Bonn Econ Discussion Papers

Discussion Paper 03/2012

Optimal Private Good Allocation: The Case for a Balanced Budget

by

Drexl, Moritz and Kleiner, Andreas

September 2012

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Kaiserstrasse 1
D-53113 Bonn

Financial support by the
Deutsche Forschungsgemeinschaft (DFG)
through the
Bonn Graduate School of Economics (BGSE)
is gratefully acknowledged.

Deutsche Post World Net is a sponsor of the BGSE.

Optimal Private Good Allocation: The Case for a Balanced Budget

Moritz Drexl and Andreas Kleiner*

First version: October 2011
This version: 26th September 2012

Abstract

In an independent private value auction environment, we are interested in strategy-proof mechanisms that maximize the agents' residual surplus, that is, the utility derived from the physical allocation minus transfers accruing to an external entity. We find that, under the assumption of an increasing hazard rate of type distributions, an optimal deterministic mechanism never extracts any net payments from the agents, i. e. it will be budget-balanced. Specifically, optimal mechanisms have a simple "posted price" or "option" form. In the bilateral trade environment, we obtain optimality of posted price mechanisms without any assumption on type distributions, thereby providing a rationale for confining attention to budget-balanced mechanisms.

*We wish to thank seminar participants at the University of Bonn and especially Benny Moldovanu for numerous insightful discussions. Drexl, Kleiner: Bonn Graduate School of Economics, University of Bonn, Kaiserstr. 1, 53113 Bonn, Germany; drexl@uni-bonn.de, akleiner@uni-bonn.de

1 Introduction

Most parts of the mechanism design literature studying welfare maximization problems focus on mechanisms implementing the efficient allocation. However, in general it is not possible to implement the efficient allocation in dominant strategies using budget-balanced mechanisms (Green and Laffont 1979). The same holds in the bilateral trade setting: No efficient and budget-balanced mechanism exists that also respects individual rationality, even if one considers the weaker notion of Bayes-Nash implementation (Myerson and Satterthwaite 1983).

Given these results, we study the question of how to choose among different mechanisms that cannot attain both allocative efficiency and budget-balancedness. Since we are concerned with welfare maximization, the social planner's objective function should consist of the agents' aggregate utility and therefore include aggregate transfers. In other words, it is natural to maximize what we call the residual surplus. This is the surplus, or utility, the agents derive from the chosen allocation, reduced by the amount of money that is burnt or lost to an external agency.

More specifically, we consider the auction of an indivisible good among two agents with independent private values, which are distributed according to prior type distributions. We consider *strategy-proof* mechanisms, where it is a dominant strategy for the agents to reveal their valuation truthfully. Since the incentives do not depend on the priors of the other agents, this ensures that truth-telling is robust to informational disturbances.[1] In order to achieve incentive compatibility, monetary payments may be imposed on the agents, as long as neither *individual rationality* nor *no deficit constraints* are violated. The aim is to identify among all strategy-proof mechanisms the one that provides the largest ex-ante residual surplus for the agents. This means that payments which cannot be redistributed to other agents are wasted. Our first result is that, under an increasing hazard rate assumption on type distributions, the trade-off between allocative efficiency and budget-balancedness is resolved completely in favour of a balanced budget (Theorem 1). Hence, the optimal deterministic mechanism will never waste any payments, thereby giving up large amounts of allocative efficiency. In fact, our proof method reveals that all mechanisms that allocate efficiently are worse than the simple mechanism where the object is always given to one of the agents (Corollary 1). The optimal mechanisms can be implemented either as a "posted price" or an "option" mechanism: The object is assigned to one of the agents unless both agents agree to trade at a prespecified price (posted price mechanism) or unless the second agent uses his option to buy the object at a fixed price from the first agent (option mechanism). Also, by imposing individual rationality in the bilateral trade setting, we are able to establish optimality of posted price mechanisms without any restrictions on type distributions (Theorem 2). This provides an argument for the focus on budget-balanced mechanisms (see Myerson and Satterthwaite 1983, Hagerty and Rogerson 1987). These results are interesting because the solution to this complex optimization problem is given by simple mechanisms that are easy to implement in practice. Moreover, they do not rely on the burning of money which is considered to be an unrealistic feature.

Our findings are related to a recent branch of the literature. Miller (2012) studies a

[1] Note that the set of strategy-proof mechanisms is independent of the type distributions; they only affect what the ex-ante optimal mechanism will be.

model of firms colluding in a Bertrand oligopoly. Efficiency considerations require the oligopoly to allocate the market to the firm with lowest costs. Since costs are private information, payments are required to incentivize their truthful revelation. On the other hand, money that flows out of the cartel decreases the welfare of the firms. Miller studies the question how a mechanism should be optimally designed to map costs into market share allocations. He shows that under general conditions it is never optimal to allocate market shares efficiently. We take up this question in our model and provide a simple mechanism that improves upon efficient mechanisms. Lacking general statements about the form of an optimal mechanism, Miller gives numerical evidence that for some type distributions it is optimal to give up large amounts of efficiency in order to obtain a balanced budget. However, other examples indicate that this observation does not hold for general distributions. For the special case of a uniform distribution, Shao and Zhou (2008) are able to completely characterize the optimal mechanism analytically, even allowing for stochastic mechanisms. They show that a mechanism is optimal if and only if it is a convex combination of simple posted price and option mechanisms.[2] Athey and Miller (2007) examine a similar question in the repeated bilateral trade setting and obtain numerical results suggesting that for many type distributions the optimal mechanism is a posted price mechanism.

Another strand of the literature studies the expected residual surplus of Bayesian incentive compatible mechanisms, but with the additional assumption that it is not possible to redistribute any payments among the agents (Hartline and Roughgarden 2008, Condorelli 2012). This simplifies the analysis to a great extent since methods similar to those in Myerson (1981) can be applied. Interestingly, one of the results is that for a large class of type distributions (those which exhibit an increasing hazard rate) it is optimal to always assign the object to the same agent. However, the assumption of no redistribution of payments is crucial, since otherwise the results do not hold and (to the best of our knowledge) no method for tackling the optimization problem analytically is known so far.

Focusing on the efficient allocation, an interesting question is which part of the Vickrey payments can be redistributed to the agents without losing incentive compatibility (Cavallo 2006). Starting with a budget-balanced mechanism, Tatur (2005) examines by how much allocative inefficiency can be reduced when allowing a given budget deficit. He does this in a bilateral trade setting with a large number of buyers and sellers.

We present our basic model for the auction environment in Section 2 and characterize incentive compatible mechanisms in Section 3. The optimization problem is solved in Section 4. In Section 5, we study this mechanism design problem in the bilateral trade context and conclude in Section 6. All proofs are relegated to the appendix.

[2]After the completion of this paper, we were made aware of a new version of the paper by Shao and Zhou (2012) in which they tackle the general problem. In contrast to our paper, they restrict themselves to environments where both agents' types are distributed according to the same distribution function, which allows them to focus on symmetric mechanisms. While we restrict attention to individually rational mechanisms, they do not impose this but emphasize that the optimal mechanism is individually rational. However, in the general case with ex ante asymmetric agents, this observation does not hold and individual rationality has to be imposed. In addition, they require a slightly more restrictive assumption on the distribution of types. Corollary 1 as well as Theorem 2 regarding the bilateral trade setting analyzed by Myerson and Satterthwaite (1983) have no counterpart in their paper.

2 Model

An indivisible object is auctioned among two agents. Each agent $i = 1, 2$ has a valuation x_i for the object. Valuations are drawn independently from $X_i = [0, \bar{x}_i]$ according to distribution functions F_i with corresponding densities f_i. We denote by $X = X_1 \times X_2$ the product type space and by F the joint distribution on X. For notational convenience, when concentrating on agent i, we will write (x_i, x_{-i}) for $x = (x_1, x_2) \in X$.

If agent i is given a payment of p_i (usually negative), his utility is $x_i + p_i$ for winning the object, and p_i if the other agent gets the object.

Mechanisms

Due to the Revelation Principle we shall focus on truthfully implementable direct revelation mechanisms for selling the object.

Definition 1. *A* mechanism *M is a tuple (d, p), where $d : X \to \{0, 1\}^2$ and $p : X \to \mathbb{R}^2$ are measurable functions, such that $d_1(x) + d_2(x) = 1$.*

The interpretation is that $d_i(x) = 1$ if and only if agent i gets the object. If the agents report x, then agent i receives as payment the component $p_i(x)$ of $p(x)$. Note that with this definition we restrict attention to deterministic allocation rules.

Equilibrium Concept

We consider strategy-proof mechanisms where truthful reporting is a dominant strategy for both agents. Therefore, we define the following notion of incentive compatibility:

Definition 2. *A mechanism M is* incentive compatible (IC) *if for every agent i and for each $x_i \in X_i$, $r_i \in X_i$,*

$$d_i(x_i, r_{-i}) \cdot x_i + p_i(x_i, r_{-i}) \geq d_i(r_i, r_{-i}) \cdot x_i + p_i(r_i, r_{-i})$$

holds for each $r_{-i} \in X_{-i}$.

This definition is independent of the distribution of valuations, which reflects the robustness of strategy-proof mechanisms as compared to mechanisms that are Bayes-Nash incentive compatible. Although the set of mechanisms we consider does therefore not depend on F, the next section shows that the distributions determine which mechanism is optimal.

Objective and Further Constraints

We aim at finding the mechanism that maximizes the sum of agents' ex ante (expected) residual surplus, that is, utility derived from the physical allocation minus aggregate payments. A natural constraint is that the mechanism has to be *ex post no-deficit (ND)*, that is, for every type profile x, we require $p_1(x) + p_2(x) \leq 0$. Also, the mechanism has to be *ex post individually rational (IR)*, that is, for all type profiles x, we require $d_i(x)x_i + p_i(x) \geq 0, i = 1, 2$. Summarizing, we want to solve the following optimization problem:

$$\max_{M=(d,p)} \int_X \Big[d_1(x)x_1 + d_2(x)x_2 + p_1(x) + p_2(x)\Big] dF(x) \tag{1}$$
$$\text{s. t. } M \text{ satisfies IC, ND and IR.}$$

We say that a mechanism is optimal if it solves problem (1).

3 Characterization of Incentive Compatibility

The aim of this section is to give a characterization of incentive compatibility in order to simplify problem (1). The conditions characterizing incentive compatible mechanisms involve a monotonicity and an integrability condition. We first define monotonicity.

Definition 3. *The allocation function d is* monotone *if d_i is non-decreasing in x_i for $i = 1, 2$.*

Now given a monotone allocation function d, define the following functions for $i = 1, 2$:

$$g_i(x_{-i}) := \inf\{x_i : d_i(x_i, x_{-i}) = 1\}.$$

If there is no x_i such that $d(x_i, x_{-i}) = 1$, then we set $g_i(x_{-i}) = \bar{x}_i$. Note that if d is monotone, these functions define d almost everywhere. The following lemma gives a characterization of incentive compatibility.

Lemma 1. *A mechanism $M = (d, p)$ is incentive compatible, if and only if the following two conditions are satisfied:*

1. *The allocation rule d is monotone.*
2. *For $i = 1, 2$ let $x_{-i} \in X_{-i}$ be given. Then for all $x_i \leq x_i' \in X_i$,*

$$p_i(x_i, x_{-i}) - p_i(x_i', x_{-i}) = \begin{cases} g_i(x_{-i}) & \text{if } d_i(x_i, x_{-i}) = 0 \text{ and } d_i(x_i', x_{-i}) = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The straightforward proof can be found in Appendix C. The interpretation of condition (2) is that an agent who receives the object is punished by paying a higher amount compared to the case where he would not have gotten the object. The punishment has to make the agent's marginal type $g_i(x_{-i})$ indifferent between receiving and not receiving the object.

It follows from Lemma 1 that if a mechanism satisfies IC, payments have the following form:

$$p_i(x_i, x_{-i}) = q_i(x_{-i}) - g_i(x_{-i}) d_i(x_i, x_{-i})$$

with some functions $q_i : X_{-i} \to \mathbb{R}$. This can be interpreted as a payoff-equivalence result: Payments are completely determined by the allocation as soon as one fixes the payment for some type x_i. Or, in other words, once the allocation is fixed, the only freedom that is left regarding the payment scheme, is to give the agent an additional payment

that is independent of his type. These additional payments can serve as a possibility to redistribute certain amounts of payments to another agent, as e. g. in Cavallo (2006).[3]

The simplified formulation of problem (1) is the following:

$$\max_{M=(d,q)} \int_X \Big[d_1(x)[x_1 - g_1(x_2)] + d_2(x)[x_2 - g_2(x_1)] + q_1(x_2) + q_2(x_1) \Big] dF(x) \qquad (3)$$
$$\text{s. t. } M \text{ satisfies IR and ND, and } d \text{ is monotone.}$$

We will write $U(M) = U(d, q)$ for the above integral and from now on only consider mechanisms that are IC, IR and ND.

4 The Optimal Auction

In this section, we present the first main result of this paper: if we impose an increasing hazard rate condition on the type distributions, then the optimal mechanism is always budget-balanced. Specifically, it turns out that the optimal mechanism takes one of two simple forms:

Either it is a posted price mechanism which by default allocates the object to one of the agents (agent 1, say) and changes the allocation if and only if both agents agree to trade at a prespecified price a, i. e. agent 1 reports a valuation below a fixed price a and agent 2 reports a valuation above a. If agent 2 is allocated the object, he makes a payment a to agent 1, otherwise no transfers accrue.

Or it is an option mechanism where the good is allocated by default to agent 1, but agent 2 has the option to buy the object at price a. Hence, if agent 2's valuation is above the strike price a, he buys the object and pays a to agent 1 (see also Shao and Zhou 2008).

Formally, these two mechanisms are defined as follows:

Definition 4. *A mechanism $M = (d, p)$ is a* posted price mechanism with default agent 1 and price a, *if*

$$\begin{array}{ll} d_2(x) = 1,\ p(x) = (a, -a) & \text{if } x_1 \leq a \text{ and } x_2 \geq a, \\ d_2(x) = 0,\ p(x) = (0,\ \ 0) & \text{otherwise.} \end{array}$$

M is an option mechanism with default agent 1 and price a, *if*

$$\begin{array}{ll} d_2(x) = 1,\ p(x) = (a, -a) & \text{if } x_2 \geq a, \\ d_2(x) = 0,\ p(x) = (0,\ \ 0) & \text{otherwise.} \end{array}$$

Similarly, one can define posted price and option mechanisms with default agent 1. If we do not specify the agent or price we just say that M is option or posted price.

[3]We note here that the additional assumption that $p_i(x) \leq 0$ for $i = 1, 2$, which is implicitly assumed in Hartline and Roughgarden (2008) and Condorelli (2012), together with ND and IR implies that $q_i(x_{-i}) \equiv 0$ for $i = 1, 2$. This reduces the complexity of the payment-scheme to a great extent, since then payments are completely determined through the allocation. In fact, by looking at the proof of our main result, one can see that with this simplification results analogous to those in the papers just mentioned can be proved. Thus, if positive transfers to the agents were not possible, then the increasing hazard rate condition implies that it is optimal to always assign the object to the same agent, thereby completely ignoring valuations.

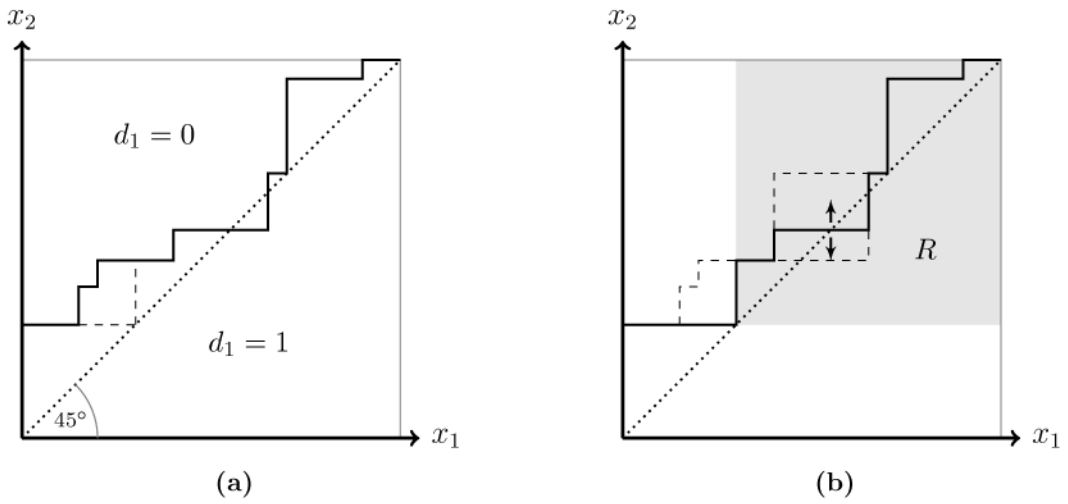


Figure 1: Illustration of the proof of Theorem 1

Both classes of mechanisms are parameterized by the price a and it is easy to check that all these mechanisms are budget-balanced as well as incentive compatible and individually rational.

Our assumption on type distributions is the following:

Condition (HR)**.** *The hazard rates of the type distributions are monotone. That is, the functions* $h_i(x_i) = \frac{f_i(x_i)}{1-F_i(x_i)}$ *are non-decreasing in* $x_i \in [0, \bar{x}_i)$ *for* $i = 1, 2$.

Theorem 1. *Given condition (HR), there is a posted price or an option mechanism which solves the optimization problem* (3)*.*

The proof (which can be found in Appendix A) can be sketched as follows: We first show the important auxiliary result that either an option mechanism or a posted price mechanism is optimal in $\mathcal{M}_{st}$, the class of mechanisms whose allocation functions are step functions (Lemma 2). We then argue that the welfare of a given mechanism can be approximated arbitrarily well by a mechanism in $\mathcal{M}_{st}$ (Lemma 3). The Theorem then follows by the following observation: Suppose there is a mechanism $\bar{M}$ being strictly better than the best option or posted price mechanism, and denote the welfare difference by ε. It follows from Lemma 3 that there is a mechanism in the class $\mathcal{M}_{st}$ whose welfare is within $\frac{\varepsilon}{2}$ of $U(\bar{M})$, thus being better than the best option or posted price mechanism. But this contradicts Lemma 2, hence there cannot be a mechanism being better than the best option or posted price mechanism.

To show Lemma 2, i. e. that an option or posted price mechanism is optimal within $\mathcal{M}_{st}$, we start with an arbitrary step function mechanism and manipulate it to end up with an option or posted price mechanism that is welfare-superior. To illustrate the arguments, we consider the mechanism shown in Figure 1a. The Figure shows the product type space and an allocation function that assigns the object to agent 1 for report profiles below the solid line and to agent 2 for profiles above the solid line. In *Step 1*, we argue that the maximal possible redistribution payments q_i are easy to determine: in the example shown, agent 2 receives no redistribution, while the redistribution to agent 1 is completely determined through the allocation. Using this observation, we argue in *Step 2* that

changing the allocation to the one shown in Figure 1b does not increase money burning, but increases allocative efficiency and hence aggregate welfare. To finish the argument, we study the effects of shifting steps in the set R, shown as the shaded area in Figure 1b, while fixing redistribution payments (*Step 3*). Our condition on the hazard rate ensures that each step should optimally be moved to either the lowest or the highest possible position. Hence, proceeding iteratively, we obtain either an option mechanism or a posted price mechanism, proving Lemma 2.

A consequence of the theorem is that, given the increasing hazard rates of the agents' type distributions, finding the best mechanism reduces to finding the best posted price and option mechanisms and comparing these two. For example, if the agents have the same distribution function, all option and posted price mechanisms with the same strike price yield the same welfare and therefore the best mechanism is characterized by the strike price a^* satisfying

$$a^* = \mathbb{E}[x_1] = \mathbb{E}[x_2].$$

Our intermediate results (see the proof of Lemma 2) also allow for a refined judgement of the welfare implied by the efficient allocation. Miller (2012) showed, under very general conditions, that the efficient allocation rule is never part of the optimal mechanism. We can strengthen this statement in our context by providing a mechanism that improves upon all efficient mechanisms. Surprisingly, this improvement can be achieved using an extremely simple mechanism:

Corollary 1. *Given condition (HR), every mechanism that allocates efficiently is dominated by a mechanism that always allocates the good to one of the agents.*

More precisely, a mechanism that is better than every efficiently allocating mechanism can be found simply by comparing the agents' type distributions, giving the good to the agent with the higher expected valuation and completely ignoring any reported types.

While optimal mechanisms for distributions obeying condition (HR) are very simple, the following example shows that if the condition is not satisfied the optimal mechanism need not be of the form stated in Theorem 1. The example also illustrates the role of (HR) in establishing the result.

Example 1. *Let the distribution function of two symmetric agents be given as*

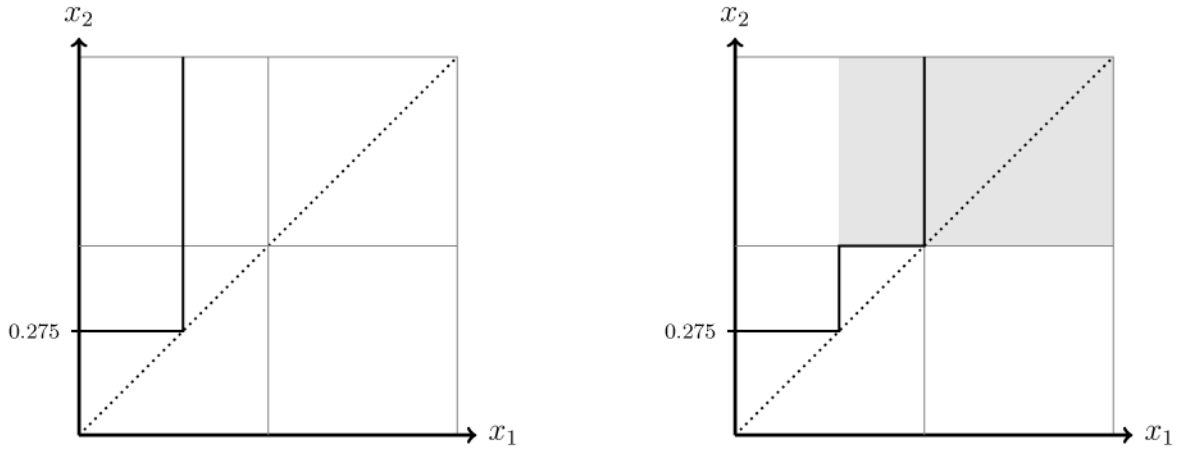
$$f(x_i) = \begin{cases} 0.9 & \text{if } x_i \leq 0.5 \\ 0.1 & \text{otherwise.} \end{cases}$$

Due to the sharp jump downwards at 0.5*, f does not satisfy condition (HR). The optimal posted price mechanism (which is as good as the optimal option mechanism) has a strike price of $a^* = 0.275$, attaining a social welfare of* 0.0718*. However, the following mechanism M attains a higher social welfare of* 0.0741*: Set*

$$d_2(x) = 1 \quad \Leftrightarrow \quad (x_2 \geq a^* \text{ and } x_1 \leq a^*) \text{ or } (x_2 \geq 0.5 \text{ and } x_1 \leq 0.5),$$

and set $q_2(x_1) \equiv 0$, as well as

$$q_1(x_2) = \begin{cases} 0 & \text{if } x_2 \leq a^* \\ a^* & \text{otherwise.} \end{cases}$$



(a) Optimal posted price mechanism. **(b)** Mechanism M.

Figure 2: Mechanisms presented in Example 1

This mechanism and the best option mechanism are depicted in Figure 2. One can see that the allocation of mechanism M is more efficient. Because the induced higher payments cannot be redistributed, payments of $(0.5 - 0.275) = 0.225$ *are lost for type profiles in the shaded area in Figure 2b. But still, since type profiles x with* $x_1, x_2 \geq 0.5$ *appear so rarely (with density* 0.01*), this does not counter the positive effect due to the better allocation. In this sense, an increasing hazard rate ensures that lost payments can never be weighed out by an improved efficiency of the allocation.*

5 Bilateral Trade

Myerson and Satterthwaite (1983) showed that one cannot implement the efficient allocation in the bilateral trade setting in a budget-balanced and individually rational way. They provide a characterization of the optimal budget-balanced and individually rational mechanism. In the same environment, Hagerty and Rogerson (1987) study the set of dominant-strategy implementable mechanisms that are budget-balanced and individually rational, showing that essentially only posted price mechanisms fulfill these conditions. However, a priori it is not clear why one should restrict the search for the optimal mechanism to mechanisms with a balanced budget. After all, it is conceivable that deviating from a balanced budget could improve incentives and therefore lead to higher welfare of the mechanism. However, the results in this section show that in the bilateral trade environment with general prior type distributions, the restriction to budget-balanced mechanisms does not reduce aggregate welfare.

Let the model and notation be as in Section 2, but assume now that agent 1 (called the "seller" from now on and indexed by S) is the owner of the good before participating in the mechanism (whereas agent 2 is called the "buyer" and indexed by B). By a buyer posted price mechanism (B-PP) we denote a posted price mechanism in which the buyer gets the object if and only if he announces a type high enough, and the seller a type that is low enough. Again, we are looking for a mechanism that maximizes the sum of the expected utilities of the agents, taking monetary transfers into account. The fact that in the bilateral trade setting the seller initially owns the good requires a stronger condition

for a mechanism to be individually rational: now the outside option for a seller is to not participate in the mechanism and to keep the object. Hence, for a mechanism to be individually rational,

$$d_S(x)\ x_S + p_S(x) \geq x_S \qquad \text{and} \qquad d_B(x)\ x_B + p_B(x) \geq 0 \tag{IR'}$$

must hold for all $x \in X$.

Thus, a mechanism is optimal if it solves

$$\max_{M=(d,p)} \int_X \Big[d_S(x)x_S + d_B(x)x_B + p_S(x) + p_B(x) \Big] dF(x) \tag{4}$$
$$\text{s. t. } M \text{ satisfies IC, ND and IR'.}$$

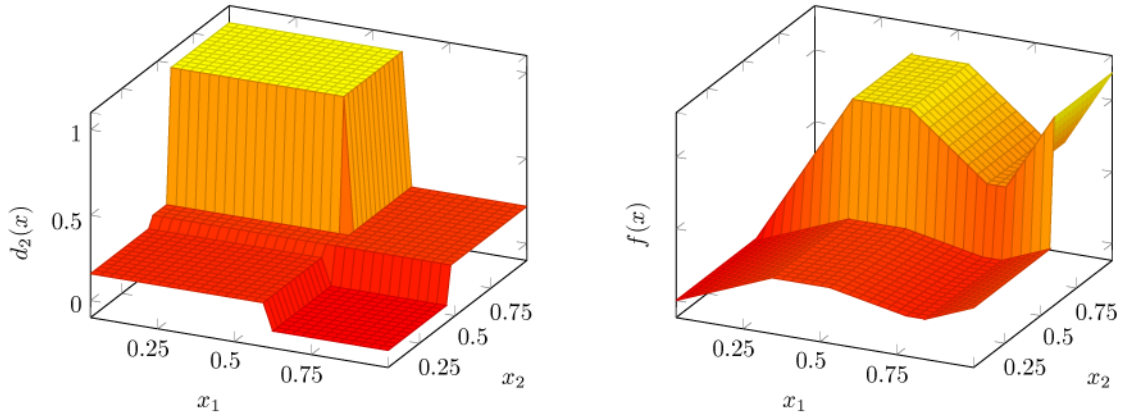
Theorem 2. *There is a B-PP mechanism that solves problem* (4).

The proof can be found in Appendix B and is similar to the proof of Theorem 1 in that for every given mechanism we construct a posted price mechanism that is welfare-superior. Theorem 2 implies that the optimal mechanism is budget-balanced. Thus, focusing on dominant-strategy implementation, this result provides a justification for confining attention to budget-balanced mechanisms.

6 Discussion

We have studied the trade-off between efficiency and budget-balancedness in an independent private values auction model. We incorporated this into the model by letting the social welfare objective function include all payments, that is, by maximizing residual surplus. We showed that, if one focuses on robust implementation in dominant strategies, an increasing hazard rate condition on agents' type distributions guarantees a resolution of the trade-off completely in favor of a balanced budget. In addition, budget-balanced mechanisms have a very simple form and can easily be implemented as posted price or option mechanisms. Further, we showed that without any assumption on the prior distribution of types a posted price mechanism is optimal in the bilateral trade setting. This provides a strong rationale for focusing on budget-balanced mechanisms, as done by Myerson and Satterthwaite (1983) and Hagerty and Rogerson (1987).

The findings in this paper could be extended in a number of important ways. First, allowing for stochastic mechanisms in our analysis is an important research avenue. Numerical examples strongly suggest that in the auction environment, even if the distributions of types satisfy condition (HR), stochastic mechanisms can be strictly better than the best deterministic ones: Figure 3 shows a distribution of types and the corresponding optimal mechanism (obtained by numerically solving a discretized version of our model[4]). This mechanism is not budget-balanced and improves strictly upon the best deterministic mechanism. Unfortunately, the structure of this optimal mechanism suggests that an easy analytical description of optimal stochastic mechanisms is not likely to be found with our assumptions. Therefore, one approach is to find general additional conditions on type distributions such that the designer cannot improve using stochastic mechanisms. Note



(a) Optimal allocation. **(b)** Joint density.

Figure 3: An instance where the optimal mechanism is not deterministic, although the densities exhibit an increasing hazard rate.

here, that Shao and Zhou (2008) have proved the sub-optimality of stochastic mechanisms if types are distributed uniformly.

Another interesting question is how the result generalizes to a model including more than two agents. We strongly believe that the optimal mechanism will still be budget-balanced. An important argument for this is that, as the number of agents gets large, the efficient allocation can be approximated in a budget-balanced way: in the spirit of McAfee (1992), allocate efficiently while ignoring one agent who then receives all payments from the other agents. This can be implemented by tentatively giving the object to one of the agents and then simulating a second price auction with reserve price where this agent sells the object to the remaining agents.

Studying optimal Bayesian incentive compatible mechanisms without imposing budget-balancedness is another interesting topic for future research.

A Proof of Theorem 1

The theorem follows from two lemmas. First we define what a step function is.

Definition 5. *A* step function *is a function $\varphi_2 : X_1 \to X_2$ that can be written as*

$$\varphi_2 = \sum_{j=1}^{n} \beta_j \chi_{A_j},$$

where $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n \in X_2$, $A_1, \ldots, A_n$ is an ordered partition of X_1 and χ_A denotes the indicator function of a set A. A step function is simple, *if the sets A_j have the form $A_1 = [0, \alpha_1], A_2 = (\alpha_1, \alpha_2], \ldots, A_n = (\alpha_{n-1}, \alpha_n]$.*

The allocation rule d associated with *the step function φ_2 is defined through*

$$d_2(x_1, x_2) = 1 \quad \Leftrightarrow \quad x_2 \geq \varphi_2(x_1).$$

[4]The instance as well as the source code for computing the optimal mechanism can be obtained from the authors upon request.

Note that if d is associated with the step function φ_2, then $g_2 = \varphi_2$. Similarly we can then define a step function for agent 1 through $\varphi_1 = g_1$. This will be used in the following proofs.

The first lemma says that, under (HR), option or posted price mechanisms are optimal step function mechanisms, that is, within the class of all mechanisms whose allocation rule is associated with a step function ($\mathcal{M}_{st}$), posted price and option mechanisms yield the highest residual surplus.

Lemma 2. *Assume condition (HR) and let $M = (d, q)$ be any mechanism where d is associated with the step function φ_2. Then there exists a mechanism M' that is posted price or option such that $U(M') \geq U(M)$.*

Proof. The proof consists of three steps, where we constructively manipulate M in order to end up with the desired mechanism M'.

Step 1: Let β_j and A_j $(j = 1, \ldots, n)$ be given by φ_2. We first argue that, without loss of generality, we can assume $\beta_1 > 0$. For if $\beta_1 = 0$ and $A_1 = \{0\}$, we could set $\beta_1 = \beta_2$ without reducing $U(M)$, where without loss of generality $\beta_2 > 0$. If $\beta_1 = 0$ and $A_1 \neq \{0\}$, we could switch the roles of the agents and then describe d in terms of a step function with $\beta_1 > 0$.

It then follows that $q_2(x_1) = 0,\ \forall x_1 \in X_1$. To see this, pick some x_1. Then $\beta_1 > 0$ implies $\varphi_1(0)d_1(x_1, 0) = \varphi_2(x_1)d_2(x_1, 0) = 0$. From ND it follows that $q_1(0) + q_2(x_1) \leq 0$. Also, IR for agent 2 at $(x_1, 0)$ implies $q_2(x_1) \geq 0$, and IR for agent 1 at $(0, 0)$ implies $q_1(0) \geq 0$, and therefore $q_2(x_1) = 0$.

Next, we can assume that

$$q_1(x_2) = \min_{x_1} \left\{ \varphi_1(x_2)d_1(x_1, x_2) + \varphi_2(x_1)d_2(x_1, x_2) \right\} \tag{5}$$

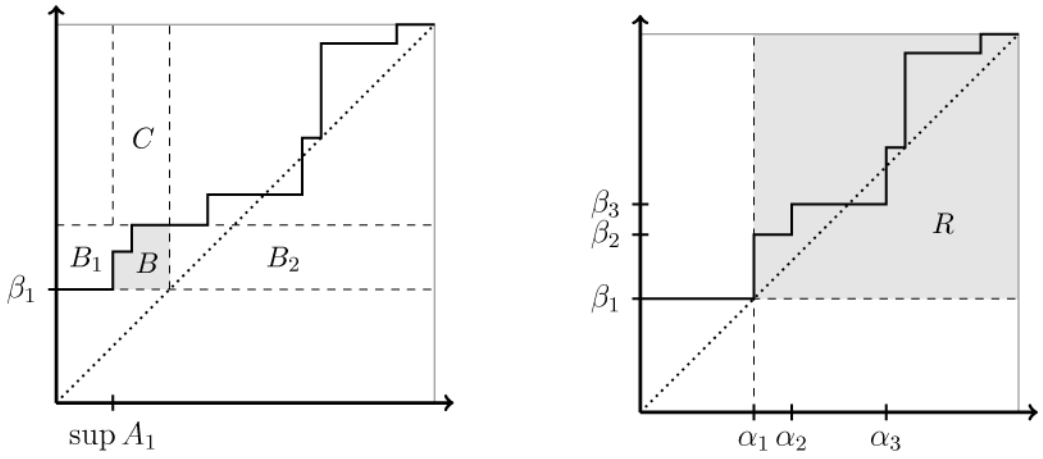
always holds, since by ND this relation always holds with $\leq$ and changing it to equality does not reduce $U(M)$. In this way, the complete payment-scheme is determined through the allocation d.

Step 2: We distinguish two cases. If $\sup A_1 \geq \beta_1$, we do nothing and go directly to Step 3. Otherwise, we modify and improve the mechanism in the following way:

We set $d_2(x_1, x_2) = 1$ and $d_1(x_1, x_2) = 0$ for $x_1 \leq \beta_1, x_2 \geq \beta_1$. This will change the functions φ_i, while q_2 remains unchanged, and q_1 adjusts according to (5). We claim that $U(M)$ did not decrease during this operation. To see this, take some $x \in X$ and look at the effects due to the change in allocation and payments. Let $B \subseteq X$ be the set of types where the allocation changed,

$$\begin{aligned} B_1 &:= \{(x_1, x_2) \in X \setminus B : q_1(x_2) \text{ changed},\ x_1 \leq \beta_1\} \\ B_2 &:= \{(x_1, x_2) \in X \setminus B : q_1(x_2) \text{ changed},\ x_1 \geq \beta_1\} \\ C &:= \{(x_1, x_2) \in X : \varphi_2(x_1) \text{ changed},\ x_2 \geq \tilde{x}_2\ \forall \tilde{x}_2 \text{ s.t. } q_1(\tilde{x}_2) \text{ changed}\}. \end{aligned}$$

These sets are depicted in Figure 4a. Then we know that the integrand in (3) did not change for types outside these sets. The allocation only changed in B, where we have $x_1 \leq x_2$ and $d_2(x) = 1$, and therefore the allocation effect is positive. So we only need to concentrate on payments. Consider the case where $x \in C$. Here, $q_1(x_2)$ did not change and $\varphi_2(x_1)$ was reduced. Therefore, the effect is non-negative. The case $x \in B_1$ is like the



(a) Sets as defined in the proof of Lemma 2.

(b) Situation after the object is assigned to agent 2 in the set B.

Figure 4

one above, except that $q_1(x_2)$ increased, so the effect is non-negative. If $x \in B_2$, nothing changed for agent 2 and $q_1(x_2) - \varphi_1(x_2) = 0$ before and after the change, so there is no effect in this case. Finally, for $x \in B$, agent 2 had to pay nothing before and now has to pay β_1, whereas agent 1 used to pay nothing and now receives β_1, so the net effect is zero.

Step 3: From now on, we only change $d(x)$ for $x \in R := \{(x_1, x_2) \in X : x_1, x_2 \geq \beta_1\}$. Looking at (5), one sees that $q_1(x_2)$ will never be affected by those changes and therefore we can completely ignore the functions q_i from now on. We show that either always giving the object to agent 1 (posted price) or always giving it to agent 2 (option) in R does not decrease $U(M)$. This will complete the proof.

Since we can ignore the functions q_i, changing d on a null set will not affect the value $U(M)$ and therefore we can from now on assume that φ is simple and that $\alpha_1 = \beta_1$ (c. f. Figure 4b). Leaving α_1 and β_1 fixed, we give a procedure that removes the second step (i. e. the first step contained in R) without decreasing $U(M)$. The procedure distinguishes two cases.

(a) $\beta_2 = \beta_1$. We vary α_2 on the interval $[\alpha_1, \alpha_3]$ and look at how $U(M)$ changes. The part of $U(M)$ that depends on α_2 is the following:

$$\int_{\beta_2}^{\beta_3} \left[\int_{\alpha_1}^{\alpha_2} (x_2 - \beta_2) dF_1(x_1) + \int_{\alpha_2}^{\bar{x}_1} (x_1 - \alpha_2) dF_1(x_1) \right] dF_2(x_2) \\ - \int_{\beta_3}^{\bar{x}_2} \left[\int_{\alpha_1}^{\alpha_2} \beta_2 dF_1(x_1) + \int_{\alpha_2}^{\alpha_3} \beta_3 dF_1(x_1) \right] dF_2(x_2)$$

Differentiating with respect to α_2 using Leibniz' rule yields

$$\int_{\beta_2}^{\beta_3} \Big[f_1(\alpha_2)(x_2 - \beta_2) - \big[1 - F_1(\alpha_2)\big] \Big] dF_2(x_2) + \int_{\beta_3}^{\bar{x}_2} f_1(\alpha_2)[\beta_3 - \beta_2] dF_2(x_2).$$

If we pull the factors that do not depend on x_1 and x_2 out of the integrals and write constants C_1, C_2 and C_3 for the integrals (which do not depend on α_2) we get

$$C_1 f_1(\alpha_2) - C_2\big[1 - F_1(\alpha_2)\big] + C_3 f_1(\alpha_2).$$

Assuming $C_2[1 - F_1(\alpha_2)] > 0$ (if either $C_2 = 0$ or $1 - F_1(\alpha_2) = 0$, we set $\alpha_2^* = \alpha_3$ without reducing U), we can divide by $C_2[1 - F_1(\alpha_2)]$ and get that the derivative is non-negative if and only if

$$C \cdot h_1(\alpha_2) - 1 \geq 0,$$

where $C = (C_1 + C_3)/C_2 > 0$. Because $h_1(\alpha_2)$ is non-decreasing by condition (HR), it follows that $U(M)$ is increased by either setting $\alpha_2^* = \alpha_1$ or $\alpha_2^* = \alpha_3$. In either case, delete step 2 from the step function, and in the latter case set $\beta_3 = \beta_1$. We then have decreased the number of steps by one and the procedure ends.

(b) $\beta_2 > \beta_1$. We vary β_2 on the interval $[\beta_1, \beta_3]$. Similar arguments as above establish that $U(M)$ is increased by setting $\beta_2^* = \beta_1$ or $\beta_2^* = \beta_3$. In the former case we can go to case *(a)* above, and in the latter case we can drop step 2 from the step function. Thus we either end up in case *(a)* or have decreased the number of steps by one.

Iteratively applying this procedure establishes the lemma. □

The next lemma enables us to approximate any mechanism with mechanisms from the class $\mathcal{M}_{st}$.

Lemma 3. *For every mechanism $M = (d, q)$ and for every $\varepsilon > 0$ there exists a mechanism $\tilde{M} = (\tilde{d}, \tilde{q})$ characterized by a step function such that $U(M) - U(\tilde{M}) < \varepsilon$.*

Proof. Let the mechanism $M = (d, q)$ and $\varepsilon > 0$ be given and let $g_1(x_2)$ and $g_2(x_1)$ be defined as above. Define $D_i := \{x \in X : d_i(x) = 1\}$ as the set of type profiles where agent i gets the object. Since g_2 is a monotone function it can be approximated uniformly by a monotone step function φ_2. Denote the allocation rule associated with φ_2 by $\tilde{d}$. By choosing the step width small enough the approximation can be done such that for given $\delta > 0$,

$$||g_1 - \varphi_1||_\infty < \delta \quad \text{and} \quad ||g_2 - \varphi_2||_\infty < \delta$$

holds. The approximation can be chosen such that $g_i(x_{-i}) = \bar{x}_i$ implies $\varphi_i(x_{-i}) = \bar{x}_i$ and φ_2 can be chosen such that $\varphi_2 \leq g_2$, implying that $\tilde{D}_1 \subset D_1$.

Without loss of generality, we can assume that $q_2(x_1) \equiv 0$ (see Step 1 in the proof of Lemma 2). By construction of φ_2 and since M satisfies ND, we can define functions $\tilde{q}_i(x_{-i})$ such that $\tilde{q}_2(x_1) \equiv 0$, $0 \leq \tilde{q}_1(x_2) \leq \inf_{x_1}\{\varphi_1(x_2)\tilde{d}_1(x_1, x_2) + \varphi_2(x_1)\tilde{d}_2(x_1, x_2)\}$ $\forall x_2 \in X_2$ and $||\tilde{q}_1 - q_1||_\infty < \delta$. We then have:

$$\begin{aligned}
U(d,q) - U(\tilde{d},\tilde{q}) \leq & \int_X q_1(x_2) - \tilde{q}_1(x_2)\ dF(x) \\
& + \int_{D_1} x_1 - g_1(x_2)\ dF(x) - \int_{\tilde{D}_1} x_1 - \varphi_1(x_2)\ dF(x) \\
& + \int_{D_2} x_2 - g_2(x_1)\ dF(x) - \int_{\tilde{D}_2} x_2 - \varphi_2(x_1)\ dF(x) \\
\leq \delta + & \int_{D_1 \setminus \tilde{D}_1} x_1 - g_1(x_2)\ dF(x) \\
& + \int_{\tilde{D}_1} \delta\ dF(x) + \int_{\tilde{D}_2 \setminus D_2} x_2 - g_2(x_1)\ dF(x) + \int_{D_2} \delta\ dF(x) \\
\leq & \ 5\delta.
\end{aligned}$$

Hence, by choosing $\delta < \frac{\varepsilon}{5}$, it follows that $U(d, q) - U(\tilde{d}, \tilde{q}) < \varepsilon$. □

Now we can put everything together, proving the theorem.

Proof of Theorem 1. Without loss of generality, we restrict ourselves to posted price mechanisms for agent 2. We first establish that U maps the set of all posted price mechanisms to a compact subset of $\mathbb{R}$. Let $\bar{a} = \min\{\bar{x}_1, \bar{x}_2\}$ and let $a \in [0, \bar{a}]$ be some price for a posted price mechanism M_a. Then $U(M_a)$ can be written as

$$U(M_a) = \int_0^a \int_a^{\bar{x}_2} x_2 dF(x) + \int_0^{\bar{x}_1} \int_0^a x_1 dF(x) + \int_a^{\bar{x}_1} \int_a^{\bar{x}_2} x_1 dF(x).$$

Due to the continuity of F, this function is continuous with respect to a. Since $[0, \bar{a}]$ is compact, so is $\{U(M_a) \mid a \in [0, \bar{a}]\}$ and therefore there exists an a^* such that $U(M_{a^*})$ is maximal among all posted prices.

Next, assume that the theorem is false, i. e. there exists a mechanism M and $\varepsilon > 0$ such that $U(M) > U(M_{a^*}) + \varepsilon$. Then apply Lemma 3 to M and ε to get a mechanism $\tilde{M} = (\tilde{d}, \tilde{q}) \in \mathcal{M}_{st}$ with $U(\tilde{M}) > U(M_{a^*})$. This contradicts Lemma 2, establishing the theorem. □

B Proof of Theorem 2

The proof consists of two steps. First, we set the stage by showing properties that any mechanism M always has. Then, we improve M, making it a B-PP mechanism.

Step 1: Preparation Note first that (IR') implies that $d_S(x_S, x_B) = 1$ for all $x_S > x_B$: Summing the two individual rationality constraints we get $(1 - d_S(x_S, x_B))(x_B - x_S) + p_S(x_S, x_B) + p_B(x_S, x_B) \geq 0$. The no deficit constraint together with $x_S > x_B$ imply that the inequality can hold only if $d_S(x_S, x_B) = 1$.

It follows that $q_B(x_S) = 0$, $\forall x_S \in X_S$. To see this, pick some x_S. Then $g_S(0)d_S(x_S, 0) = g_B(x_S)d_B(x_S, 0) = 0$ because $d_S(x_S, x_B) = 1$ for all $x_S > x_B$. From ND it then follows that $q_S(0) + q_B(x_S) \leq 0$. Also, IR' for the buyer at $(x_S, 0)$ implies $q_B(x_S) \geq 0$, and IR for the seller at $(0, 0)$ implies $q_S(0) \geq 0$, and therefore $q_B(x_S) = 0$.

Next, we can assume that

$$q_S(x_B) = \min_{x_S} \left\{ g_S(x_B)d_S(x_S, x_B) + g_B(x_S)d_B(x_S, x_B) \right\} \tag{6}$$

always holds, since by ND this relation always holds with $\leq$ and changing it to equality does not reduce $U(M)$. In this way, the complete payment-scheme is determined through the allocation d.

We now show that because of IR', $d_S(x_S, x_B) = 1$ whenever $x_S, x_B > g_B(0)$. So take some x with $x_S, x_B > g_B(0)$ and assume on the contrary that $d_S(x) = 0$. At the profile $(0, x_B)$ we have $g_S(x_B)d_S(0, x_B) + g_B(0)d_B(0, x_B) = g_B(0)$ which, using (6), implies that $q_S(x_B) \leq g_B(0)$. So the payoff for the seller at x is

$$0 \cdot x_S + q_S(x_B) \leq g_B(0) < x_S,$$

which contradicts individual rationality for the seller at point x.

Step 2: Improving the mechanism If the allocation function of this mechanism does not already correspond to a B-PP mechanism, we modify and improve the mechanism in the following way, making it a B-PP mechanism and completing the proof.

We set $d_B(x_S, x_B) = 1$ and $d_S(x_S, x_B) = 0$ for $x_S \leq g_B(0), x_B \geq g_B(0)$. Letting q_S adjust according to (6), this new mechanism corresponds to a B-PP mechanism and we claim that $U(M)$ did not decrease during this operation. To see this, note that $p_S(x) + p_B(x) \leq 0$ by ND and this inequality holds with equality in a B-PP mechanism. Hence, the effect of changing the mechanism on aggregate payments is non-negative. Since the allocation only changed for x such that $x_B \geq x_S$ and the new allocation rule prescribes $d_B(x) = 1$ for such x the allocation effect of changing the mechanism is positive. Hence, this operation increased $U(M)$, completing the proof. □

C Proof of Lemma 1

Necessity: Let any allocation d be given. We first prove that d_1 (and similarly d_2) is non-decreasing. Fix x_2 and let $x_1 < x_1'$ with $d_1(x_1, x_2) = 1$ and assume to the contrary that $d_1(x_1', x_2) = 0$. Then, by incentive compatibility for agent 1, we get

$$x_1 + p_1(x_1, x_2) \geq 0 + p_1(x_1', x_2) \geq x_1' + p_1(x_1, x_2),$$

or equivalently $x_1 \geq x_1'$, which is a contradiction.

Now we prove that (2) holds for agent 1. Take any $x_1,\ x_1' \in X_1$ such that $d_1(x_1, x_2) = 1$ and $d_1(x_1', x_2) = 1$, and assume without loss of generality that $p_1(x_1', x_2) > p_1(x_1, x_2)$. Then we get by incentive compatibility for type x_1

$$x_1 + p_1(x_1, x_2) \geq x_1 + p_1(x_1', x_2) > x_1 + p_1(x_1, x_2),$$

a contradiction. Therefore, $p_1(x_1', x_2) = p_1(x_1, x_2)$ must hold. The argument is the same for the case where $d_1(x_1, x_2) = 0$ and $d_1(x_1', x_2) = 0$. So take any $x_1,\ x_1' \in X_1$ such that $d_1(x_1, x_2) = 0$, $d_1(x_1', x_2) = 1$ and assume without loss of generality that $p_1(x_1, x_2) - p_1(x_1', x_2) > g_1(x_2)$. Then there exists a type $x_1'' \in X_1$ with the property that $p_1(x_1, x_2) - p_1(x_1', x_2) > x_1'' > g_1(x_2)$. From $x_1'' > g_1(x_2)$ it follows that $d(x_1'', x_2) = 1$ and therefore by the case shown above $p_1(x_1'', x_2) = p_1(x_1', x_2)$. Rearranging, this gives us $p_1(x_1, x_2) > x_1'' + p_1(x_1'', x_2)$, a contradiction to incentive compatibility for type x_1''.

Sufficiency: Let $x_2 \in X_2$. We are done if we can show that for any pair $x_1 < x_1'$, type x_1 does not want to report x_1' and vice versa. The cases $d_1(x_1, x_2) = 1$, $d_1(x_1', x_2) = 1$ and $d_1(x_1, x_2) = 0$, $d_1(x_1', x_2) = 0$ are obvious, so let $d_1(x_1, x_2) = 0$, $d_1(x_1', x_2) = 1$ (the other case is not possible due to monotonicity). Monotonicity also tells us that $x_1 \leq g_1(x_2) \leq x_1'$. We then have the following inequalities:

$$\begin{aligned} x_1' + p_1(x_1', x_2) &\geq g_1(x_2) + p_1(x_1', x_2) = p_1(x_1, x_2) \\ p_1(x_1, x_2) &= g_1(x_2) + p_1(x_1', x_2) \geq x_1 + p_1(x_1', x_2) \end{aligned}$$

Here, the equalities stem from (2). These inequalities imply that no type x_1' wants to report x_1 and vice versa. □

References

Athey, S. and Miller, D. A. (2007). Efficiency in Repeated Trade with Hidden Valuations, *Theoretical Economics* **2**(3): 299–354.

Cavallo, R. (2006). Optimal Decision-Making with Minimal Waste: Strategyproof Redistribution of VCG Payments, *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '06, ACM, New York, NY, USA, pp. 882–889.

Condorelli, D. (2012). What Money can't buy: Efficient Mechanism Design with Costly Signals, *Games and Economic Behavior* **75**(2): 613–624.

Green, J. R. and Laffont, J.-J. (1979). *Incentives in Public Decision-Making*, North-Holland Pub. Co., Amsterdam, NL.

Hagerty, K. M. and Rogerson, W. P. (1987). Robust Trading Mechanisms, *Journal of Economic Theory* **42**(1): 94–107.

Hartline, J. D. and Roughgarden, T. (2008). Optimal Mechanism Design and Money Burning, *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, STOC '08, ACM, pp. 75–84.

McAfee, R. P. (1992). A Dominant Strategy Double Auction, *Journal of Economic Theory* **56**(2): 434–450.

Miller, D. A. (2012). Robust Collusion with Private Information, *Review of Economic Studies* **79**(2): 778–811.

Myerson, R. B. (1981). Optimal Auction Design, *Mathematics of Operations Research* **6**(1): 58–73.

Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory* **29**(2): 265–281.

Shao, R. and Zhou, L. (2008). Optimal Allocation of an Indivisible Good. Working Paper, Yeshiva University and Shanghai Jiao Tong University.

Shao, R. and Zhou, L. (2012). Optimal Allocation of an Indivisible Good. Working Paper, Yeshiva University and Shanghai Jiao Tong University.

Tatur, T. (2005). On the Trade off Between Deficit and Inefficiency and the Double Auction with a Fixed Transaction Fee, *Econometrica* **73**(2): 517–570.