of the agents reports that all his three valuations are bad. That is, each agent is allowed to trade his bad-valuation objects only if he reported at least one good valuation.

To compute the required rents, we first pay the bad-bad-bad type the minimal amount that makes him participate, which is 0. Now consider the types with exactly one good valuation. If they lie regarding their good object and pretend that its valuation is also bad, the mechanism will not let them trade it. Thus, we don't need to pay them any information rent, so they, also, are paid 0. Next, consider types with two good valuations. If they lie once, pretending that one of their good objects is bad, they will gain the valuation difference (11) time the probability that they will then trade this object (0.5 – the probability that the opponent has a good valuation for that object). The rent needed to prevent them from lying is thus $0.5 \cdot 11 = 5.5$. Note that these types gain nothing by lying twice, as then they will be treated as a bad-bad-bad type and not trade at all. Finally, the good-good-good type has to be paid twice the rent of 5.5, as he can lie regarding the valuations of two objects. Again, he cannot lie three times as then he will be banned from trade.

Since, out of an agent's 8 possible types, there are 3 with two good valuations and 1 with three good valuations, each agent's expected rent is $(3 \cdot 5.5 + 1 \cdot 11)/8$. Thus, the sum of the two rents is 6.875. This is less than the surplus created by the mechanism. To see this, let us compute the average probability of trade of bad objects. As there is one type with three bad valuations who trades bad objects with probability 0, three types with two bad valuations who trade bad objects with probability 1, and three with one bad valuations who trade with probability 1, the average probability is $\frac{1 \cdot 3 \cdot 0 + 3 \cdot 2 \cdot 1 + 3 \cdot 1 \cdot 1}{1 \cdot 3 + 3 \cdot 2 + 3 \cdot 1} = 0.75$. This probability is higher than 0.6 – the probability that an object with a bad valuation is traded with the one-object optimal mechanism.

The total surplus of this mechanism, $3 \cdot (12 + 0.75 \cdot 1 + 0.75 \cdot 1)/4 = 10.125$, is already higher than that of trading each object separately ($3 \cdot 3.3 = 9.9$). Moreover, we can increase it even more by allowing the bad-bad-bad types of the two agents a probability of trade that exactly equates the surplus and the rents. In the optimal mechanism, this probability is 26/74, and the expected surplus is approximately 10.25.

The mechanism that we devised maximizes the total surplus, which is the sum of the agents' utilities. To span the entire Pareto frontier of the trading problem, we should consider all the mechanisms that maximize weighted sums of the agents' payoffs, with weights $\alpha^s$ and $\alpha^b = 1 - \alpha^s$ on the seller's and buyer's utilities. How does the second-best mechanism change along the Pareto frontier?

First note that our choice above, to use the remaining surplus and let each agent's bad-bad-bad types to trade with probability 26/74, was only one of many optimal choices for the case where $\alpha^s = \alpha^b = 1/2$. In fact, the remaining surplus could have been used for any combination of trade probabilities for the seller's and the buyer's bad-bad-bad types, as long as the sum of the probabilities is $2 \cdot 26/74$. The way this probability is split determines the agents' rents, which are their utilities, and thus the utility-pair (seller and buyer, respectively) can be at any point between $(3.44, 6.82)$ and $(6.82, 3.44)$.

The remaining parts of the Pareto frontier are found by devising the optimal mechanisms for non-equal weights $\alpha^s$ and $\alpha^b$. Let us thus gradually increase $\alpha^s$ from 1/2 to 1 (a symmetric construction applies for $\alpha^b$ in $(1/2, 1]$). Given this problem's parameters, there are three additional relevant thresholds for $\alpha^s$. For $1/2 < \alpha^s < 148/281$, the only $\alpha$-optimal solution is to completely ban the buyer's bad-bad-bad type from trading, and allocating the probability of 52/74 solely to the seller's bad-bad-bad type. For all these $\alpha$'s we obtain the same point on the Pareto frontier, $(6.82, 3.44)$. At $\alpha^s = 148/281$, we are indifferent between the buyer's bad-bad-good types and the sellers bad-bad-bad. We can now transfer trade probability from the former to the latter, until the seller's bad types all trade with probability 1. Doing so, we move from the vertex $(6.82, 3.44)$ to the vertex $(8.25, 1.84)$. For $\alpha^s$ above 148/281, we continue to decrease the buyer's trade probabilities of bad-valuation objects, first from his bad-bad-good type (until $\alpha^s = 44/82$), and then from his bad-good-good type (until $\alpha^s = 11/19$). The net savings (reduced rent minus reduced surplus) is paid up-front to the seller. This way we span two more edges of the Pareto frontier. Finally, for larger $\alpha^s$, the buyer only trades good-valuation objects and receives no rent at all. All the surplus goes to the seller.[5]

# 3 Model and Preliminary Results

Consider a buyer-seller problem with $N$ objects, labeled $i \in I = \{1, ..., N\}$. Both agents are risk neutral. They value each object as either "*high*" or "*low*": the possible valuations of the seller $(s)$ are $v_h^s > v_l^s \geq 0$, and for the buyer $(b)$ they are $v_h^b > v_l^b > 0$. For non-triviality of the problem we assume that $v_h^b > v_h^s > v_l^b > v_l^s$. We also denote the difference between the possible valuations, for each agent $j \in \{b, s\}$, by $\Delta^j = v_h^j - v_l^j$.

To allow treating the buyer and seller in a similar way, for each object $i$ we say that the buyer's valuation of that object is "good" if it is high $(v_h^b)$ and "bad" if low $(v_l^b)$. For the seller, we say that his valuation is "good" if it is low $(v_l^s)$ and "bad" if it is high $(v_h^s)$. In other words, good valuations are associated with the stronger desire to trade (for a given price).

A type for agent $j \in \{b, s\}$ is a vector of valuations for each object, $w^j = \left(w_1^j, ..., w_N^j\right) \in W^j$. Valuations are independent across objects and agents. Specifically, each valuation $w_i^j$ takes the values $v_{good}^j$ and $v_{bad}^j$ with probabilities $q^j \in [0, 1]$ and $\left(1 - q^j\right)$, respectively. Agents are risk-neutral, and their utilities are additive: denoting the set of objects that are traded (transferred

[5] When there are sufficiently many objects, there are types of the buyer (those with many good valuations) whose trade of bad objects generates more surplus then rent to higher types who might mimic them. In this case these types are allowed to trade in their bad objects, even if $\alpha^s$ is 1; they thus receive a positive rent.

from the seller to the buyer) by $I_{TR} \subseteq I$ and the transfer to agent $j$ (which may be negative) by $t^j$, the agents' preferences are represented by

$$\begin{aligned} u^b\left(I_{TR}, t^b\right) &= \sum\nolimits_{i \in I_{TR}} w_i^b + t^b, \\ u^s\left(I_{TR}, t^s\right) &= \sum\nolimits_{i \in I_{TR}} -w_i^s + t^s. \end{aligned}$$

### 3.1 Direct Mechanism

By the revelation principle (Myerson 1979), we restrict our attention, without loss of generality, to incentive compatible direct mechanisms. A direct trade mechanism receives as input a vector of objects' valuations from each agent. It then determines for each object $i \in I$ whether it should be transferred from the seller to the buyer, and sets a monetary transfer (possibly negative) to each agent.

More formally, a mechanism $\Gamma$ specifies, for each pair of announcements $w^s, w^b$, the probability $p\left(i; w^s, w^b\right)$ of trade for each object $i$, and the monetary transfers to each of the agents, $t^s\left(w^s, w^b\right)$ and $t^b\left(w^s, w^b\right)$. In particular, the probability of trade in one object is allowed to depend on the valuations of other objects. Under this notation, agents' ex post utilities are given by:

$$\begin{aligned} U^b(w^s, w^b) &= \sum\nolimits_{i=1}^{N} p\left(i; w^s, w^b\right) w_i^b + t^b\left(w^s, w^b\right), \\ U^s(w^s, w^b) &= \sum\nolimits_{i=1}^{N} -p\left(i; w^s, w^b\right) w_i^s + t^s\left(w^s, w^b\right). \end{aligned}$$

For each agent $j$ and each of his types $w^j$, denote the interim expected probability of trade in each object $i$ and the expected transfer, under the mechanism $\Gamma = \langle p, t^s, t^b \rangle$, by:

$$\begin{aligned} \bar{p}^j\left(i, w^j\right) &= E_{w^{-j}}\left[p\left(i; w^s, w^b\right)\right], \\ \bar{t}^j\left(w^j\right) &= E_{w^{-j}}\left[t^j\left(w^s, w^b\right)\right]. \end{aligned}$$

where $E_{w^{-j}}$ denotes the expectation over all types of the rival agent $-j$ (if $j$ is the seller then $-j$ is the buyer, and vice-versa). The interim expected utilities of the seller or buyer of type $w^j$ who announces $\hat{w}^j$, are then:

$$\begin{aligned} EU^b(w^b, \hat{w}^b) &= \sum\nolimits_{i=1}^{N} \bar{p}^b\left(i, \hat{w}^b\right) w_i^b + \bar{t}^b\left(\hat{w}^b\right), \\ EU^s(w^s, \hat{w}^s) &= \sum\nolimits_{i=1}^{N} -\bar{p}^s\left(i, \hat{w}^s\right) w_i^s + \bar{t}^s\left(\hat{w}^s\right). \end{aligned} \tag{1}$$

[6]

The mechanism must satisfy incentive-compatibility:

$$EU^j(w^j, w^j) \geq EU^j\left(w^j, \hat{w}^j\right) \quad \text{for any } w^j, \hat{w}^j \in W^j,$$

and individual rationality:

$$EU^j\left(w^j\right) \equiv EU^j\left(w^j, w^j\right) \geq 0 \;\; \text{for any } w^j \in W^j.$$

Moreover, we require ex-post budget balance of the mechanism:

$$t^s\left(w^s, w^b\right) + t^b\left(w^s, w^b\right) \leq 0 \;\; \text{for any } w^s \in W^s,\; w^b \in W^b.$$

## 3.2 Allocation-Neutral Utility Representation

Note that the utility of both agents comes from two sources. One is the utility (positive or negative) that is obtained directly from the exchange of the objects, that is $\sum_{i=1}^{N} p\left(i, w^s, w^b\right) w_i^b$ for buyer and $\sum_{i=1}^{N} -p\left(i, w^s, w^b\right) w_i^s$ for seller. The second is the utility from the transfers, namely $t^b\left(w^b\right)$ and $t^s\left(w^s\right)$.

It is convenient to decompose the transfers $t^j$ to two parts, as follows:

$$\begin{aligned} t^b(w^s, w^b) &= \sum\nolimits_{i=1}^{N} -p\left(i, w^s, w^b\right) w_i^b + \tau^b\left(w^s, w^b\right) \\ t^s(w^s, w^b) &= \sum\nolimits_{i=1}^{N} p\left(i, w^s, w^b\right) w_i^s + \tau^s\left(w^s, w^b\right). \end{aligned} \tag{2}$$

The first term is just minus the utility from objects' exchange; by definition, the utility from the exchange itself plus this term is zero (assuming the agents report their true valuations). The second part, $\tau$ (which is implicitly defined by equation 2), is now the net utility of the agents. Denoting $\bar{\tau}^j\left(w^j\right) \equiv E_{w^{-j}}\left[\tau^j\left(w^s, w^b\right)\right]$, we have:

$$EU\left(w^j\right) = \bar{\tau}^j\left(w^j\right) \tag{3}$$

This decomposition of the transfers can be described as follows: if trade takes place, the agents immediately pay (or get paid) their full announced valuations. On top of these "offsetting payments", they pay or receive additional transfers $\tau^j$ which – given truth telling – now capture their net utility.

Importantly, as long as an agent reports his true valuations, he is indifferent regarding the mechanism's decisions over trades – his utility depends only on the (new) transfer function $\tau^j$. In case of a deviation from reporting the true valuation, he gains or loses from the trade (plus offsetting payment), and the transfer $\tau^j$ is also changed. Under this new notation the incentive compatibility constraint of agent $j$ can be written as:

$$\begin{aligned} \sum\nolimits_{i=1}^{N} \bar{p}^b\left(i, \hat{w}^b\right)\left(\hat{w}_i^b - w_i^b\right) &\geq \bar{\tau}^b\left(\hat{w}^b\right) - \bar{\tau}^b\left(w^b\right), \\ \sum\nolimits_{i=1}^{N} \bar{p}^s\left(i, \hat{w}^s\right)\left(\hat{w}_i^s - w_i^s\right) &\leq \bar{\tau}^s\left(w^s\right) - \bar{\tau}^s\left(\hat{w}^s\right). \end{aligned} \tag{4}$$

That is, the net gain from trading when the true valuation is $w_i^j$ and the offset payment is according to the reported valuation $\hat{w}_i^j$, must be less then the resulting change in the transfer $\tau^j$.

The individual rationality constraint, in the new notation, is simply:

$$\bar{\tau}^j\left(w^j\right) \geq 0 \text{ for any } w^j \in W^j.$$

### 3.3 Ex Post and Ex Ante Budget Balance

By Proposition 2 in Borgers and Norman (2008), since the agents are risk neutral and their types are independent, for any mechanism satisfying ex ante budget balance there exists a mechanism satisfying ex post budget balance with the same trading rule and same interim utilities to both agents. Intuitively, the two risk-neutral agents can insure the mechanism while keeping their interim utilities unchanged. We can, therefore, look for an optimal mechanism in the domain of ex ante budget balanced ones:

$$E_{w^b, w^s}\left[t^s\left(w^s, w^b\right) + t^b\left(w^s, w^b\right)\right] \leq 0.$$

Denoting the "revenue" of the mechanism by:

$$\begin{aligned} R\left(w^s, w^b\right) &= -\left(t^s(w^s, w^b) + t^b(w^s, w^b)\right) = \\ &= \sum\nolimits_{i=1}^{N} p\left(i, w^s, w^b\right)\left(w_i^b - w_i^s\right) - \tau^b\left(w^s, w^b\right) - \tau^s\left(w^s, w^b\right), \end{aligned} \tag{5}$$

ex ante budget balance is simply:

$$E_{w^b, w^s}\left[R\left(w^s, w^b\right)\right] \geq 0. \tag{6}$$

### 3.4 Symmetric Mechanisms

Since our trading problem is symmetric across objects (each agent's valuations of objects are i.i.d.), we can restrict attention to symmetric mechanisms. Such mechanisms are agnostic to changing the names of the objects.

Let $\pi : I \to I$ denote a permutation mapping, and let $\Pi$ be the set of all $N!$ possible permutations of the set $I$. Let $M_\pi$ denote the corresponding permutation operator (matrix) on vectors, so that $M_\pi w$ is a vector of valuations in which the $i^{th}$ element is the $\pi(i)^{th}$ element in $w$.[7]

**Definition 1** *A mechanism* $\Gamma = \langle p, \tau^s, \tau^b \rangle$ *is symmetric if for all* $w^s \in W^s$, $w^b \in W^b$, *and* $\pi \in \Pi$*:*

1. $p(i; w^s, w^b) = p(\pi(i); M_\pi w^s, M_\pi w^b)$
2. $\tau^s(w^s, w^b) = \tau^s(M_\pi w^s, M_\pi w^b)$
3. $\tau^b(w^s, w^b) = \tau^b(M_\pi w^s, M_\pi w^b)$

**Proposition 1** *For every incentive compatible mechanism* $\Gamma = (p, \tau^s, \tau^b)$*, there exists a mechanism* $\hat{\Gamma} = (\hat{p}, \hat{\tau}^s, \hat{\tau}^b)$ *that is incentive compatible, symmetric and provides the same ex ante utilities for the agents and same ex ante revenue for the mechanism.*

In a symmetric mechanism, the probability of trade in each object depends only on the agents' valuations of that object and on the numbers of good valuations in each agent's type. Moreover, the transfers to agents depend only on the number of good valuations in their types. More formally, let $g(w^j)$ denote the number of "good" valuations in $w^j$. Then the probability of trade $p(i, w^s, w^b)$ depends only on $w_i^s$, $w_i^b$, $g(w^s)$ and $g(w^b)$, while the transfers $\tau^j(w^s, w^b)$ depend only on $g(w^s)$ and $g(w^b)$. It is thus convenient to partition the set $W^j$ of all types of agent $j$ to $N+1$ equivalence classes $\left\{G_m^j\right\}_{m=0}^N$ such that all types $w^j$ in $G_m^j$ have the same number $g(w^j) = m$ of "good" elements. An arbitrary type in $G_m^j$ is denoted by $w^{j,m}$.

---

[7] Formally, $M_\pi$ is a matrix whose $i^{th}$ row is the $\pi^{-1}(i)^{th}$ row of the unit matrix of size $N \times N$, i.e., $[M_\pi]_{(i,j)} = [I_{N\times N}]_{(\pi^{-1}(i),j)}$. For example, if $\pi(1) = 3$, $\pi(2) = 1$, $\pi(3) = 2$, then the corresponding permutation matrix is $M_\pi = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

# 4 Optimal Mechanisms

In this section we characterize, step by step, a collection of simple mechanisms that spans the Pareto frontier of our trading problem. That is, for any pair of weights $\alpha = \left(\alpha^s, \alpha^b\right)$ (non-negative and sum to 1), we look for a mechanism that maximizes the $\alpha$-weighted sum of the agents' ex ante utilities:

$$\alpha^s E_{w^s,w^b} U^s(w^s, w^b) + \alpha^b E_{w^s,w^b} U^b(w^s, w^b) \tag{7}$$

in the domain of incentive compatible, individually rational and (ex ante) budget balanced mechanisms. We call such mechanisms "$\alpha$-optimal".

Recall from the example in section 2 that, whenever a first-best outcome is not attainable, the construction of a second-best mechanism involves finding the best compromise between generating trade surplus vs. increasing the rents to agents. Put differently, in these cases our aim is to restrict trade as efficiently as possible so as to reduce rents until the mechanism's surplus (which is also reduced by restricting trade) equals the sum of rents. Importantly, rents are paid to good types so as to deter them from pretending to be bad ones and obtain better trading prices. These rents are thus proportional to the probability that they will be still allowed to trade if they lie, i.e., on the probability that the bad types trade.

## 4.1 Good-Good and Bad-Bad Encounters

There are two cases in which there is no tension between trade surplus and rents. When two bad types meet, it is optimal not to let them trade, as this generates negative gains from trade and induces positive rents. When two good types meet, they must trade with probability 1, as this generates positive gains from trade without requiring any rents (as bad types have no reason to pretend to be good). This intuition, straightforward in the case of one object, remains true also with many objects, as the following proposition states:[8]

**Proposition 2** *In any $\alpha$-optimal mechanism, for every object $i \in I$:*

1. If both agents value $i$ as "good" – the object is traded with probability 1.

2. If both agents value $i$ as "bad" – the objects is not traded.

[8]Note that the proof of the proposition does not rely on the symmetry of the mechanism.

## 4.2 Good-Bad encounters and Minimal Rents

Knowing how good-good and bad-bad encounters should be dealt with, we now need to devise the most efficient way of limiting trade in good-bad encounters.

By Proposition 1 we can restrict our attention to symmetric mechanisms. Any symmetric mechanism induces, for each agent $j$, $2N$ constants $\left\{\xi_m^j\right\}_{m=1}^N$ and $\left\{\phi_m^j\right\}_{m=0}^{N-1}$, where $\xi_m^j$ and $\phi_m^j$ are the expected probabilities that a type $w^{j,m}$ of agent $j$ (a type with exactly $m$ good valuations) trades his good and bad objects, respectively:[9]

$$\begin{aligned}\xi_m^j &= \bar{p}^j\left(i,w^j\right) \text{ where } w_i^j = v_{good}^j \text{ and } g\left(w^j\right)=m,\ m=1...N\\ \phi_m^j &= \bar{p}^j\left(i,w^j\right) \text{ where } w_i^j = v_{bad}^j \text{ and } g\left(w^j\right)=m,\ m=0...N-1.\end{aligned}$$

These constants, accompanied by a set of transfers $\tau$, are sufficient to determine whether the mechanism satisfies IC, IR and BB. If these hold, then the transfers $\tau$ are the agents' utilities and they determine the value of the $\alpha$-weighted social welfare function.

The standard way of devising the transfers in unidimensional settings, in which types can be linearly ordered (See, e.g., Bolton and Dewatripont (2005), pp. 78-80), involves an inductive construction by which each type is paid the minimal rent that makes him not mimic the type that is just one step below.[10] That is, the mechanism pays the minimal rents subject to satisfying "local downward incentive compatibility" (LDIC). In the unidimensional case LDIC (plus the minimal rents) implies full incentive compatibility, since in any incentive-compatible mechanism the probabilities of trade must be weakly increasing in the agent's type[11].

In our multidimensional setting, this last implication is no longer the case. Incentive compatibility *alone* does not guarantee that the probabilities are increasing, that is $\phi_{m+1}^j \geq \phi_m^j$ and $\xi_{m+1}^j \geq \xi_m^j$[12]. Thus, local IC constraints do not ensure IC. We therefore employ the following proof strategy. We construct the optimal mechanism under the local IC constraint. We then observe that the *optimal* mechanism (under the local IC) is, in fact, monotone ($\phi_{m+1}^j \geq \phi_m^j$ and $\xi_{m+1}^j \geq \xi_m^j$ for

[9] The symmetry of the mechanism directly imply that the $\xi$'s and $\phi$'s are well defined.

[10] In continuous settings, such as Myerson and Satterthwaite's, the inductive construction is replaced by integration of a derivative.

[11] Provided that utilities are quasi-linear, as in our model.

[12] Incentive compatibility does imply that $\xi_{m+1}^j \geq \phi_m^j$ (which is the same kind of monotonicity as in the unidimensional case), but this does not help us.

all $m$). The monotonicity implies full incentive compatibility. This implies that the mechanisms that we construct, which are optimal in the larger set of local IC mechanisms, belong also to the smaller set of IC mechanisms. Thus, they are indeed optimal.

#### 4.2.1 Local Incentive Compatibility

We now define the set of SLIC mechanisms – symmetric and locally incentive compatible – in which we will search for optimal ones. We also define monotonicity and show that a monotone SLIC mechanism satisfies full IC.

**Definition 2** *A symmetric mechanism is SLIC if, for any $m \in \{0, ..., N-1\}$ and any $w^{j,m} \in G_m^j$ and $w^{j,m+1} \in G_{m+1}^j$*

1. $\Delta^j \xi_{m+1}^j \geq \bar{\tau}\left(w^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right)$
2. $\Delta^j \phi_m^j \leq \bar{\tau}\left(w^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right)$
3. $\xi_m^j \geq \phi_m^j$

The three conditions is are direct applications of Equation 4 (the IC condition). The first states that cheating upwards (reporting that one of the bad valuations is good) is not beneficial. The second states that pretending to have one less good valuation is not beneficial. The third states that within a given class $G_m^j$ the expected probability of trade of bad objects is less than that of good objects, which directly implies that an agent prefers to report his true type rather than pretending to be a different type in $G_m^j$ (i.e., cheat horizontally).

**Definition 3** *A symmetric mechanism is monotone if, for any $m \in \{1, ..., N-1\}$:*

1. $\phi_m^j \geq \phi_{m-1}^j$
2. $\xi_{m+1}^j \geq \xi_m^j$

**Proposition 3** *A monotone and SLIC mechanism is incentive compatible.*

**Naive utilities and revenue**

When we then look at SLIC mechanisms, utility and revenue calculations are misleading. A SLIC mechanism is not necessarily incentive compatible and truth-telling is, therefore, not necessarily an equilibrium. Thus, the equilibrium utilities of the agents may differ from their $\bar{\tau}$'s and the equilibrium revenue to the mechanism might not be $R$. We can however refer to the $\bar{\tau}$'s and $R$ as "naive" utilities and revenue: those that a naive mechanism designer, who trusts that the agents will report true valuations even in the absence of incentive compatibility, expects. We proceed by identifying the optimal mechanism from a naive designer point of view. Since, eventually, the naive-optimal mechanism turns out to satisfy IC, the $\bar{\tau}$'s and $R$ will turn to be the true utilities and revenue.

### 4.2.2 Tight Rents

Since in a SLIC mechanism we only need to prevent an agent from pretending to be of neighboring types, the minimal rents are obtained when each type in $G^j_m$ is paid to make him exactly indifferent between reporting the truth and pretending to be a type in $G^j_{m-1}$. This amounts to setting the difference between the payments $\bar{\tau}\left(w^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right)$ to be exactly $\Delta^j \phi^j_m$. Note that setting the $\tau$'s this way guaranties that the upward local IC condition, $\Delta^j \xi^j_{m+1} \geq \bar{\tau}\left(w^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right)$, is also satisfied, since $\xi^j_{m+1} \geq \phi^j_m$ in a SLIC mechanism.

**Definition 4** *A SLIC mechanism $\Gamma$ is tightly SLIC if, for any $m \in \{1, ..., N-1\}$ and any $w^{j,m} \in G^j_m$ and $w^{j,m+1} \in G^j_{m+1}$,*

$$\bar{\tau}\left(w^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right) = \Delta^j \phi^j_m \tag{8}$$

The following proposition states that we can pursue the optimal mechanisms within the set of tightly SLIC ones:

**Proposition 4** *For every SLIC mechanism $\Gamma = \left(p, \tau^s, \tau^b\right)$, there exists a tightly SLIC mechanism $\hat{\Gamma} = \left(p, \hat{\tau}^s, \hat{\tau}^b\right)$ that provides the same naive ex-ante utilities for the agents and same naive revenue for the mechanism.*

By proposition 2, in an optimal mechanism the probability of trade in objects for which both agents have bad valuation is 0. Thus, the expected probability that agent $j$ trades when he has a bad valuation, $\phi^j_m$, comes only from trade when opponent has a good valuation, an event with probability $q^{-j}$. Let then

$$\mu^j_m = \frac{\phi^j_m}{q^{-j}}$$

denote the expected probabilities that $j$ trades his bad objects (given that he has $m$ good valuations) conditional on $-j$ having a good valuation.

Proposition 4 lets us relate the payments that agents receive to their expected trade probabilities. We fix the payment to $w^{j,0}$, the worst-off type of agent $j$, at $c^j$. That is, independently of the rival's type $w^{-j}$, $\tau\left(w^{j,0}, w^{-j}\right) = c^j$. By (8), in an optimal mechanism, the expected payment to $j$'s other types are uniquely determined. For any $m = 0...N-1,\ j \in \{b, s\}$,

$$\bar{\tau}\left(w^{j,m}\right) = c^j + q^{-j}\Delta^j \sum\nolimits_{k=0}^{m-1} \mu_m^j. \tag{9}$$

We can now compute the ex ante naive utilities of the agents and the ex ante naive revenue of the mechanism. Recall that $j$'s valuations are i.i.d. across objects. Thus, the probability that $j$ is of type $w^{j,m} \in G_m^j$ is the corresponding density of the binomial distribution, $f_{BD}(k, N, q^j) = \binom{N}{m}\left(q^j\right)^m (1-q^j)^{N-m}$. The ex-ante naive utility of $j$ is, therefore:

$$\begin{aligned} NU^j\left(c^j, \mu_0^j...\mu_{N-1}^j\right) &= E_{w^j}\left[\sum\nolimits_{m=0}^{N} f_{BD}(m, N, q^j)\bar{\tau}\left(w^{j,m}\right)\right] \\ &= c^j + q^{-j}\Delta^j \sum\nolimits_{m=0}^{N}\left[f_{BD}(m, N, q^j)\sum\nolimits_{k=0}^{m-1} \mu_m^j\right] \end{aligned} \tag{10}$$

Define the naive income, $NI$, of the mechanism as the total surplus from trade. (The naive revenue is the naive income minus the payments $NU^s$ and $NU^b$ to the agents.) The surplus comes from good-good encounters (where trade occurs with probability 1), and from good-bad encounters (where the expected probability of trade is the appropriate $\mu_m^j$). The ex ante naive income is then:

$$\begin{aligned} NI\left(c^s, \mu_0^s...\mu_{N-1}^s, c^b, \mu_0^b...\mu_{N-1}^b\right) &= Nq^s q^b(v_h^b - v_l^s) \\ &\quad + q^s(v_l^b - v_l^s)\sum\nolimits_{m=0}^{N-1}\left[(N-m)\, f_{BD}(m, N, q^b)\mu_m^b\right] \\ &\quad + q^b(v_h^b - v_h^s)\sum\nolimits_{m=0}^{N-1}\left[(N-m)\, f_{BD}(m, N, q^s)\mu_m^s\right] \end{aligned} \tag{11}$$

Note that, by (10), the individual rationality constraint now reduces to requiring that $c^j \geq 0$. Ex ante budget balance requires that $NI \geq NU^s + NU^b$.

### 4.2.3 Optimal Surplus to Rents Trade-off

In a tightly SLIC mechanism, the expected payments to agents and the revenue to the mechanism are uniquely determined by the fixed payments to agents, $\left\{c^j\right\}$, and the set of relevant expected trade probabilities $\left\{\mu_m^j\right\}$. To devise the optimal limitations on trade, we need to compare the cost-to-benefit rations of the different variables. The cost, or shadow price, is the effect of increasing the variable on the budget-balance condition. The benefit is its effect on the agent's utility, weighted by the agent's $\alpha^j$. While the $c^j$ enter both the cost and the benefit side directly, the effect of the increasing the probabilities of trade $\mu_m^j$ is more subtle. They determine the agents rent, and thus enter both the benefit and the cost. But they also determine the surplus from trade, which is collected by the mechanism. This surplus relaxed the budget-balance condition. We will see that the cost-to-benefit ratio of $\mu_m^j$ decreases in $m$. In fact, it may even negative when $m$ is sufficiently large, as then the required rent can be lower than the surplus generated by the trade. For such $m$, there is no reason to restrict trade.

To prove our main result, we thus proceed to writing the linear maximization program that identifies the best way of restricting trade. Let $Z$ be the set of decision variables, that includes the constants $c^j$ and probabilities $\mu_m^j$:

$$Z = \left\{ c^s, \mu_0^s ... \mu_{N-1}^s, c^b, \mu_0^b ... \mu_{N-1}^b \right\}$$

Within the set of tightly SLIC mechanisms, we are looking, given weights $\left(\alpha^s, \alpha^b\right)$, for the one that maximizes the $\alpha$-weighted sum of naive utilities:

$$\begin{aligned} & MAX_Z \qquad \alpha^s NU^s + \alpha^b NU^b \qquad (12) \\ & \textit{subject to} \\ & \quad NU^s + NU^b - NI \leq 0 \qquad \text{(Budget Balance)} \\ & \quad c^s, c^b \geq 0 \qquad \text{(Individual Rationality)} \\ & \quad 1 \geq \mu_m^j \geq 0 \qquad \forall m \in \{0, ..., N-1\},\ j \in \{s, b\} \end{aligned}$$

For any decision variable $z$ in $Z$, let $r(z)$ be its cost-to-benefit ratio.[13] That is, $r(z)$ is the ratio of the derivative of the budget balance condition with respect to $z$ to the derivative of the objective function:[14]

$$r(z) = \frac{\partial \left(NI - NU^s - NU^b\right) / \partial z}{\partial \left(\alpha^s NU^s + \alpha^b NU^b\right) / \partial z}.$$

The following lemma states that the cost-to-benefit ratio of each agent's variables is monotone:

[13] We slightly abuse notation here, as $z$ in $r(z)$ refes to the name of the variable rather than to its value.

[14] Because our problem is linear, the cost-to-surplus ratio for each variable is a constant, i.e., $r(z)$ is a real number.

**Lemma 1** *For each agent $j$, the cost-to-benefit ratio of the trade probabilities $\mu_m^j$ is decreasing in $m$. Moreover, the cost-to-benefit ratio of $c^j$, the payment to $j$'s worst-off type $w^{j,0}$, is greater than that of all of his $\mu_m^j$.*

**Remark 1** *The key to the proof of this lemma is that for the binomial distribution (with p.d.f. $f_{BD}$ and c.d.f. $F_{BD}$), the function $\frac{(N-m)f_{BD}(m,N,q)}{1-F_{BD}(m,N,q)}$ is increasing in $m$. This function captures the three factors that affect the surplus-to-cost ratio of letting types with $m$ good valuations trade their bad objects. The numerator is the proportion of types who enjoy the increased probability of trade of bad objects, times the number bad objects per type. The denominator is the proportion of types higher than $m$ that must be paid higher rents when the $m$-types are allowed to trade their bad objects. While the denominator decreases with $m$ as needed, $f_{BD}$ increases first, but then decreases, and $(N-m)$ decreases all along. When these decrease, the numerator decreases – the opposite to the direction we need. Nonetheless, the lemma proves that the combination of the three factors is still increasing.*[15]

By standard linear programming arguments,[16] for any two decision variables $z_1$, $z_2$ with $r(z_2) < r(z_1)$, optimality requires that either $z_2$ attains its upper bound, or $z_1 = 0$. The optimal solution thus constitutes of ordering all the decision variables in $Z$ by their cost-to-benefit ratios, from the lowest to the to highest, and sequentially setting each to its maximal value, until the budget-balance condition is exhausted. The last variable that equates the mechanism's budget to 0 can either be one of the $\mu$'s, in which case this trade probability is set to some $\sigma \in (0,1)$ so as to exactly balance the budget; alternatively the process can terminate with one of the $c$'s, in which case that $c$ is set to positive amount that exhausts the budget (recall that the $c$'s are not bounded from above).[17] Finally, all the remaining variables are set to 0.[18]

**Remark 2** *This process, of gradually allowing trades in good-bad encounters, starts with a strictly positive revenue to the mechanism, which is the surplus from good-good encounters. Then, if there*

[15] While the lemma proves the monotonicity for our case of i.i.d. valuations, it is easy to see in the proof that there is some slack, as the monotonicity is strict. Thus, the monotonicity would still hold also for non independent valuations, as long as the correlation is not too high.

[16] See, e.g., the continuous knapsack problem in Dantzig (1963).

[17] For a finite number of values of $(\alpha^s, \alpha^b)$, there are two variables with the same $r(z)$, such that setting both to their maximal value would exceed the remaining budget. In these cases, any way of dividing the budget between the two variables (which must belong each to a different agent) is optimal.

[18] Unless the first-best outcome is attainable, the process ends before all the $\mu$'s are 1.

*are $\mu$'s with a negative cost-to-benefit ratio, they are set to* 1 *and relax even more the budget-balance constraint (such $\mu$'s exist if for one or both agents, for high enough $m$, the surplus generated from trading their bad objects exceeds the required rent to the types with more that $m$ good objects – see the exact condition in Proposition 5 below). As we continue to set more and more variables to their upper bound, the net income starts decreasing, until it is exhausted.*

As the cost-to-benefit ratios of the $\mu_m^j$'s and $c^j$ decrease with $\alpha^j$ (see the proof of Lemma 1), then as $\alpha^j$ increases, the $\mu_m^j$'s all decrease and the rival's $\mu_m^{-j}$'s all increase. As a result, the order of the variables in $Z$ can only change in one direction – some of the $\mu_m^j$'s may now outrank some of $-j$'s variables. Consequently, more of the $\mu_m^j$'s may "make it" before the budget is exhausted, implying that the threshold $m$ above which $j$ can trade his bad objects decreases.

To conclude the proof of our main result, note that any SLIC mechanism constructed by the inductive process above is monotone (the monotonicity of the $\mu$'s implies that $\phi_m^j \geq \phi_{m-1}^j$ and $\xi_{m+1}^j \geq \xi_m^j$ for all $m$). We thus apply Proposition 3 and conclude that it is, after all, incentive compatible. Thus, the naive utilities and naive income are, in fact, true utilities and income. We thus have:

**Theorem 1** *For any $\alpha^b + \alpha^s = 1$, there exists an $\alpha$-optimal mechanism, characterized by constants $M^s, M^b \in \{0, ..., N-1\}$ and $\sigma^s, \sigma^b \in [0, 1)$, in which, for any pair of announcements $w^s$, $w^b$, the probability of trade of object $i$ is:*

0 – *if both agents valuate $i$ as "bad"*

1 – *if both agents valuate $i$ as "good"*

$\mu_m^j$ – *if agent $j$ valuates $i$ as "bad" and agent $-j$ valuates $i$ as "good", where $m$ denotes the number of good valuations in $j$'s bid, and*

$$\mu_m^j = \begin{cases} 0 & \text{if} \quad m < M^j \\ \sigma_j & \quad\quad m = M^j \\ 1 & \quad\quad m > M^j \end{cases}$$

*Moreover, each $M^j$ weakly increases in $\alpha^j$.*

**Calculation of the transfers to agents:**

Recall that, by (9), the trade probabilities uniquely determine the interim transfers to agents. Substituting for the optimal trade probabilities above, and fixing an agent's ex post payment independently of his rival's type, we obtain the following simple transfer scheme:

$$\tau^j\left(w^j, w^{-j}\right) = \bar{\tau}^j\left(w^j\right) = Max\left\{0, \left(m - M_j - 1 + \sigma_j\right)\Delta^j q_{-j}\right\}.$$

Intuitively, an agent's is paid information rent for the number of good valuations in his announcement that exceeds his threshold $M_j$ (at the threshold itself the rent is $\sigma_j \Delta^j q_{-j}$, rather than $\Delta^j q_{-j}$ above the threshold, as its trade probability is only $\sigma_j$).

Note that we constructed the optimal mechanism under an ex ante budget balance constraint. To obtain ex post budget balanced, we can simply redefine the payments to be:

$$\hat{\tau}^j\left(w^j, w^{-j}\right) = \tau^j\left(w^j, w^{-j}\right) + \frac{1}{2}\left[R\left(w^j, w^{-j}\right) - E_{w^{-j}}\left[R\left(w^j, w^{-j}\right)\right] + E_{w^j}\left[R\left(w^j, w^{-j}\right)\right]\right],$$

where $R\left(w^j, w^{-j}\right)$ is the revenue of the mechanism, as defined in Equation 5.

**Construction of the Pareto frontier**

The set of utility pairs that can be achieved by IR, IC and BB mechanisms is convex.[19] To construct the Pareto Frontier of the trade problem recall that in the construction of the optimal trade probabilities, for all but a finite number of values of $\alpha$, exactly one variable exhausts the budget (it is one of the $\mu_m^j$'s or one of the $c^j$'s). This variable is the same for a range of $\alpha$'s, and corresponds to a vertex of the Pareto frontier.

For a finite number of $\alpha$'s, two variables exhausts the budget together, having exactly the same cost-to-benefit ratio. Such two variables must belong each to a different agent. By splitting the remaining budget between the two, the $\alpha$-weighted utility sum is unchanged, and we can pass utility between the agents in fixed ratio. In this way we obtain edges of the Pareto frontier. (One can easily check that any way of splitting the remaining budget between the agents is identical to a lottery between the mechanisms that yield the vertices of the corresponding edge.)

**Monopoly and consumer surplus:**

When $\alpha^s = 1$ and $\alpha^b = 0$, we can view the seller as a monopolist who sets the mechanism in his best interest. Sometimes the monopolist finds that selling each object at $v_h^b$ is optimal. In this case, there is no trade when the buyer has a bad valuation, and therefore he receives no rents. In other cases, the monopolist finds it optimal to sell some objects to buyers with low valuation, provided it

[19] This can be easily seen by noting that a simple lottery between any two such mechanisms also satisfies all the constraint, and yields each agent a weighted average of his utilities under the two mechanisms

has low cost for the object (the monopolist may be willing to sell all the objects to a low-valuation buyer, or prefer mixed bundling, by to selling some objects at the buyers low valuation only if sufficiently many objects are bought at the buyer's high valuation). In these cases, the buyer ends up with a positive (expected) consumer surplus.[20]

The following corollary states the necessary and sufficient condition for agent $j$ with $\alpha^j = 0$ to have a positive probability to trade his bad objects, or, equivalently, obtain a positive utility form $-j$'s preferred mechanism. The condition amount to verifying whether $j$'s types who are most likely to trade their bad objects, i.e. the type in $G^{N-1}$, have a negative cost-to-benefit ratio.

**Proposition 5** *In every second best mechanism, agent $j$ expects to receive a strictly positive ex ante utility iff:*

$$N \geq \frac{q^j}{1-q^j} \cdot \frac{\Delta^j}{v_h^b - v_l^s - \Delta^j}$$

# 5 Numerical Simulations

In this section we present numerical simulations that provide some insights for the properties of $a$-second best mechanisms. The following figure shows a Pareto-frontier of a trading problem with eight objects:

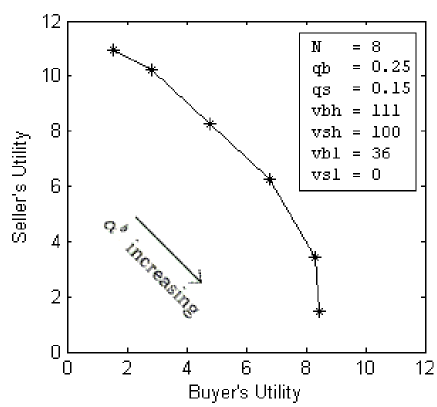


Figure 1

Observe that the seller's utility is maximal at $\alpha^s = 1$ and decreases along the curve, as the designer leans more towards the the buyer. Observe also that when the utility of one agent is maximal (that is $\alpha^j = 1$), the utility of its rival is non-zero. This is a direct implication of

[20]Note that our problem is different than the standard monopolistic bundling model, as the monopoly's type is not common knowledge. Instead, it devises its optimal mechanism at the ex ante stage, and then both the buyer and the monopoly submit their reports to the mechanism. In other words, the monopoly here is an "informed principal" (see Myerson 1983 and Maskin and Tirole 1990).

Proposition 5. It demonstrates that when there are sufficiently many objects, some types of each agent are allowed to trade their bad objects whatever the designer's bias is; the agents thus obtain a positive ex ante utility.

Figure (2) shows how the per-object utilities change as the number of objects increases:

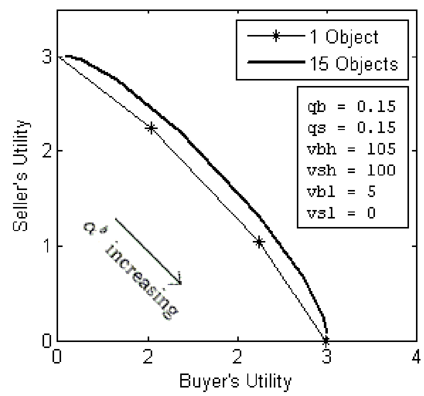


Figure 2

The thin curve is the Pareto frontier in the case of one object. It is also the one achieved with many objects if the mechanism is just a replication of the one-object optimal mechanism and does not exploit the possibility to link the trade problems. With fifteen objects, our mechanism increases both agents' utilities considerably if $\alpha^b$ and $\alpha^s$ are close to 1/2. But as the $\alpha$'s become far apart, the gain from linking the trade problems is reduced. In the extreme (the monopoly or monopsonist cases of $\alpha^s = 1$ or $\alpha^b = 1$), the two curves meet, implying that linking the trading problems does not help. This conclusion holds here because, given the problem's parameters, the number of objects is not large enough and the monopolist sell only to good types, leaving no rent for the adversary (see Proposition 5). Note, however, that for sufficiently many objects, an agent's utility when his $\alpha$ is 0 is no longer 0 (as in the first figure). In such cases the two Pareto frontiers do not meet at the extremes, implying that linking the trade problems is helpful.

Figure 3 show how the trade limitations for the two agents change together as $\alpha$ change. It shows the different combinations of thresholds, such that each agent is allowed to trade his bad objects only when the number of good valuations in his announcement is above his threshold. In this example with eight objects, if agents are treated equally, both thresholds are 1. But if the mechanism leans sufficiently towards one of them, his threshold becomes 0, while the threshold of the less favored agent increases gradually as his $\alpha$ decreases (but never reaches 8).

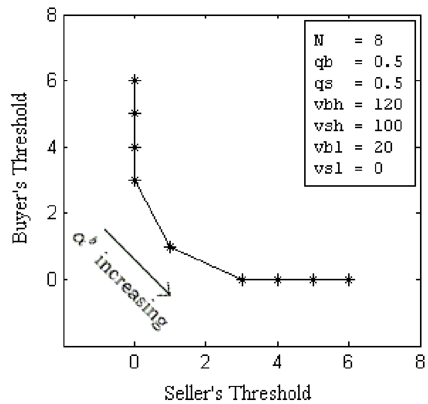


Figure 3

Finally, Figures 4 and 5 show the Pareto frontier and trade limitations when the valuations are such that the first-best surplus is attainable. Figure 4 shows that at $\alpha^s = \alpha^b$ there is no gain in linking the trade problems (as the one-object mechanism is already optimal). As we move away from equal treatment of the agents, the Pareto frontier is expanded when the optimal multiple-object mechanism is employed. Figure 5 shows that, as we shift away from equal treatment (were no trade limitations are imposed on any player), only the agent with lesser weight is restricted from trading (and the saving in his information rent, minus the forgone surplus, goes to the opponent and is added to his fixes payment $c$).

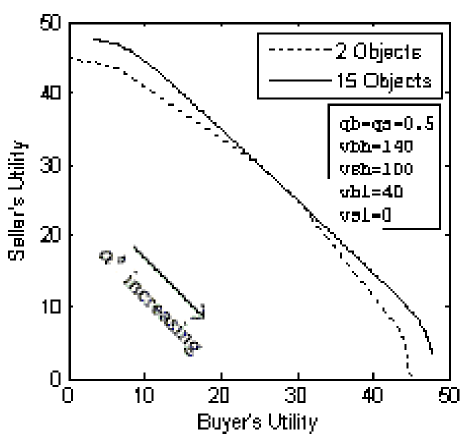


Figure 4

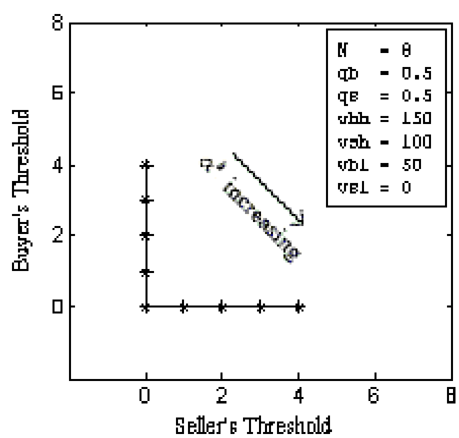


Figure 5

## References


[1] Athey, S. and D. Miller (2007), Efficiency in repeated trade with hidden valuations, *Theoretical Economics* 2, 299–354

[2] Bolton, P. and M. Dewatripont (2005) *Contract Theory*, MIT Press, Cambridge, MA, USA.

[3] Borgers, T. and P. Norman (2008), A Note on Budget Balance under Interim Participation Constraints: The Case of Independent Types *Economic Theory* 34,

[4] Chechile, R. A. (2003), Mathematical tools for hazard function analysis, *Journal of Mathematical Psychology* 47, 478-494.

[5] Dantzig, G. B. (1963), *Linear Programming and Extensions*, Princeton University Press, Princeton, NJ, USA.

[6] Jackson, M. O. and H. O. Sonnenschein (2007), Overcoming Incentive Constraints by Linking Decisions *Econometrica* 75, 241-258.

[7] E. Maskin and J. Tirole (1992) The Principal-Agent Relationship with an Informed Principal: The Case of Private Values *Econometrica* 58, 379-409.

[8] Myerson, R. B. (1979), Incentive compatibility and the bargaining problem, *Econometrica* 47, 61-74.

[9] Myerson, R. B. (1983), Mechanism Design by an Informed Principal, *Econometrica* 51, 1767-1798.

[10] Myerson, R. B. and M. A. Satterthwaite (1983), Efficient Mechanisms for Bilateral Trading, *J. Economic Theory* 28, 265–281.

[11] Fang, H. and P. Norman (2006), Optimal Provision of Multiple Excludable Public Goods, working paper.

[12] Fang, H. and P. Norman (2008), Optimal Provision of Multiple Excludable Public Goods, working paper.

Appendix: Proofs

**Proposition 1**[21]

Assume that $\Gamma \equiv \langle p, \tau^s, \tau^b \rangle$ is an incentive compatible mechanism. Let $\Gamma^\pi \equiv \langle p_\pi, \tau^s_\pi, \tau^b_\pi \rangle$ denote a new mechanism that is defined for every $w^j \in W^b$, $j \in \{b, s\}$ and $\pi \in \Pi$ as follows:

$$\begin{aligned} \tau^j_\pi \left( w^s, w^b \right) &= \tau^j \left( M_\pi w^s, M_\pi w^b \right) \\ p_\pi \left( i; w^s, w^b \right) &= p \left( \pi(i), M_\pi w^s, M_\pi w^b \right) \end{aligned}$$

**Lemma 2** *For every $j \in \{b, s\}$ and every arbitrary function $h : W^j \to \mathbb{R}$:*

$$E_{w^j} \left[ h \left( w^j \right) \right] = E_{w^j} \left[ h \left( M_\pi w^j \right) \right]$$

**Proof.** The expected value of $h\left(w^j\right)$ is given by:

$$\begin{aligned} E_{w^j} \left[ h \left( w^j \right) \right] &= \sum\nolimits_{w^j} \left[ prob(w^j) \cdot h \left( w^j \right) \right] \\ &= \sum\nolimits_{w^j} \left[ prob(M_\pi w^j) \cdot h \left( M_\pi w^j \right) \right]. \end{aligned}$$

The elements of $W^j$ are i.i.d. random variables, hence $pr(w^j) = pr(M_\pi w^j)$. This therefore equals:

$$\sum\nolimits_{w^j} \left[ prob(w^j) \cdot h \left( M_\pi w^j \right) \right] = E_{w^j} \left[ h \left( M_\pi w^j \right) \right]$$

■

**Lemma 3** *if $\Gamma$ is incentive-compatible then $\Gamma^\pi$ is incentive-compatible*

**Proof.** Consider an agent $j \in \{b, s\}$ of an arbitrary type $w^j \in W^j$. Since $\Gamma$ is incentive-compatible, then truth-telling is always at least as profitable to $j$ as any other announcement $\hat{w}^j$. That is,

$$E_{w^{-j}} \left[ \sum_{i=1}^{N} p \left( i, \hat{w}^j, w^{-j} \right) \left( w^j_i - \hat{w}^j_i \right) + \tau^j \left( \hat{w}^j, w^{-j} \right) \right] \leq E_{w^{-j}} \left[ \tau^j \left( w^j, w^{-j} \right) \right]$$

for every $w^j, \hat{w}^j \in W^j$. In particular, this holds for $\left( M_\pi w^j \right) \in W^j$ and $\left( M_\pi \hat{w}^j \right) \in W^j$:

$$E_{w^{-j}} \left[ \sum_{i=1}^{N} p \left( i, M_\pi \hat{w}^j, w^{-j} \right) \left( \left( M_\pi w^j \right)_i - \left( M_\pi \hat{w}^j \right)_i \right) + \tau^j \left( M_\pi \hat{w}^j, w^{-j} \right) \right] \leq E_{w^{-j}} \left[ \tau^j \left( M_\pi w^j, w^{-j} \right) \right]$$

[21] The proof of this proposition is inspired by Norman and Fang (2006).

By lemma (2), this implies:

$$
\begin{aligned}
&E_{w^{-j}}\left[\sum_{i=1}^{N} p\left(i, M_\pi \hat{w}^j, M_\pi w^{-j}\right)\left(\left(M_\pi w^j\right)_i - \left(M_\pi \hat{w}^j\right)_i\right) + \tau^j\left(M_\pi \hat{w}^j, M_\pi w^{-j}\right)\right] \\
\leq\ & E_{w^{-j}}\left[\tau^j\left(M_\pi w^j, M_\pi w^{-j}\right)\right]
\end{aligned}
$$

Note that by definition $\left(M_\pi w^j\right)_i = w^j_{\pi^{-1}(i)}$ and $\left(M_\pi \hat{w}^j\right)_i = \hat{w}^j_{\pi^{-1}(i)}$. Thus,

$$
\begin{aligned}
&E_{w^{-j}}\left[\sum_{i=1}^{N} p\left(i, M_\pi \hat{w}^j, M_\pi w^{-j}\right)\left(w^j_{\pi^{-1}(i)} - \hat{w}^j_{\pi^{-1}(i)}\right) + \tau^j\left(M_\pi \hat{w}^j, M_\pi w^{-j}\right)\right] \\
\leq\ & E_{w^{-j}}\left[\tau^j\left(M_\pi w^j, M_\pi w^{-j}\right)\right]
\end{aligned}
$$

or equivalently, by changing the order of summation:

$$
\begin{aligned}
&E_{w^{-j}}\left[\sum_{i=1}^{N} p\left(\pi\left(i\right), M_\pi \hat{w}^j, M_\pi w^{-j}\right)\left(w^j_i - \hat{w}^j_i\right) + \tau^j\left(M_\pi \hat{w}^j, M_\pi w^{-j}\right)\right] \\
\leq\ & E_{w^{-j}}\left[\tau^j\left(M_\pi w^j, M_\pi w^{-j}\right)\right]
\end{aligned}
$$

using the definitions of $p_\pi$ and $\tau^b_\pi$, this is:

$$
\begin{aligned}
&E_{w^{-j}}\left[\sum_{i=1}^{N} p_\pi\left(i, \hat{w}^j, w^{-j}\right)\left(w^j_i - \hat{w}^j_i\right) + \tau^j_\pi\left(\hat{w}^j, w^{-j}\right)\right] \\
\leq\ & E_{w^{-j}}\left[\tau^j_\pi\left(w^j, w^{-j}\right)\right]
\end{aligned}
$$

Which implies that $\Gamma_\pi$ is Incentive Compatible. ■

**Lemma 4** *If* $\Gamma = \left\langle p, \tau^s, \tau^b\right\rangle$ *and* $\Gamma_\pi = \left\langle p_\pi, \tau^s_\pi, \tau^b_\pi\right\rangle$ *are incentive compatible, then both yield the same ex-ante utility for the agents.*

**Proof.** The ex ante utility of agent $j$ from bidding truthfully under $\Gamma$ is:

$$
E_{w^j} E_{w^{-j}}\left[\tau^j\left(w^j, w^{-j}\right)\right].
$$

using lemma (2) twice, this equals:

$$
E_{w^j} E_{w^{-j}}\left[\tau^j\left(M_\pi w^j, M_\pi w^{-j}\right)\right]
$$

and using the definition of $\tau_{\pi}^{j}$, this is:

$$E_{w^j} E_{w^{-j}} \left[\tau_{\pi}^{j}\left(w^j, w^{-j}\right)\right],$$

which is the ex-ante utility of $j$ under $\Gamma_\pi$. ■

**Lemma 5** *If $\Gamma = \left\langle p, \tau^s, \tau^b \right\rangle$ and $\Gamma_\pi = \left\langle p_\pi, \tau_\pi^s, \tau_\pi^b \right\rangle$ are incentive compatible, then both yield the same revenue to the mechanism.*

**Proof.** Let $ER$ and $ER_\pi$ denote the expected revenue of the mechanism under $\Gamma$ and $\Gamma_\pi$, respectively. Then:

$$ER = E_{w^j} E_{w^{-j}} \left[ \begin{array}{c} \sum_{i=1}^{N} \left[p\left(i, w^s, w^b\right) \times \left(w_i^b - w_i^s\right)\right] - \\ \tau^b\left(w^s, w^b\right) - \tau^s\left(w^s, w^b\right) \end{array} \right].$$

Using lemma (2) twice, this equals:

$$E_{w^j} E_{w^{-j}} \left[ \begin{array}{c} \sum_{i=1}^{N} p\left(i, M_\pi w^s, M_\pi w^b\right) \left(\left(M_\pi w^b\right)_i - \left(M_\pi w^s\right)_i\right) - \\ \tau^b\left(M_\pi w^s, M_\pi w^b\right) - \tau^s\left(M_\pi w^s, M_\pi w^b\right) \end{array} \right]$$

By definition $\left(M_\pi w^j\right)_i = w_{\pi^{-1}(i)}^j$, that is:

$$E_{w^j} E_{w^{-j}} \left[ \begin{array}{c} \sum_{i=1}^{N} p\left(i, M_\pi w^s, M_\pi w^b\right) \left(w_{\pi^{-1}(i)}^b - w_{\pi^{-1}(i)}^s\right) - \\ \tau^b\left(M_\pi w^s, M_\pi w^b\right) - \tau^s\left(M_\pi w^s, M_\pi w^b\right) \end{array} \right],$$

or equivalently, by changing the order of summation:

$$E_{w^j} E_{w^{-j}} \left[ \begin{array}{c} \sum_{i=1}^{N} p\left(\pi\left(i\right), M_\pi w^s, M_\pi w^b\right) \left(w_i^b - w_i^s\right) - \\ \tau^b\left(M_\pi w^s, M_\pi w^b\right) - \tau^s\left(M_\pi w^s, M_\pi w^b\right) \end{array} \right]$$

and by definition of $p_\pi, \tau_\pi^b$ and $\tau_\pi^s$, this is

$$E_{w^j} E_{w^{-j}} \left[ \begin{array}{c} \sum_{i=1}^{N} \delta^\pi\left(i, w^s, w^b\right) \left(w_i^b - w_i^s\right) - \\ \tau_\pi^b\left(w^s, w^b\right) - \tau_\pi^s\left(w^s, w^b\right) \end{array} \right] = ER_\pi$$

■

We use these lemmas to specify the symmetric mechanism. Given $\Gamma = \left\langle p, \tau^s, \tau^b \right\rangle$, let us define a new mechanism denoted by $\hat{\Gamma}$. In $\hat{\Gamma}$ the agents are first asked to submit their vectors of valuations:

$w^b \in W^b, w^s \in W^b$. Then, the mechanism *randomly* selects a permutation $\pi$ from $\Pi$, and implements the trade probabilities and transfers implied by $\Gamma_\pi$. $\hat{\Gamma}$ is symmetric since for every object $i \in I$, the probability for trade depends only upon the agent's valuation of $i$ and the *distribution* of valuations of all objects other than $i$.

According to lemma (3), both agents would reveal their true types in $\Gamma_\pi$ for every possible $\pi \in \Pi$. They would therefore also reveal their true types under $\hat{\Gamma}$, i.e. before they know the result of the randomization. This implies that $\hat{\Gamma}$ is incentive compatible. Furthermore, lemma (4) and lemma (5) imply that the agents have the same ex-ante utilities under $\hat{\Gamma}$ and $\Gamma$, and the bank has the same budget balance.

**Proposition 2**

We show that if any of the properties does not hold for an incentive compatible mechanism $\Gamma = \left\langle p, \tau^s, \tau^b \right\rangle$ then there exists another incentive compatible mechanism $\tilde{\Gamma}$ for which the ex-ante utility of (at least) one of the agents is higher, and the ex-ante budget balance does not change.

**Good-Good trades:** Suppose that $\Gamma = \left\langle p, \tau^s, \tau^b \right\rangle$ is an incentive compatible mechanism in which objects that are valuated as "good" by both agents are not always traded with probability 1. That is, there exist two vectors $\tilde{w}^j \in W^j$ $(j \in \{b, s\})$ and an object $k \in I$ for which $\tilde{w}_k^j = v_{good}^j$ with $p\left(k, \tilde{w}^s, \tilde{w}^b\right) < 1$.

Let us consider an alternative mechanism $\tilde{\Gamma} = \left\langle \tilde{p}, \tau^s, \tau^b \right\rangle$ with $\tilde{p} = p$ except $\tilde{p}\left(k, \tilde{w}^s, \tilde{w}^b\right) = 1$. Since this change affects only agents that announce "good" in the $k^{th}$ object, it does not violate incentive compatibility (bad types only lose more by pretending to be "good"). It does, however, increase the expected surplus of the mechanism by

$$\left[1 - p(k, \hat{w}^s, \hat{w}^b)\right] \cdot prob(\tilde{w}^b) \cdot prob(\tilde{w}^s) \cdot \left(v_{good}^b - v_{good}^s\right).$$

This quantity is strictly positive and can be transferred as a lump sum to the agents in order to increase their ex-ante utility. Doing so keeps the mechanism budget balanced, while not affecting the incentives of the agents.

**Bad-Bad trades:** Suppose that $\Gamma = \left\langle p, \tau^s, \tau^b \right\rangle$ is an incentive compatible mechanism in which objects that are valuated as "bad" by both agents have a positive probability of being traded. That is, there exist two vectors $\tilde{w}^j \in W^j$ $(j \in \{b, s\})$ and an object $k \in I$ for which $\tilde{w}_k^j = v_{bad}^j$ with $p\left(k, \tilde{w}^s, \tilde{w}^b\right) > 0$.

Consider an alternative mechanism $\tilde{\Gamma} = \left\langle \tilde{p}, \tau^s, \tau^b \right\rangle$ with $\tilde{p} = p$, except that $\tilde{p}\left(k, \tilde{w}^s, \tilde{w}^b\right) = 0$.

Again, this change does not violate incentive-compatibility of the agents under $\tilde{\Gamma}$, as good types lose now more if they pretend to be bad. However, by decreasing the probability that objects that are valuated as "bad" by both agents are traded, the expected surplus is increased by:

$$p\left(k, \tilde{w}^s, \tilde{w}^b\right) \times pr(\tilde{w}^b) \times pr(\tilde{w}^s) \times \left(v^s_{bad} - v^b_{bad}\right)$$

which is strictly positive, and can be transferred to the agents as a lump sum in order to increase their ex-ante utility.

**Proposition 3**

Suppose that $\Gamma$ is a monotone and SLIC mechanism. By definition (2) we know that for any $m \in \{1, ..., N-1\}$ and any $\hat{w}^{j,m+1} \in G^j_{m+1}$ and $w^{j,m} \in G^j_m$ the following inequalities hold:

$$\begin{aligned} \Delta^j \phi^j_m &\leq \bar{\tau}\left(\hat{w}^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right) \\ \Delta^j \xi^j_{m+1} &\geq \bar{\tau}\left(\hat{w}^{j,m+1}\right) - \bar{\tau}\left(w^{j,m}\right) \end{aligned}$$

Using the inequalities recursively it is straightforward to show that for any two vectors $\hat{w}^{j,m+k} \in G^j_{m+k}$ and $w^{j,m} \in G^j_m$ where $1 \leq k \leq (N-m)$:

$$\begin{aligned} \Delta^j \sum_{l=0}^{k-1} \phi^j_{m+l} &\leq \bar{\tau}\left(\hat{w}^{j,m+k}\right) - \bar{\tau}\left(w^{j,m}\right) \\ \Delta^j \sum_{l=0}^{k-1} \xi^j_{m+l+1} &\geq \bar{\tau}\left(\hat{w}^{j,m+k}\right) - \bar{\tau}\left(w^{j,m}\right) \end{aligned}$$

and by monotonicity:

$$\Delta^j \cdot k \cdot \phi^j_m \leq \bar{\tau}\left(\hat{w}^{j,m+k}\right) - \bar{\tau}\left(w^{j,m}\right), \tag{13}$$

$$\Delta^j \cdot k \cdot \xi^j_{m+k} \geq \bar{\tau}\left(\hat{w}^{j,m+k}\right) - \bar{\tau}\left(w^{j,m}\right). \tag{14}$$

Let the *distance* between $w^{j,m}$ and $\hat{w}^{j,m+k}$, $\left\|w^{j,m}, \hat{w}^{j,m+k}\right\|$, to be the number of elements by which the two vectors differ. Observe that:

$$\left\|w^{j,m}, \hat{w}^{j,m+k}\right\| \geq k. \tag{15}$$

Note also that there are exactly $\frac{1}{2}\left(\left\|w^{j,m}, \hat{w}^{j,m+k}\right\| + k\right)$ elements that $\hat{w}^{j,m+k}$ valuates as *good* and $w^{j,m}$ valuates as *bad*. Similarly, there are exactly $\frac{1}{2}\left(\left\|w^{j,m}, \hat{w}^{j,m+k}\right\| - k\right)$ elements that $\hat{w}^{j,m+k}$ valuates as *bad* and $w^{j,m}$ valuates as *good*. Using the properties of SLIC and equation (15) we know that:

$$\begin{aligned}\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\left(\phi_m^j-\xi_m^j\right)\Delta^j &\leq 0,\\ \frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\left(\xi_{m+k}^j-\phi_{m+k}^j\right)\Delta^j &\geq 0.\end{aligned}$$

Add both sides of the upper equations to (13) and both sides of the lower to (14) to get:

$$\begin{aligned}\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\left(\phi_m^j-\xi_m^j\right)\Delta^j+\Delta^j\cdot k\cdot\phi_m^j &\leq \bar{\tau}\left(\hat{w}^{j,m+k}\right)-\bar{\tau}\left(w^{j,m}\right),\\ \frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\left(\xi_{m+k}^j-\phi_{m+k}^j\right)\Delta^j+\Delta^j\cdot k\cdot\xi_{m+k}^j &\geq \bar{\tau}\left(\hat{w}^{j,m+k}\right)-\bar{\tau}\left(w^{j,m}\right)\end{aligned}$$

or equivalently:

$$\bar{\tau}\left(\hat{w}^{j,m+k}\right)\geq\bar{\tau}\left(w^{j,m}\right)+\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|+k\right)\Delta^j\phi_m^j-\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\Delta^j\xi_m^j \tag{16}$$

$$\bar{\tau}\left(w^{j,m}\right)\geq\bar{\tau}\left(\hat{w}^{j,m+k}\right)-\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|+k\right)\Delta^j\xi_{m+k}^j+\frac{1}{2}\left(\left\|w^{j,m},\hat{w}^{j,m+k}\right\|-k\right)\Delta^j\phi_{m+k}^j \tag{17}$$

For every non-negative $k$, equation (16) implies that $\hat{w}^{j,m+k}$ does not gain from pretending to be $w^{j,m}$, and equation (17) implies that $w^{j,m}$ does not gain from pretending to be $\hat{w}^{j,m+k}$. This means that the mechanism is incentive compatible.

**Proposition 4**

Let $w^{j,m}\in G_m^j$ and $w^{j,m+1}\in G_{m+1}^j$ denote two types of agent $j\in\{b,s\}$ and let $w^{-j}\in W^{-j}$ denote an arbitrary type of agent $-j$. SLIC implies that:

$$\Delta^j\phi_m^j\leq\bar{\tau}^j\left(w^{j,m+1}\right)-\bar{\tau}^j\left(w^{j,m}\right)\leq\Delta^j\xi_{m+1}^j \quad \text{for} \quad m\in\{0,...,N-1\} \tag{18}$$

Note that (18) consists of $N$ equations with two inequalities each. If $\Gamma$ is tightly SLIC then we are done. Otherwise, let $k$ denote the index of the first equation for which the left inequality is strict. That is:

$$\begin{aligned}&\Delta^j\phi_m^j=\bar{\tau}^j\left(w^{j,m+1}\right)-\bar{\tau}^j\left(w^{j,m}\right)\leq\Delta^j\xi_{m+1}^j \quad \text{for } m\in\{0,...,k-1\}\\ &\Delta^j\phi_m^j<\bar{\tau}^j\left(w^{j,m+1}\right)-\bar{\tau}^j\left(w^{j,m}\right)\leq\Delta^j\xi_{m+1}^j \quad \text{for } m=k\\ &\Delta^j\phi_m^j\leq\bar{\tau}^j\left(w^{j,m+1}\right)-\bar{\tau}^j\left(w^{j,m}\right)\leq\Delta^j\xi_{m+1}^j \quad \text{for } m\in\{k+1,...,N-1\}\end{aligned} \tag{19}$$

Let $\hat{\Gamma} = \left(p, \hat{\tau}^s, \hat{\tau}^b\right)$ denote a mechanism with the same trade probabilities as $\Gamma$ but with transfers:

$$\hat{\tau}^j \left(w^{j,m}, w^{-j}\right) = \begin{cases} \tau\left(w^{j,m}, w^{-j}\right) + d\left[1 - F_{BD}\left(k, N, q^j\right)\right] & \text{for } m \in \{0, ..., k\} \\ \tau\left(w^{j,m}, w^{-j}\right) + d\left[1 - F_{BD}\left(k, N, q^j\right)\right] - d & \text{for } m \in \{k+1, ..., N-1\} \end{cases} \tag{20}$$

where $d$ is the following constant:

$$d \equiv E_{w^{-j}}\left[\tau\left(w^{j,k+1}, w^{-j}\right) - \tau\left(w^{j,k}, w^{-j}\right)\right] - \Delta^j \phi_k^j$$

Thus, as compared to $\Gamma$, under $\hat{\Gamma}$ all types of agent $j$ receive a constant addition of size $d\left[1 - F_{BD}\left(k, N, q^j\right)\right]$ to their utility. However, types with *more* than $k$ *"good"* valuations also lose a constant of size $d$. Since in ex-ante terms the probability to have more than $k$ *"good"* valuations is exactly $\left[1 - F_{BD}\left(k, N, q^j\right)\right]$, then the ex-ante utility is the same under $\Gamma$ and $\hat{\Gamma}$, and the revenue of the mechanism is the same under $\Gamma$ and $\hat{\Gamma}$.

Observe also that under $\hat{\Gamma}$ the difference between the transfers to $w^{j,m+1}$ and $w^{j,m}$ is given by

$$\begin{aligned} & E_{w^{-j}}\left[\hat{\tau}^j\left(w^{j,m+1}, w^{-j}\right) - \hat{\tau}^j\left(w^{j,m}, w^{-j}\right)\right] \\ = & \begin{cases} E_{w^{-j}}\left[\tau^j\left(w^{j,m+1}, w^{-j}\right) - \tau^j\left(w^{j,m}, w^{-j}\right)\right] & m \neq k \\ \Delta^j \phi_k^j & m = k \end{cases} \end{aligned}$$

and thus:

$$\begin{aligned} \Delta^j \phi_m^j = \bar{\tau}^j\left(w^{j,m+1}\right) - \bar{\tau}^j\left(w^{j,m}\right) \leq \Delta^j \xi_{m+1}^j \quad & \text{for } m \in \{0, ..., k\} \\ \Delta^j \phi_m^j \leq \bar{\tau}^j\left(w^{j,m+1}\right) - \bar{\tau}^j\left(w^{j,m}\right) \leq \Delta^j \xi_{m+1}^j \quad & \text{for } m \in \{k+1, ..., N-1\} \end{aligned}$$

Therefore, $\hat{\Gamma}$ preserves local incentive compatibility and has exactly one more binding constraint compared to $\Gamma$. If $\hat{\Gamma}$ is tightly SLIC - we are done, otherwise - repeat the process of eliminating a non-binding constraint until the mechanism is tightly SLIC.

**Lemma 1**

Let $a_m^j$ denote the (marginal) effect of increasing $\mu_m^j$ on the objective function:

$$\begin{aligned} a_m^j \equiv \partial\left(\alpha^s N U^s + \alpha^b N U^b\right)/\partial \mu_m^j &= \alpha^j \cdot q^{-j} \cdot \Delta^j \cdot \sum\nolimits_{k=m+1}^{N} \left[\binom{N}{k} \left(q^j\right)^k \left(1 - q^j\right)^k\right] \\ &= \alpha^j \cdot q^{-j} \cdot \Delta^j \cdot \left(1 - F_{BD}\left(m, N, q^j\right)\right) \end{aligned}$$

and let $b_m^j$ denote the (marginal) effect of increasing $\mu_m^j$ on the budget constraint:

$$\begin{aligned} b_m^b &= \frac{\partial\left[NU^b(\cdot)+NU^s(\cdot)-NI(\cdot)\right]}{\partial\mu_m^b} = \left(\frac{a_m^b}{\alpha^b}\right) - q^s(v_l^b - v_l^s)\times(N-m)\times f_{BD}(m,N,q^b) \\ b_m^s &= \frac{\partial\left[NU^b(\cdot)+NU^s(\cdot)-NI(\cdot)\right]}{\partial\mu_m^s} = \left(\frac{a_m^s}{\alpha^s}\right) - q^b(v_h^b - v_h^s)\times(N-m)\times f_{BD}(m,N,q^s) \end{aligned}$$

Thus, the cost-to-benefit rate of $\mu_m^s$ and $\mu_m^b$ is explicitly given by:

$$\begin{aligned} \frac{b_m^b}{a_m^b} &= \frac{1}{\alpha^b}\left[1-\frac{(v_l^b-v_l^s)}{(v_h^b-v_l^b)}\times\frac{(N-m)\times f_{BD}(m,N,q^b)}{1-F_{BD}(m,N,q^b)}\right] \\ \frac{b_m^s}{a_m^s} &= \frac{1}{\alpha^s}\left[1-\frac{(v_h^b-v_h^s)}{(v_h^s-v_l^s)}\times\frac{(N-m)\times f_{BD}(m,N,q^s)}{1-F_{BD}(m,N,q^s)}\right] \end{aligned}$$

Note that for $c^j$ the cost-to-benefit rate is $\frac{1}{\alpha^j}$, which is greater than $\frac{b_m^j}{a_m^j}$ for all $m\in\{0,...,N-1\}$. To show that the ratios $\frac{b_m^j}{a_m^j}$ decrease in $m$, it is sufficient to prove the following claim:

**Claim 1** $(N-m)\frac{[f_{BD}(m,N,q)]}{[1-F_{BD}(m,N,q)]}$ *is strictly increasing in $m$ for every $q\in[0,1]$*

**Proof.** Let $p_m = f_{BD}(m,N,q) = \binom{N}{m}q^m(1-q)^{N-m}$. Following Chechile (2003)[4], define $W \equiv \frac{p_{k+j}p_{\hat{k}}}{p_{\hat{k}+j}p_k}$ for every and $\hat{k},k,j\in\{0,...,N\}$ and observe that:

$$W \equiv \frac{p_{k+j}p_{\hat{k}}}{p_{\hat{k}+j}p_k} = \frac{(\hat{k}+j)...(\hat{k}+1)(n-k)...(n-k-j+1)}{(k+j)...(k+1)(n-\hat{k})...(n-\hat{k}-j+1)}.$$

Multiply both sides by $\frac{(n-\hat{k})}{(n-k)}$ to get:

$$\frac{p_{k+j}p_{\hat{k}}}{p_{\hat{k}+j}p_k}\frac{(n-\hat{k})}{(n-k)} = \frac{(\hat{k}+j)...(\hat{k}+1)(n-k-1)...(n-k-j+1)}{(k+j)...(k+1)(n-\hat{k}-1)...(n-\hat{k}-j+1)}. \tag{21}$$

Clearly, if $\hat{k}>k$ then each of the terms in the numerator of the RHS of (21) is larger than the corresponding term in the denominator; thus $\frac{p_{k+j}p_{\hat{k}}}{p_{\hat{k}+j}p_k}\frac{(n-\hat{k})}{(n-k)}>1$ or, equivalently,

$$\frac{1}{(n-k)}\frac{p_{k+j}}{p_k} > \frac{p_{\hat{k}+j}}{p_{\hat{k}}}\frac{1}{\left(n-\hat{k}\right)}.$$

Thus,

$$\frac{1}{(n-k)}\frac{\sum_{i=k+1}^{i=N-\left(\hat{k}-k\right)}p_i}{p_k} > \frac{\sum_{i=\hat{k}+1}^{i=N}p_i}{p_{\hat{k}}}\frac{1}{\left(n-\hat{k}\right)} \qquad \forall\hat{k}>k,$$

and therefore also:

$$\frac{1}{(n-k)}\frac{\sum_{i=k+1}^{i=N} p_i}{p_k} > \frac{\sum_{i=\hat{k}+1}^{i=N} p_i}{p_{\hat{k}}}\frac{1}{\left(n-\hat{k}\right)} \qquad \forall \hat{k} > k,$$

and since $\sum_{i=0}^{i=N} p_i = 1$ then

$$\frac{1}{(n-k)}\frac{1-\sum_{i=0}^{i=k} p_i}{p_k} > \frac{\sum_{i=0}^{i=\hat{k}} p_i}{p_{\hat{k}}}\frac{1}{\left(n-\hat{k}\right)}.$$

Recall that $F_{BD}(k, N, q) = 1 - \sum_{i=0}^{i=k} p_i$, and therefore:

$$(n-k)\frac{f_{BD}(k, N, q)}{1-F_{BD(k,N,q)}} < \left(n-\hat{k}\right)\frac{f_{BD}(\hat{k}, N, q)}{1-F_{BD(k,N,q)}} \qquad \forall \hat{k} > k,$$

which implies that $(n-m)\frac{f_{BD}(m,N,q)}{1-F_{BD(m,N,q)}}$ is increasing in $m$. ■

**Proposition** 5

By Equation (9) and the specification of the second best mechanism we know that agent $j$ receives a positive ex ante utility if, and only if, $\mu^j_{N-1} > 0$. There are two cases: 1. The cost-to-benefit ratio of $\mu^j_{N-1}$ is positive; in this case, as $\alpha^j$ decrease, its cost-to-benefit ratio grows to infinity and in the optimal solution $\mu^j_{N-1}$ is set to 0. 2. The cost-to-benefit ratio of $\mu^j_{N-1}$ is negative; in this case optimality ensures that $\mu^j_{N-1} = 1$, regardless of $\alpha^j$.

Using the notations defined in the proof of theorem (1) the cost to surplus of $\mu^j_{N-1}$ is given by $b^j_{N-1}/a^j_{N-1}$. Since $a^j_{N-1}$ is always positive, the sign of this ratio depends only upon the sign of $b^j_{N-1}$, which is explicitly given by:

$$b^j_{N-1} = q^{-j}\Delta^j\left(1 - F_{BD}(N-1, N, q^j)\right) - q^{-j}(v^b_h - v^s_l - \Delta^j)\cdot(N-(N-1))\cdot f_{BD}(N-1, N, q^j).$$

Thus, the ratio is negative if, and only if,

$$\frac{\Delta^j}{\left(v^b_h - v^s_l\right) - \Delta^j} \leq \frac{f_{BD}(N-1, N, q^j)}{1 - F_{BD}(N-1, N, q^j)},$$

or equivalently:

$$N \geq \frac{q^b}{1-q^b}\cdot\frac{\Delta^j}{v^b_h - v^s_l - \Delta^j}.$$