Serdica J. Computing **1** (2007), 293–312

Serdica
Journal of Computing

Bulgarian Academy of Sciences
Institute of Mathematics and Informatics

AUCTIONS WITH UNTRUSTWORTHY BIDDERS

Sviatoslav Braynov, Radoslav Pavlov

Abstract. The paper analyzes auctions which are not completely enforceable. In such auctions, economic agents may fail to carry out their obligations, and parties involved cannot rely on external enforcement or control mechanisms for backing up a transaction. We propose two mechanisms that make bidders directly or indirectly reveal their trustworthiness. The first mechanism is based on discriminating bidding schedules that separate trustworthy from untrustworthy bidders. The second mechanism is a generalization of the Vickrey auction to the case of untrustworthy bidders. We prove that, if the winner is considered to have the trustworthiness of the second-highest bidder, truthfully declaring one's trustworthiness becomes a dominant strategy. We expect the proposed mechanisms to reduce the cost of trust management and to help agent designers avoid many market failures caused by lack of trust.

1. Introduction Auctions have been extensively used in e-commerce as a means for price determination for multilateral trading without market intermediaries. They are particulary useful in markets with incomplete and asymmetric information, where the bidders' private information is the main factor determining strategic behavior.

ACM Computing Classification System (1998): I.2.11
Key words: auctions, e-commerce, mechanism design, trust.

As mechanisms for distributed optimization, auctions can offer several computational challenges. Determining the winners in combinatorial auctions, for example, is a complex optimization problem that has been recently studied [30, 27, 26, 13]. Several bidding languages have been proposed in an effort to reduce the communication overhead [15, 28]. Another important thread tries to identify auction protocols limiting the preferences that are to be revealed by bidders [1, 33, 10].

Most of the literature on auction theory has focused on well defined and enforceable auctions. It is usually assumed that auction results are binding for the auctioneer and bidders. That is, each party behaves as expected, and carries out their obligations. Many on-line auctions do not usually meet this assumption. Economic agents can fail to perform their tasks or to meet their commitments due to lack of incentives, lack of ability, or circumstances beyond their control. Internet users still fear the possibility of fraud, identity change, misuse of private information, etc. Complaints about online auction transactions have skyrocketed, accounting for 87 percent of the Internet fraud reports made to the Internet Fraud Watch in the first six months of 2002, compared to 70 percent in 2001 (`www.fraud.org`).

This paper analyzes auctions which are not completely enforceable. In such auctions, economic agents may fail to carry out their obligations and parties involved cannot rely on external enforcement or control mechanisms for backing up a transaction. An important characteristic of these settings is the risk of losses due to failure, fraud, or inability of other parties to fulfill their contractual obligations. Another important characteristic is the presence of asymmetric information. That is, untrustworthy agents may not communicate private information concerning their contractual abilities or intentions.

In the paper, we analyze a multidimensional auction in which a trustworthy buyer faces many sellers with varying degree of trustworthiness. The buyer does not know the bidders' trustworthiness and has to move first after the auction has been closed. That is, the buyer has to make the payment without having any guaranties of delivery. Such a setting raises several important questions. The first one is how to evaluate bids. The profitability of each bid depends on the bidder's trustworthiness, which is privately known to the bidder. Another question is what kind of incentive-compatible schemes are possible and what is their economic efficiency.

Many applications of mechanism design [24, 21, 25, 23] consider schemes that provide sufficient incentives to parties to reveal privately known information. The problem in our case is that the auctioneer faces uncertain profits and has to move first without being able to condition his payment on contractual performance. For example, the standard Vickrey auction fails to provide bidders with sufficient incentives to truthfully declare their trustworthiness.

In the paper, we study two mechanisms that make agents truthfully reveal their trustworthiness. The first mechanism is based on constrained bidding, in which the auctioneer offers different bidding schedules for different types of bidders. The schedules are designed to separate trustworthy from untrustworthy bidders. That is, all trustworthy bidders choose one schedule, while untrustworthy bidders choose another. This eliminates information asymmetry, and allows the auctioneer to evaluate bids using the actual bidders' trustworthiness.

The second mechanism is a generalization of the Vickrey auction to the case of untrustworthy bidders. In the auction, the highest bidder wins and the terms of trade are chosen as if the winner had the trustworthiness of the second-highest bidder.

The auction analyzed in the paper is three-dimensional, where sellers bid on price and quantity, besides reporting their trustworthiness. Multidimensional auctions arise frequently and have been extensively studied [8, 5]. For example, many defense procurement auctions involve multidimensional bids on promised technical characteristics, delivery date, estimated project costs, etc.

The paper is organized as follows. Section 2 provides a brief formalization of trust in the context of e-commerce. Section 3 defines the problem setting in which an auctioneer faces many bidders with varying degree of trustworthiness. A discriminating auction based on several bidding schedules is described in Section 4. In the auction, agents reveal their trustworthiness by choosing different schedules. Section 5 presents a generalization of the Vickrey auction to the case of untrustworthy bidders. Finally, the paper concludes by summarizing the results and providing directions for future research.

2. A formal framework of trust. The concept of trust has been a subject of continuous interest in different research areas, including multi-agent systems [20, 6, 4, 7, 2], game theory and economics [18, 11, 32], sociology [9, 19], risk-analysis [14], and psychology [31]. The notion of trust is also closely related to the design and implementation of multi-stage safe exchanges [29, 22].

Trust has different connotations and has been used in different meanings in different contexts by different authors. Many authors [7, 12] consider trust as a belief or cognitive stance that could eventually be quantified by a subjective probability [14, 16]. We give a brief conceptualization of trust that will help avoid confusion and will facilitate further exposition.

We assume that trust is a bilateral relation that involves an entity manifesting trust called the trustor and an entity being trusted called the trustee. Further, we assume that

— There is an event Γ that the trustor cannot control and that depends on the trustee. The trustee may have partial or full control over Γ.

— The trustor voluntarily decides to put himself in a position dependent on Γ in the sense that the trustor will benefit if Γ occurs, otherwise he will lose.

In other words, the trustor depends on the trustee for some event Γ which is controlled by the trustee. Since it is Γ what affects the trustor, we assume that trustworthiness could be measured by the probability of Γ. For example, the trustee could be an untrustworthy seller and $\Gamma = \{$The seller delivers promised merchandize after it has been paid for$\}$. In another example, the trustor could depend on the trustee for some information and $\Gamma = \{$The trustee delivers accurate and truthful information$\}$. Another interpretation is $\Gamma = \{$The quality of the merchandize meets the buyer's expectation$\}$.

In general, two types of trustworthiness can be identified: perceived and actual. Perceived trustworthiness is defined as the trustor's subjective belief in Γ which could be different from the objective trustworthiness, that is, the objective probability of Γ. For example, an agent might delegate a task to another agent, believing that the task can be successfully executed with probability $\hat{\alpha}$, while the actual success rate of the performing agent is α.

Formally, the trustor's utility function can be denoted by:

$$U(\hat{\alpha}, \Gamma(p_1, \ldots, p_n)) \tag{1}$$

where U is the trustor's utility, $p_1, \ldots, p_n$ are parameters describing the event Γ, and $\hat{\alpha}$ is the degree of perceived trustworthiness, i.e., the degree in which Γ is expected to happen.

The event Γ is favorable to the trustor:

$$\frac{\partial U(\hat{\alpha}, \Gamma(p_1, \ldots, p_n))}{\partial \hat{\alpha}} \geq 0$$

That is, the trustor benefits from higher trustworthiness. The case of complete trustworthiness is represented by $\hat{\alpha} = 1$, and vice versa, the trustee is completely untrustworthy when $\hat{\alpha} = 0$:

$$U(1, \Gamma(p_1, \ldots, p_n)) > 0$$

$$U(0, \Gamma(p_1, \ldots, p_n)) < 0$$

If we assume that utility is a continuous function of trustworthiness, then there is a threshold level $\hat{\alpha}_0 \in [0, 1]$ that separates trustworthiness from untrustworthiness:

$$U(\hat{\alpha}, \Gamma(p_1, \ldots, p_n)) \geq 0 \qquad \text{for all} \qquad \hat{\alpha} \geq \hat{\alpha}_0$$

That is, he trustor is always better off if the other agent's trustworthiness exceeds the threshold $\hat{\alpha}_0$ which depends on the event Γ and its parameters $p_1, \ldots, p_n$. This defines a natural *participation constraint*: the trustor will place trust on the trustee (or will voluntarily agree to depend on the trustee) if the the trustee's perceived trustworthiness exceeds $\hat{\alpha}_0$. The participation constraint corresponds to the intuition that an agent will only engage in an interaction if the trustworthiness of the other party exceeds some threshold (the level of acceptable trustworthiness), which depends on the interaction context (through parameters $p_1, \ldots, p_n$) and on the trustor (through the trustor's utility function U). In other words, the threshold $\hat{\alpha}_0$ is both objectively and subjectively determined.

Such a formalization of trust is domain independent and captures a wide range of applications where the trustor believes that the trustee will behave in some expected way specified by the event Γ. Depending on the context the subjective (or objective) trustworthiness can be given different interpretations. For instance, it could be the probability of delivery, the probability of high product quality, probability that an agent will follow contract terms, etc.

By choosing probability $\hat{\alpha}$ (or α) as a measure of trustworthiness we do not mean that trust always depends on a single factor. The event Γ may have a complex structure represented by parameters $p_1, \ldots, p_n$. In another work of ours [17] we experimentally validated a multidimensional model of trust in online exchanges. We showed that the following six factors affect trust: information content, product, transaction, technology, institutions, and consumer-behavior. We assume that all these factors could be combined so as to produce a single measure of an agent's trustworthiness. In other words, we can think of α as a measure of the combined effect of different constituents and determinants of trust.

3. Problem setting. This section describes an auction with untrustworthy bidders. A buyer solicits bids from sellers with two different levels of trustworthiness α and $\beta, \alpha < \beta; \alpha, \beta \in [0, 1]$. Both α and β are normalized measures of a bidder's commitment to back up his bids. For the ease of interpretation, α and β could be thought of as probability of delivery, measure of quality, ability, etc. For example, in one interpretation, a less trustworthy bidder will deliver with probability α if he wins the auction, while a more trustworthy bidder will deliver with probability β. Each bidder knows only his own type (α or β) and the set of possible types is common knowledge among the buyer and the sellers. Throughout the paper, we refer to bidders of type α and β as untrustworthy and trustworthy bidders, respectively, assuming that the variation in trustworthiness is significant enough to make a difference.

The buyer is completely trustworthy and he makes the first move after the auction has been closed. That is, the buyer pays first without knowing the probability of delivery. Since by moving first the buyer explicitly discloses his type, the assumption of complete buyer's trustworthiness does not limit the generality of the model.

Each bid specifies an offer of promised quantity q and price p. The buyer and the sellers are risk-neutral, and the buyer derives utility from a bid, $(p, q) \in \mathbb{R}^2_+$:

$$U(p, q, \theta) = V(q, \theta) - p \tag{2}$$

where θ is the bidder's trustworthiness, $\theta \in \{\alpha, \beta\}$, and $V(q, \theta)$ is the buyer valuation function, $V_q > 0, V_{qq} < 0$, and $V_q(0, p) = 0$ to ensure an interior solution. Subscripts denote partial derivatives.

A bidder, upon winning, earns from a bid (p, q) the following profits:

$$W(p, q, \theta) = p - C(q, \theta) \tag{3}$$

where W and C are the bidder's utility and cost functions, respectively. We assume $C_{qq} > 0, C_\theta > 0$, and $C_{q\theta} > 0$. Thus, both the total and the marginal cost increase with θ. To understand the intuition behind these assumptions, it is convenient to view one's trustworthiness as a measure of quality or probability of delivery. Production costs usually increase with quality, all other things being equal. In addition, trustworthy agents may incur added costs for establishing and keeping a good reputation. In economic literature, [3] trustworthiness is often modelled as a financial asset requiring a certain level of capital investment.

The problem with untrustworthy bidders is that the buyer's utility depends on the trustworthiness of the winner, which is only privately known. In this case, the informed winner's trading decision depends on privately held information in a manner that negatively affects the uninformed auctioneer. Without knowing bidders' types, the auctioneer cannot precisely evaluate the utility of a bid, and therefore determine the auction winner. Since the buyer is moving first, he cannot condition his payment on the seller compliance. In addition, we assume that the buyer does not have access to indirect indicators of a seller's trustworthiness such as reputation database or history of previous interactions. If the buyer moved second, after verifying, directly or indirectly, the delivery, the principal-agent theory could be used [24, 21] to design optimal auction rules.

If the auctioneer uses a scoring function equal to his utility, defined by Equation (3), and asks bidders to reveal their types, then untrustworthy bidders ($\theta = \alpha$) may have un incentive to report a higher type ($\theta = \beta$). The problem is that the scoring function (and the auctioneer's utility) increases in θ. For example, in a standard Vickrey auction, the winner has to match the price and the quality of the second-score bidder. This, however, does not prevent an untrustworthy bidder from reporting higher trustworthiness. Reporting a higher type increases the chance of winning the auction without affecting a bidder's utility.

Proposition 1. *Truthfully declaring an agent's trustworthiness is not a dominant strategy in a standard Vickrey auction, where agents bid on price and quantity.*

Proof. Since agents bid on price and quantity along with declaring their types, every bid is a triple (p, q, θ). Note that the reservation level of untrustworthy bidders lies below the reservation level of trustworthy bidders for fixed quantity and price. Figure 1 shows the zero-utility indifference curves, for trustworthy (I_{trust}) and untrustworthy (I_{untr}) bidders as functions of price and quantity. As long as every untrustworthy bidder submits a bid above the line I_{untr}, he earns nonnegative utility. Since I_{trust} is not below I_{untr}, every bid which is individually rational for a trustworthy bidder is also individually rational for an untrustworthy bidder. For example, bid A earns zero utility for a trustworthy bidder and strictly positive utility, $p_A - p_B$, for an untrustworthy bidder. If the auction winner is an untrustworthy agent, who misrepresented his type, the second-score bid will always be individually rational for him. Therefore, by declaring a higher type, an untrustworthy bidder does incur the risk of receiving negative utility.

Figure 1 also shows the auctioneer's indifference curves for trustworthy (J_{trust}) and untrustworthy (J_{untr}) bidders for bid A. That is, J_{trust} is the auctioneer's indifference curve that passes trough A under the assumption that the bidder is trustworthy. Since J_{trust} is steeper than J_{untr}, the auctioneer receives greater utility if bid A is submitted by a trustworthy bidder. Therefore, by declaring higher trustworthiness, a bidder can increase his chances of winning the auction with no risk of receiving negative payoff.

Note that bidders reveal their type by submitting bids between lines J_{trust}

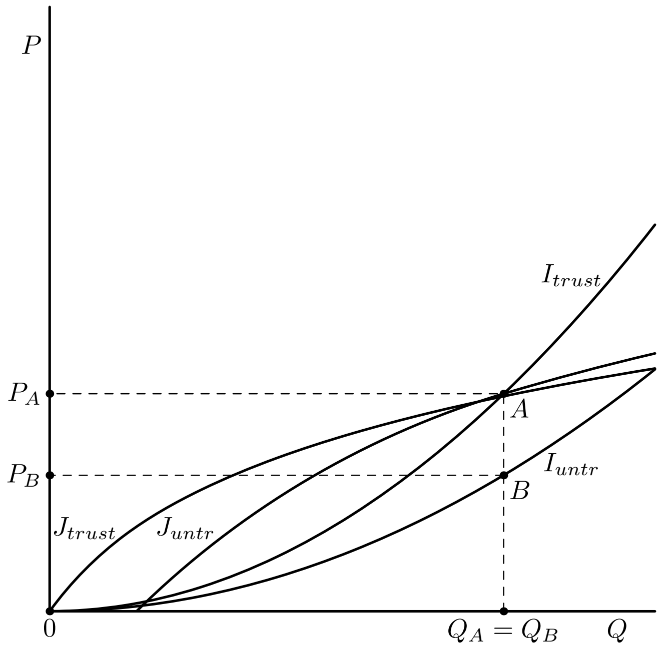

Fig. 1. Zero-utility indiffernce curves for trustworthy and untrustworthy bidders

and J_{untr}. Such bids are not individually rational for trustworthy bidders, and therefore can be submitted only by untrustworthy agents. □

4. A separating auction. In this section we study the problem of how to perform bid evaluation and winner determination based solely on information contained in bids. The problem with untrustworthy bidders is that, by declaring higher trustworthiness, they can manipulate the way bids are evaluated. We discuss discriminating auction rules that separate trustworthy from untrustworthy agents.

A natural way to approach the problem with untrustworthy bidders is to assume that the auctioneer adopts a play-safe strategy and decides to insure himself against the worst case possible.

Definition 1. *In a **distrust-based auction**, every bidder submits a bid on price and quality. The auctioneer uses a scoring function S that treats each bidder as untrustworthy:*

$$S(p, q) = V(q, \alpha) - p$$

Unfortunately, being overcautious does not help auctioneers to avoid untrustworthy bidders.

Proposition 2. *There is a strictly positive probability that an untrustworthy bidder wins in a distrust-based Dutch, English, first-score, and a second-score sealed bid auctions. If the difference in trustworthiness, $\beta - \alpha$, between the two agent types is sufficiently large, then only untrustworthy bidders win.*

Proof. A trustworthy bidder can get a maximal score by submitting a bid $b_0 = (C(q_0, \beta), q_0)$, where

$$q_0 = \underset{q}{\operatorname{argmax}} \left(V(q, \alpha) - C(q, \beta)\right)$$

Apparently,

$$\begin{aligned} U(b_0, \alpha) &= V(q_0, \alpha) - C(q_0, \beta)) \\ W(b_0, \beta) &= 0 \end{aligned}$$

That is, a trustworthy seller receives zero utility from b_0, while the auctioneer receives $V(q_0, \alpha) - C(q_0, \beta)$. On the other hand, an untrustworthy bidder receives a strictly positive utility from b_0:

$$W(b_0, \alpha) = C(q_0, \beta) - C(q_0, \alpha) > 0$$

Therefore, an untrustworthy bidder can always underbid a trustworthy bidder by sending $b_1 = (C(q_0, \beta) - \epsilon, q_0)$. Note that biding below b_0 is not individually rational for a trustworthy bidder. If the two bidder types differ significantly, $\beta \gg \alpha$, then the equilibrium in which a trustworthy bidder bids b_0 and an untrustworthy bidder bids below b_0 is a Nash equilibrium in pure strategies for the Dutch, first-score and second-score sealed bid auctions auctions. An untrustworthy bidder wins in a English auction, as well.

If the difference between levels of trustworthiness is small, than both types of agents will adopt mixed strategies depending on the auction rules. Hence, there is always a strictly positive probability that an untrustworthy seller wins. □

Proposition 2 can be explained using the cost differences between agent types. If the difference in trustworthiness is sufficiently large, trustworthy bidders incur sufficiently large costs compared to untrustworthy bidders, which prevents them from submitting competitive bids, and therefore from winning an auction.

Another way to solve the problem with untrustworthy bidders is to consider trustworthiness as a random variable and to evaluate bids using its expectation, $E(\theta)$. Unfortunately, a similar proposition here as well.

Proposition 3. *Suppose that an auctioneer evaluates bids according to his expectation of agents' trustworthiness:*

$$S(p, q) = V(q, E(\theta)) - p$$

There is a strictly positive probability that an untrustworthy bidder wins in Dutch, English, first-score, and second-score sealed bid auctions. If the difference in trustworthiness, $\beta - \alpha$, between the two agent types is sufficiently large, then only untrustworthy bidders win.

Proof. Parallels the proof of Proposition 2 after substituting $E(\theta)$ for α in bid evaluation. □

Propositions 2–3 show that, in some cases, trustworthy agents will be driven out of the market, thereby causing a market inefficiency. To fix the problem, we investigate constrained-bidding mechanisms.

Definition 2. *In a **constrained-bidding** multidimensional auction, an eligible bid satisfies a set of constraints on bid parameters. That is, for every eligible bid $b(t_1, \ldots, t_n)$ we have*

$$\phi_k(t_1, \ldots, t_n) \text{ for } k = 1, \ldots, m$$

where $\{\phi_k\}_{k=1}^{m}$ is a set of constraint predicates.

For example, the auction rules can fix the quantity to q_0 and define a minimal and a maximal price:

$$q = q_0, \text{ and } p \in [p_{\min}, p_{\max}] \tag{4}$$

One possible interpretation is that the maximal price is the auctioneer's reservation level, and the minimal price is the reservation level for a bidder of a certain type. In our setting, constraints (4) reduce a two-dimensional auction on price and quantity to a unidimensional auction on price.

One important characteristic of constrained auctions is that the bidders' expected utility can be ex-ante limited by the auction rules. For example, constraints (4) impose an upper bound, $p_{\max} - C(q_0, \theta)$, and a lower bound $p_{\min} - C(q_0, \theta)$ for type θ bidders. By choosing a particular set of constraints, the auctioneer can affect the incentive structure of the auction, and therefore can provide bidders with additional incentives. We will show that in our case, the bidders could be given incentives to reveal directly or indirectly their type.

We assume that, if a seller faces a choice between two auctions, he will choose an auction which gives him a better utility range, all other things being equal. For example, if a seller must choose between an auction A_1 with a utility range [2, 10] and a auction A_2 with a utility range [0, 8], he would choose A_1, all other things being equal. The intuition behind this assumption is that every bidding strategy for a auction A_2 gives a better expected utility when applied to auction A_1.

Assumption 1. *Given an auction A_1 with a utility range $[a^1_{\min}, a^1_{\max}]$ and an auction A_2 with a utility range $[a^2_{\min}, a^2_{\max}]$, where the only difference between A_1 and A_2 is*

$$\begin{aligned} a^1_{\min} &> a^2_{\min} \\ a^1_{\max} &> a^2_{\max} \\ a^1_{\max} - a^1_{\min} &= a^2_{\max} - a^2_{\min} \end{aligned}$$

then a risk-neutral bidder prefers auction A_1 to auction A_2.

In other words, in both auctions, a bidder has the same strategy set, faces the same opponents and the same rules, with the only difference being strategy payoffs. In Assumption 1, both auctions have the same length of utility range. If utility ranges differ, then a risk-neutral bidder can still prefer auction A_1 if the lower utility bound of A_1 exceeds the upper utility bound of A_2:

$$a^1_{\min} \geq a^2_{\max}$$

The intuition is that every bidding strategy in A_1 yields a greater payoff than a bidding strategy in A_2, all other things being equal. In the next section, we will drop Assumption 1 and propose a generalized Vickrey auction in which truth-telling is a dominant strategy.

Using bidders' preferences for auctions, the auctioneer can distinguish, or screen, various types of bidders by offering different bid constraints to different types of bidders.

Definition 3. *In a **separating constrained-bidding** auction, the auctioneer offers two sets of bid constraints. A bidder chooses a set of constraints and strictly follows this set throughout the auction. All other auction rules remain the same for all bidders. A bidder is not allowed to change his set of constraints during an auction.*

In other words, there are two bidding schedules, each bidder chooses and follows only one schedule, and all bidders compete with one another. That is, each bidder competes with both the bidders from his schedule and the bidders from the other schedule. For example, in a separating constrained-bidding auction based on the first-score rule the bidder with the highest score wins. In the beginning, the auctioneer offers two sets of bid constraints. A bidder either chooses a set of constraints and submits a bid (or bids) satisfying only this set, or rejects the auction.

According to the next proposition, sometimes it is possible to design two sets of bid constraints so that all trustworthy bidders prefer one set and all untrustworthy bidders prefer the other. Thus, by choosing a set of constraints, bidders disclose their type. This allows the auctioneer to evaluate the utility of each bid and to determine the winner. Since the auctioneer knows the bidders' types, he can associate every trustworthy bid with β and every untrustworthy bid with α.

Proposition 4. *If $V_q(0,\alpha) > C_q(0,\beta)$, then there exists a constrained-bidding auction that separates trustworthy from untrustworthy bidders.*

Proof. Figure 2 shows the zero-utility indifference curves for the bidders and the auctioneer. Lines J_{trust} and J_{untr} represent the auctioneer's indifference curves for trustworthy and untrustworthy bidders, respectively. Similarly, lines I_{trust} and I_{untr} represent the indifference curves for trustworthy and untrustworthy bidders.

Point C is defined as the intersection of I_{trust} and J_{untr}. Since $V_q(0,\alpha) > C_q(0,\beta)$, A is well defined and $A \neq (0,0)$.

q_A is chosen as the quantity satisfying:

$$V(q_A,\beta) - C(q_A,\beta) = V(q_C,\alpha) - C(q_C,\alpha)$$

Point B has coordinates $(q_A, C(a_A,\beta))$, and lies on the curve I_{trust}. Point D is defined as $(q_C, C(q_C,\alpha) + C(a_A,\beta) - C(q_A,\alpha))$. The definition of the points A, B, C and D implies

$$W(A,\alpha) = W(C,\alpha) \tag{5}$$

$$W(B,\alpha) = W(D,\alpha) \tag{6}$$

That is, A and C lie on the same indifference curve for an untrustworthy bidder, and B and D lie on another indifference curve. That is, an untrustworthy bidder is indifferent between selling quantity q_A at price p_A and quantity q_C at price p_C. Similarly, he is indifferent between quantity q_C at price p_C, and quantity q_D at price p_D. Consider the sets of bid constraints C_{AB} and C_{CD} defined as follows:

$$C_{AB} = \{q = q_A, \text{ and } p \in [p_B, p_A]\}$$

$$C_{CD} = \{q = q_C, \text{ and } p \in [p_D, p_C]\}$$

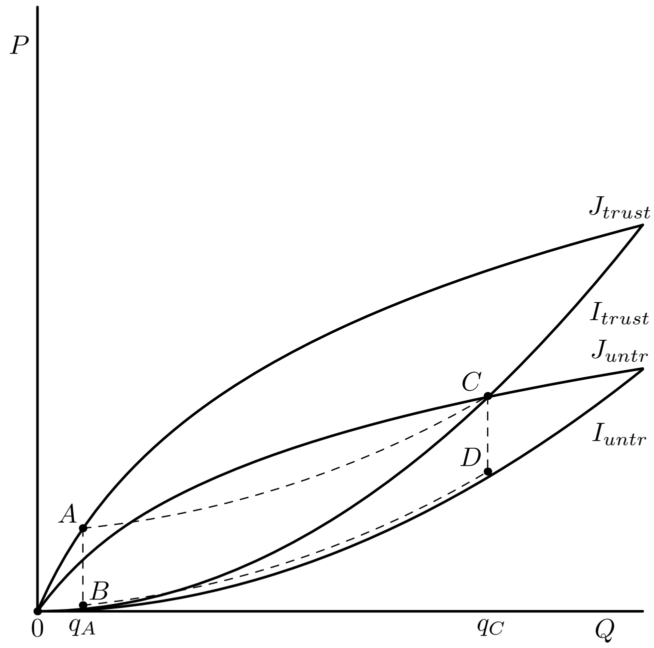

Fig. 2. A separating set of bid constraints

If a bidder chooses to submit bids satisfying C_{AB}, then he has to bid a fixed quantity, q_A, the maximal price is p_A, and the minimal price is p_B.

Equations (5) and (6) imply that an untrustworthy bidder is indifferent between two sets of constraints, C_{AB} and C_{CD} (under Assumption 1). On the other hand, Figure 2 shows that C_{CD} offers non-positive utility for a trustworthy bidder (line CD is below the zero-utility indifference curve I_{trust}), while C_{AB} offers only non-negative utility. Therefore, a trustworthy bidder prefers C_{AB} to C_{CD}.

To make the untrustworthy bidder's preferences strong, the line AB can be shifted to the left by an infinitesimal amount ϵ. This will make an untrustworthy bidder strongly prefer bidding on line CD without changing the preferences for a trustworthy bidder.

Note that the auctioneer is indifferent between C_{AB} and C_{CD}:

$$U(A, \beta) = U(C, \alpha) = 0$$

$$U(B, \beta) = U(D, \alpha) = p_C - p_D = p_A - p_B$$

That is, each set of constraints offers the same utility range, $[0, p_C - p_D]$, to the auctioneer and to all types of bidders.

The bidding schedule AB looks the same to trustworthy bidders as does schedule CD to untrustworthy bidders. The only difference between the schedules is that one of them is parallely shifted to compensate for cost differences between agents' types. Both types of bidders bid in the same price range (since $p_C - p_D = p_A - p_B$) and they receive the same utility from prices which are equally distanced from the minimal ones. In other words, each bidding schedule is equivalent to a single-unit auction with an auctioneer's reservation price $p_C - p_D$ and a bidder's reservation price 0.

Therefore, when presented with a choice between bidding schedules C_{AB} and C_{CD}, trustworthy bidders chose C_{AB}, and untrustworthy bidders chose C_{CD}. $\square$

According to Proposition 4, auction rules can be designed to eliminate the strategic consequences arising from differences in bidders' types. In such cases, the auctioneer can offer two bidding schedules so that the first type bidders choose the first schedule and the second type choose the second schedule. Both types of bidders face the same reservation utility and the same strategic choices.

Another observation is that, in order to separate trustworthy from untrustworthy bidders, the auctioneer must split a two-dimensional auction (on price and quantity) into two inidimensional (on price only) bidding schedules.

It should be pointed out that a separating auction does not prevent untrustworthy bidders from winning. What distinguishes a separating auction from distrust-based and expectation-based auctions is that the auctioneer can exactly evaluate bids and choose the most profitable bid. In addition, when the difference in trustworthiness, $\beta - \alpha$, is sufficiently large, trustworthy agents are not driven out of the market, as is the case for the other auctions.

A separating auction can have a variety of auction rules. It could be, for instance, a second-price sealed-bid auction where a bidder chooses between two predefined quantities (either q_A or q_C in Figure 2) and submits a price for that quantity. According to Proposition 4, trustworthy bidders choose q_A and bid p_B, while untrustworthy bidders choose q_C and bid p_D. In both cases, the auctioneer receives utility $p_A - p_B$.

It should be pointed out that a separating auction may not maximize the social welfare. Obviously, some price has to be paid for the possibility to separate agent types.

For example, in order to maximize his utility in a second-score auction, the auctioneer will choose bidding schedules with maximal utility range, $p_A - p_B$. That is, the auctioneer will choose quantity q''_C such that:

$$q''_C = \underset{q}{\operatorname{argmax}} \left(V(q, \alpha) - C(q, \alpha) \right)$$

If the auctioneer knew the type of each bidder, then he could fix the quantity to

$$q'_C = \underset{q}{\operatorname{argmax}} \left(V(q,\beta) - C(q,\beta)\right)$$

or to

$$q''_C = \underset{q}{\operatorname{argmax}} \left(V(q,\alpha) - C(q,\alpha)\right)$$

depending on which agent type is more profitable for him. It is apparent, that in the case where trustworthy agents offer more utility to the auctioneer, the social welfare is not maximized. If, however, untrustworthy agents are more efficient, then a separating auction is socially optimal. Whether trustworthy agents are more efficient than untrustworthy ones, depends on the value, $V(q,\theta)$, and the cost function, $C(q,\theta)$. If the social cost of trustworthiness is less than its social value, then trustworthy agents will be more efficient, and vice versa.

5. A generalization of the Vickrey auction. In this section we describe a generalization of the Vickrey auction to the case of untrustworthy bidders. We drop Assumption 1 and the restriction of having only two types of bidders. The generalized auction is applicable to situations with a continuum of bidder types.

In the generalized auction each bidder submits a bid on price, quantity, and a declaration of trustworthiness $(p, q, \hat{\theta})$. The auction uses a constrained-bidding schedule where each bidder is required to submit the maximal price for each combination of quantity and price:

$$p = C(q, \hat{\theta}) \tag{7}$$

We assume that the cost function is known to the auctioneer who can then check Condition (7) for each bid and verify its validity. The score function is equal to the auctioneer's utility, assuming that every bidder truthfully declares his type, i.e., $\hat{\theta} = \theta$. The winner is the bidder with the highest score (ties are resolved randomly). The winning bidder matches the highest rejected score by choosing a price and a quantity, which generate the same score. The central point of the auction rules is that, in matching the second-highest score, the winner is assumed to have the same type as the highest-rejected bidder. In other words, the winner is allowed to choose a price and a quantity that generate the highest-rejected score using the declared trustworthiness of the highest-rejected bidder. More formally:

Definition 4. *In the generalized Vickrey auction each bidder submits a bid* $b = (p, q, \hat{\theta})$. *Bidding is constrained and eligible bids must satisfy Equation* (7). *The score is defined as:*

$$S(p, q, \hat{\theta}) = V(q, \hat{\theta}) - p$$

The highest score wins. The price and quantity are chosen by the winner to satisfy:

$$p = V(q, \hat{\theta}_s) - V(q_s, \hat{\theta}_s) + p_s \tag{8}$$

$$V_q(q, \hat{\theta}_s) - C_q(q, \hat{\theta}) = 0 \tag{9}$$

where $(p_s, q_s, \hat{\theta}_s)$ *is the second-highest bid. As usual, subscripts denote partial derivatives.*

Condition (9) guarantees that the winner matches the second-highest score, while Equation (10) ensures that the marginal cost of the winner is equal to the marginal value which the auctioneer could have received from the second-highest bidder. Note that Condition (9) requires the winner to match the second-highest score under the assumption that he has the type of the second-highest bidder.

Proposition 5. *In the generalized Vickrey auction, it is a dominant strategy for each bidder to truthfully report his trustworthiness.*

Proof. The utility of the winner is given by:

$$W(p, q, \theta) = p - C(q, \theta)$$

where p and q are the price and quantity satisfying Equations (8)–(9). From Equation (8) it follows that:

$$W(p, q, \theta) = V(q, \hat{\theta}_s) - V(q_s, \hat{\theta}_s) + p_s - C(q, \theta)$$

The winner will choose quantity q to maximize his utility. Therefore,

$$V_q(q, \hat{\theta}_s) - C_q(q, \theta) = 0$$

On the other hand, Equation (9) requires:

$$V_q(q, \hat{\theta}_s) - C_q(q, \hat{\theta}) = 0$$

Therefore the winner will maximize his utility if and only if $\theta = \hat{\theta}$. □

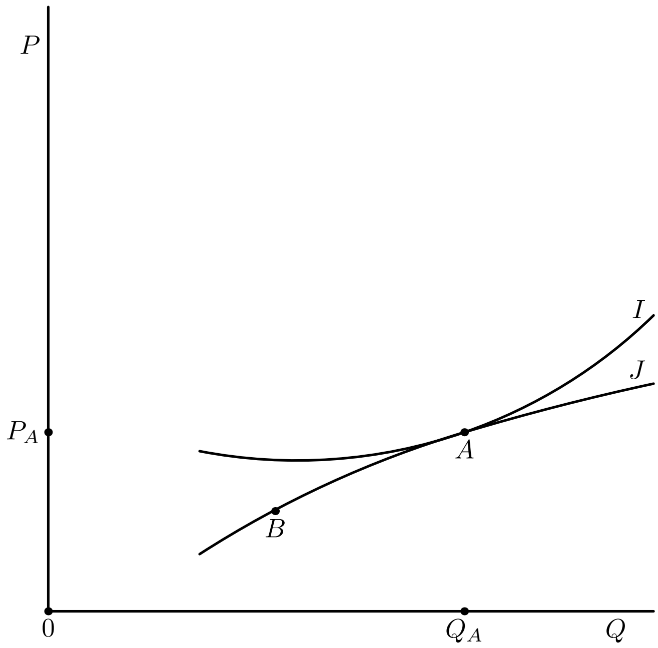

Fig. 3. Indifference curves for a bidder and an auctionner

The simple intuition behind the generalized Vickrey auction is presented in Figure 3. The winner is allowed to pick a quantity and a price that match the second-highest score, under the assumption that the winner has the type of the second-highest bidder. Point B in Figure 3 represents the second-highest bid. J is the indifference curve of the auctioneer for the type declared by the second-highest bidder. That is, the winner is allowed to move along the curve in choosing quantity and price. Clearly, the winner's profit is maximized at the tangency point between his indifference curve I and the auctioneer's indifference curve J. The angle at the tangency point is uniquely determined by the winner's type θ. Equation (9) requires the winner to pick a tangency in accordance with his declaration $\hat{\theta}$. Therefore, the winner maximizes his utility if and only if he truthfully declares his type. Note that a truthful declaration neither increases nor decreases an agent's chances of winning the auction. Since agents submit multidimensional bids, they can always make tradeoffs between bid parameters.

6. Conclusions. In the paper we have analyzed a multidimensional auction in which a trustworthy buyer faces sellers with different degrees of trustworthiness. We proposed two mechanisms that make bidders directly or indirectly reveal their trustworthiness. The first mechanism is based on discriminating bidding schedules. We have proved that, under certain conditions, it is possible to design bidding schedules that separate trustworthy from untrustworthy bidders.

The second mechanism is a generalization of the Vickrey auction to the case of untrustworthy bidders. We proved that, if the winner is considered to have the trustworthiness of the second-highest bidder, truthfully declaring one's trustworthiness becomes a dominant strategy.

We expect the proposed mechanisms to reduce the cost of trust management and to eliminate some market failures and inefficiencies caused by lack of trust. By eliminating the need to manipulate and speculate about other bidder' trustworthiness, the mechanisms could also simplify the architecture of economic software agents. In risky environments, the mechanisms could enable mutually beneficial interactions which are otherwise costly to enforce or cannot be enforced by third parties.

REFERENCES

[1] Bikhchandani S., J. Ostroy. The package assignment model. UCLA, working Paper Series, mimeo, 2001.

[2] Biswas A., S. Sen, S. Debnath. , 2000. Limiting deception in groups of social agents. *Applied Artificial Intelligence* **14** (2000) No 8, 785–797.

[3] Boot A., S. Greenbaum, V. Thakor. Reputation and discretion in financial contracting. *American Economic Review* **83** (1993) 1165–1183.

[4] Brainov S. An incentive compatible trading mechanism for trust revelation. In: Proceedings of the IJCAI'01 Workshop on Economic Agents, Models and Mechanisms, 2001, 62–70.

[5] Branco F. The design of multidimensional auctions. *Rand Journal of Economics* **28** (1997) No 1, 63–81.

[6] Braynov S., T. Sandholm. Contracting with uncertain level of trust. *Computational Intelligence* **14** 2002, No 4, 501–514.

[7] Castelfranchi C., R. Falcone. The dynamics of trust: From beliefs to action. In: Proceedings of the Second Workshop on Deception, Fraud and Trustin Agent Societies. Seattle, 1999, 41–54.

[8] Che Y. Design competition trough multidimensional auctions. Rand *Journal of Economics* **24** (1993) No 4, 668–680.

[9] Coleman J. Foundations of Social Theory. Harvard University Press, 1990.

[10] Conen W., Sandholm. Partial-revelation vcg mechanism for combinatorial auctions. In: Proceedings of the National Conference on Artificial Intelligence (AAAI), 2002.

[11] Dasgupta P. Trust as a commodity. In: Making and Breaking Cooperative Relations (Ed. D. Gambeta), Basil Blackwell, 1990.

[12] Dimitrakos T. Towards a formal model of trust in e-commerce. In: Proceedings of the The AI-2001 Workshop on Novel E-Commerce Applications of Agents. Ottawa, Canada, 2001.

[13] Fujishima Y., K. Leyton-Brown, Y. Shoham. Taming the computational complexity of combinatorial auctions: Optimal and approximate approaches. In: Proceedings of the International Joint Conference on Artificial Intelligence, 1999, 548-553.

[14] Garrik S. K. B. On the quantitative definition of risk. *Risk Analysis* **28** (1981), 11–27.

[15] Hoos H., C. Boutilier. The package assignment model. In: Proceedings of the International Joint Conference on Artificial Intelligence. Seattle, 2001, 1211-1217.

[16] Josang A. Trust-based decision making for electronic transactions. In: Proceedings of the Fourth Nordic Workshop on Secure Computer Systems(NORDSEC'99), Sweden, 1999, 496–502.

[17] Kim D., Y. Song, S. Braynov, H. Rao. A b-to-c trust model for on-line exchange. In: Americas Conference on Information Systems, 2001, 784–787.

[18] Kreps D., R. Wilson. Reputation and imperfect information. *Journal of Economic Theory* **27** (1982), 253–279.

[19] Luhmann N. Trust and Power. John Wiley and Sons, 1979.

[20] Marsh S. Formalizing trust as a computational concept. Ph.D. thesis, University of Stirling, U.K., 1994.

[21] Maskin E., J. Riley. Optimal auctions with risk-averse buyers. *Econometrica* **52** (1984) 1473–1518.

[22] Defection-free exchange mechanism for information goods. In: Proceedings of the Fourth International Conference on MultiAgent Systems. Boston, 2000, 183–190.

[23] McAfee R., J. McMillan. Competition for agency contracts. *RAND Journal of Economics* **18** (1987), No 2, 296–307.

[24] Myerson R. Optimal coordination mechanisms in generalized principle-agent problems. *Journal of Mathematical Economics* **10** (1982), 67–81.

[25] Myerson R. Mechanism design by an informed principle. *Econometrica* **51** (1983), 1767–1797.

[26] Nisan N. Bidding and allocation in combinatorial auctions. In: Proceedings of ACM Conference on Electronic Commerce. 2000, 1-12.

[27] Rothkopf M., A. Peckec, R. Harstad. Computationally manageable combinatorial auctions. *Management Science* **44** (1998) No 8, 1131-1147.

[28] Sandholm T. Algorithm for optimal winner determination in combinatorial auctions. Artificial Intelligence **135** (2002), 1-54.

[29] Sandholm T., V. Ferrandon. Safe exchange planner. In: Proceedings of the Fourth International Conference on MultiAgent Systems, Boston, 2000, 255–262.

[30] Sandholm T., S. Suri, A. Gilpin, D. Levine. Cabob: A fast optimal algorithm for combinatorial auctions. In: Proceedings of the International Joint Conference on Artificial Intelligence, Seattle, 2001, 1102-1108.

[31] Slovic P. Risk perception and trust. In: Trust: Fundamentals of Risk Analysis and Risk Management. (Ed. V. Molak), Lewis Publishers, 1997.

[32] Snijders C. Trust and Commitments, ICS, 1996.

[33] Wurman P., M. Wellman. Akba: A progressive, anonymous-price combinatorial auction. In: Proceedings of ACM Conference on Electronic Commerce, 2000, 21-29.

Sviatoslav Braynov
University of Illinois at Springfield
USA
e-mail: sbray2@uis.edu

Radoslav Pavlov
Institute of Mathematics and Informatics
Bulgaria
e-mail:radko@cc.bas.bg

Received February 15, 2007
Final Accepted September 19, 2007