

Bribing in second-price auctions

Shiran Rachmilevitch*

October 3, 2011

Abstract

An IPV 2-bidder second-price auction is preceded by two rounds of bribing: prior to the auction each bidder can try to bribe his rival to depart from the auction, so that he (the briber) will become the sole participant and obtain the good for the reserve price. Bribes are offered sequentially according to an exogenously given order—there is a first mover and a second mover. I characterize the unique efficient collusive equilibrium in monotonic strategies; in it, the second mover extracts the entire collusive gain. This equilibrium remains an equilibrium even when valuations are interdependent, and if they are separable then the full surplus extraction result continues to hold. Additionally, a family of pooling equilibria is studied, in which all the types of the first mover offer the same bribe.

Keywords: Second-price auctions, collusion, bribing, signaling, surplus extraction. *JEL Codes:* D44, D82.

*Department of Economics, University of Haifa.

1 Introduction

Collusion among participants in auctions is a serious and well-documented problem.¹ In the simplest collusive scenario, the cartel members meet prior to the auction in order to decide on side-payments and on a representative bidder, who will bid in the auction on behalf of the cartel. The very first question an economist needs to address when modeling such scenarios is: *what is it that the cartel members do before the auction? What is the pre-auction interaction?*

In the existing literature, it is typically assumed that the cartel members either play a revelation game or a “knockout auction.”² In a revelation game, each member reports a valuation, and the collusive agreement is determined as a function of the profile of reports (and possibly further information, such as the behavior of non-cartel-members). In a “knockout auction,” the cartel members run an auction among themselves, for the right to participate in the real auction.

Each of these games is a *one-shot* game. In reality, however, collusive situations are a special kind of *bargaining situations* (once a surplus is extracted from the seller, there is the question how to divide the spoils), and the latter, by their nature, are *sequential*. The colluding parties go through a “face-to-face” negotiation process, that leads to a collusive agreement (or disagreement). During such a negotiation phase there is signaling (the players’ moves are indirect signals of their private information), which gives rise to an adverse selection problem. Static models, therefore, miss an important aspect

¹See, for example, Baldwin et al. (1997), Cassady (1967), and Porter and Zona (1993).

²Leading examples include Graham and Marshall (1987), Mailath and Zemsky (1991), Marshall and Marx (2007), and McAfee and McMillan (1992).

of the strategic situation; this shortcoming leads me to study the following game.

I consider an IPV 2-bidder second-price auction which is preceded by two rounds of bribing; each bidder, in his turn, can offer his rival a bribe in exchange for the latter's departure from the auction. Bribes are offered sequentially according to an exogenously given order, and if both offers are rejected then the pre-auction phase ends, and both bidders go on to compete against one another in the auction. This game is an extension of the “take-it-or-leave-it” (TIOLI) game of Esö and Schummer (2004, henceforth ES), which consists of a second-price auction and a single pre-auction round, in which a designated player has the opportunity to offer a bribe to his rival in exchange for the latter to depart, and if it is rejected then the pre-auction phase ends. That is, the present model is obtained by adding one round of bribing to that of ES.

My goal is to address the following questions in a setting that captures the aforementioned sequential signaling: *(i)* is efficient collusion possible, and if so, what collusive strategies lead to efficient allocation, *(ii)* what is the intercartel distribution of gains, and *(iii)* what are the signaling properties of the equilibrium (i.e., pooling versus signaling). The TIOLI game is a natural first step for addressing these issues, but signaling in this game is rather limited, because, effectively, only the briber can signal his type.³ Additionally, it is highly asymmetric; many real-life situations, on the other hand, lack such

³Given an offer from the briber, low respondent types accept it and high respondent types reject it, but in either case this (coarse) signaling by the respondent is irrelevant, since either reaction effectively ends the game (a continuation game that follows a bribe's rejection is nontrivial—it is a noncooperative second-price auction—but the signal on the respondent's type is irrelevant, because it is a weakly dominant strategy for the briber to bid his valuation, no matter what he thinks about the respondent).

asymmetry. A setting where both parties play similar (though not identical) roles is therefore more suited for modeling such situations. Contrasting the current analysis with that of ES will clarify the differences between the case where one side has all the bargaining power and the case where both parties can initiate collusion.

1.1 Summary of the results

In a bribing game, the amount a briber offers depends on his valuation (type); his behavior when he bribes is summarized by a *bribing function*, defined on his type-space. Considering the TIOLI game, ES derived (under some regularity conditions on the type distribution) the unique bribery-involving equilibrium in which the bribing function is continuous. In this equilibrium, inefficiency results with a positive probability, because all the types of the briber above a certain threshold offer the same bribe, which is accepted by all the types of the respondent. The reason for this “nonseparation at the top” is that it is enough for a briber to signal that he is “sufficiently strong” in order to make sure that his bribe is accepted.

In contrast to ES, I show that with two rounds of bribing, efficiency can be achieved. Specifically, under the assumption that the type distribution of the first mover is locally convex at the minimal type, there exists an efficient equilibrium in monotonic strategies if and only if the expectation of the second mover’s type is at least as large as one half of the maximal valuation. That is, with the first and second mover being player 1 and player 2 respectively, and with i ’s valuation denoted by θ_i , the condition is $\mathbb{E}(\theta_2) \geq \frac{1}{2}$.^{4,5} When this

⁴Valuations are drawn from the unit interval. This is just a normalization.

⁵ $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ is sufficient and necessary (for an equilibrium with all the above-mentioned properties) provided that player 1’s type distribution, F_1 , satisfies $F_1''(0) > 0$. Otherwise, it

condition is met (and the distribution of the first mover's type is locally convex at the minimal type), behavior in a monotonic efficient equilibrium is unique;⁶ it is described as follows.

First, player 1 offers the difference between his valuation and his expected noncooperative payoff. More precisely, with the expected payoff in the noncooperative (dominant strategy) equilibrium of the second-price auction of type θ_i of player i denoted by $\pi_i^*(\theta_i)$, player 1's bribe is given by $b_1(\theta_1) = \theta_1 - \pi_1^*(\theta_1)$. Since b_1 is strictly increasing, player 1 reveals his type perfectly on the path. Seeing player 1's revealed type, player 2 employs an efficient acceptance rule; then, in case he rejects θ_1 's offer, he responds with the counteroffer $\pi_1^*(\theta_1)$, which player 1 accepts. In such an equilibrium, player 1's *ex post* equilibrium payoff equals his *expected* noncooperative payoff. The reason for the full surplus extraction is player 2's positional advantage: he moves second, does not reveal any private information, and learns player 1's private information before making his move. This positional advantage translates to an expected payoff of $\pi_2^*(\theta_2) + C$, where $C = \int_0^1 [\theta_1 - \pi_1^*(\theta_1)] f_1(\theta_1) d\theta_1$ is the expected surplus extracted from player 1.

This result is robust in the following sense: even if the IPV assumption is relaxed, and valuations are allowed to be interdependent, the game admits a fully-revealing efficient equilibrium which is analogous to the above mentioned one. Moreover, if the interdependent valuations satisfy a certain separability condition, then the full surplus extraction result continues to hold: the second mover extract the entire gain and the first mover's payoff equals his noncooperative payoff.

is only a sufficient condition.

⁶That is, there exist, strictly speaking, multiple efficient monotonic equilibria, but they differ only in off-path information sets.

I also describe a family of pooling equilibria. In these equilibria, all the types of player 1 make the same strictly positive offer, to which player 2 responds with a simple threshold strategy: he accepts the offer if and only if his type is below a certain threshold; otherwise, he rejects it and makes an offer of his own; moreover, all the rejecting types of player 2 counter with the same offer. In such equilibria, low types of player 1 offer a bribe which exceeds their valuation, because there is a strictly positive probability that it will be rejected and trigger a generous counteroffer. Thus, these equilibria can be thought of as representing a “bluffing” phenomenon: low first-mover types hide their identity by mimicking high types, hoping that player 2 will “take the bait” and try to eliminate player 1 in exchange for a positive bribe.

Finally, I show that the following is an equilibrium provided that the types distributions are log-concave: all the types of the first mover do not offer a bribe, and the ES equilibrium is played starting at the second round.

1.2 Related literature

This paper contributes to a growing body of literature on collusion in one-shot auctions.⁷ Most of this literature takes the mechanism-design approach to collusion and studies direct revelation mechanisms which, in many cases, the cartel operates with the help of an incentiveless third party. Seminal contributions to this literature include Graham and Marshall (1987), Mailath and Zemsky (1991), Marshall and Marx (2007), and McAfee and McMillan (1992). These papers all take the standpoint of the cartel and seek to design mecha-

⁷A related branch of research considers collusion in a repeated-game setting. Contributors to this literature include Aoyagi (2002, 2007), Athey and Bagwell (2001, 2008), Blume and Heidhues (2006), Hörner and Jamison (2007), Rachmilevitch (2009), and Skrzypacz and Hopenhayn (2004).

nisms that are desirable for the bidders.

The other side of the “mechanism-design coin” is to take the standpoint of the seller, and look for auction formats that are immune to collusion. Che and Kim (2009, henceforth CK) take this approach, and derive a collusion-proof auction. It is important to note that the work in CK does not invalidate the contribution of the current paper. First, the current paper assumes that the seller is not strategic and that he employs a standard auction format (second-price). CK considers a strategic seller who employs a nonstandard format. More importantly, the focus of the current paper is on pre-auction signaling among the bidders, an aspect which is absent from CK, since it models collusion as a one-shot signaling-free stage.⁸

From the mechanism-design literature, the setting which is closest to the one considered here is that of the informed-principal problem, which originated in the pioneering work of Myerson (1983). For example, Maskin and Tirole (1990, henceforth MT) studied the following extensive-form game. The principal (first mover) offers a mechanism to the agent (second mover), who updates his belief about the principal’s type once having seeing his offer. Then, the agent either accepts or rejects the mechanism, and the principal updates his belief about the agent’s type upon seeing his response. If the mechanism is accepted, then the players play the game it specifies, while if it is rejected, the game ends, in which case they receive their reservation utilities.⁹ The MT game resembles my game in that the first mover’s proposal conveys (potentially, at least) information about his type, and the second mover’s response

⁸Dequiedt (2007) and Pavlov (2008) also study collusion-proof auctions. In both of these papers, as in CK, there is no signaling among the bidders at the pre-auction stage.

⁹Similar extensive forms have also been utilized in more recent studies of the informed principal problem, e.g., Severinov (2008).

conveys (again, potentially) information about his type.

Finally, two works that consider ES-like models (i.e., an auction which is preceded by a TIOLI stage) are Chen and Tauman (2006) and Kivetz and Tauman (2010). The former considers a second-price auction in an environment where, in addition to the cartel members, there is a random population from which the cartel members can hire shill bidders; the latter considers a first-price auction where the bidders' valuations are commonly known among the bidders.

1.3 Organization

Section 2 lays down the model. Section 3, which is the main body of the paper, considers efficient collusive equilibria. Section 4 studies a family of equilibria with complete pooling, Section 5 shows that the ES equilibrium can be embedded as an equilibrium in the two-round game, Section 6 deals with interdependent valuations, Section 7 concludes, and the appendices collect proofs and technical details.

2 Model

There are two risk-neutral expected-utility-maximizing players, player 1 and player 2, who are about to attend a second-price auction for a single indivisible good. Player i 's valuation for the good (his type) is an independent draw from F_i , a full-support distribution on $[0, 1]$ with a strictly positive and differentiable density f_i .

The reserve price is zero.¹⁰ The auctioneer, who owns the good initially, has

¹⁰This assumption can be relaxed; see Remark 6 below.

no value for it (i.e., his valuation is zero). He awards the good at random, with probability $\frac{1}{2}$ to each player, in case the bids are tied.¹¹ Each player’s utility from nonparticipation (the outside option) is zero. Thus, given the bids (b_1, b_2) , the associated payoff for type θ_i of player i is $\mathbf{1}_{\{b_i > b_j\}}(\theta_i - b_j) + \mathbf{1}_{\{b_i = b_j\}}\frac{1}{2}(\theta_i - b_j)$. Throughout the paper, whenever player i and player j are mentioned in the same sentence, it is implicitly assumed that $j \neq i$.

Before the auction the players go through two rounds of alternating offers, where each player can try to bribe his rival (if he wishes) so that the latter will eliminate himself from the auction. Specifically, player 1 offers player 2 a nonnegative bribe, which player 2 can either accept or reject. Acceptance effectively ends the game: player 1 pays player 2 the offered amount, and in exchange player 2 eliminates himself from the auction.¹² If player 2 rejects player 1’s offer he counters with a nonnegative bribe offer of his own. Now, player 1 can accept or reject this bribe, and in case of acceptance he eliminates himself from the auction. Not offering a bribe is a feasible action for either player, modeled as “offering zero.” If both offers are rejected, then the pre-auction phase ends and the players turn to play the auction noncooperatively, in which case it is *assumed* that they bid their valuations truthfully.¹³

A (pure) strategy for player 1 in this extensive form is a specification of (i) a bribing function and (ii) a family of acceptance rules—one for every possible continuation game. For player 2, a (pure) strategy is modeled as a family of functions $\{b_2(\cdot|x)\}_{x \in \mathbb{R}_+}$, where $b_2(\cdot|x)$ is the function prescribing behavior in

¹¹The tie-breaking rule employed by the auctioneer is not important—the results of this paper hold for any tie-breaking rule.

¹²The players can commit: once a player takes a bribe, he commits to staying out of the auction.

¹³This is a reasonable assumption, because bidding truthfully is a weakly dominant auction-strategy for a player independent of his information.

the continuation game that follows the offer x . For each x , $b_2(\cdot|x): [0, 1] \rightarrow \mathbb{R}_+ \cup \{\text{“accept”}\}$, where $r \in \mathbb{R}_+$ is interpreted as “reject player 1’s offer and make the counteroffer r .” I restrict my attention to pure strategies.

I employ the notation b_1 to denote player 1’s bribing function in the first round, and σ to denote a generic strategy profile. A strategy σ is **monotonic** if (i) for player 1, the bribing function b_1 is weakly increasing in θ_1 , and (ii) for player 2, each function $b_2(\cdot|x)$ is weakly increasing, where the action “accept” is identified with the number -1 .

2.1 Solution concept

Recall that a perfect Bayesian equilibrium (PBE) is a strategy-belief pair, (σ, μ) , such that σ_i prescribes a best-response for i against σ_j in each of i ’s information sets, and beliefs are derived from Bayes’ rule whenever possible.¹⁴ Consider the following refinement:

- (A) Whenever player 1 sees an unexpected offer $x \leq 1$ from player 2, any belief he may form assigns probability 1 to the event $\{\theta_2 \geq x\}$.

(A) is found in equilibria that survive *iterated deletion of weakly dominated strategies*. Unfortunately, however, iterative dominance presents substantial problems. In Appendix C, I explain these problems and show that a weaker version of iterative dominance—*one round of deletion of weakly dominated strategies*—overcomes them and implies (A). An **equilibrium** in Sections 2-5 means an (A)-satisfying PBE. In Section 6 I depart from the IPV setting and consider interdependent valuations; there, I drop the refinement and take PBE

¹⁴I do not introduce formal notation for beliefs (as a collection of probability distributions), since it will not be needed in the sequel.

to be the solution concept. The reasons for this slight inconsistency will be explained there.

I will abuse language a little and call an equilibrium **monotonic** if its strategy profile has the associated property. The idea behind (or justification for) a monotonic equilibrium is that players who have a higher valuation for the good also have a higher willingness to pay for their rival's abstention. Monotonic equilibria are those equilibria in which the behavior of a briber can be interpreted as expressing (weakly) this willingness to pay. An equilibrium is **efficient** if it leads to a Pareto efficient allocation conditional on every type-realization.¹⁵ An equilibrium is **bribery-involving** if, under this equilibrium, bribing occurs with a strictly positive probability.

Remark 1: Our game is a signaling game with a continuum of types. In such games equilibria are typically insensitive to the behavior of a single (bordering) type. For example, suppose that there is an equilibrium in which all types $\theta_2 < \frac{1}{2}$ accept a certain bribe b^* , and all types $\theta_2 > \frac{1}{2}$ reject that bribe and counter it with the same offer. Then the threshold type $\theta_2 = \frac{1}{2}$ is indifferent between these two actions; either action can therefore be supported in equilibrium. Throughout the paper, any statement of the form “ σ is the unique PBE profile such that...” means that it is unique up to the behavior of such bordering types. Similarly, I will sometimes abuse language a little, as follows: statements of the form “type θ_j accepts a bribe b if and only if $b > x$ ” should

¹⁵The results of this paper continue to hold if instead of the pointwise-requirement of efficiency, an “efficient equilibrium” would only be taken to mean that it leads to a Pareto efficient allocation with probability 1. Nothing essential would change, except that the proof of the main result would be much more detail-heavy. I therefore adopt the slightly stronger notion of efficiency.

be interpreted as saying only that bribes $b > x$ are accepted and bribes $b < x$ are rejected.

Remark 2: I do not impose a condition, analogous to (A), that would restrict player 2's beliefs in case he sees an unexpected offer from player 1. As opposed to player 2, who is the last mover and therefore has no reason to offer a bribe in excess of his valuation, player 1 may offer such a bribe, and, in fact, such behavior can occur in equilibrium (see Section 4 below).

3 Efficient monotonic equilibria

Recall that $\pi_i^*(\theta_i)$ denotes θ_i 's expected payoff in the dominant-strategy equilibrium of the noncooperative auction. It is readily verified (through integration by parts, for example) that $\pi_i^*(\theta_i) = \int_0^{\theta_i} F_j(t)dt$. Let $\tilde{\pi}_1(\theta_1, x)$ denote the expected payoff of type θ_1 from competing in the auction against player 2 whose type is distributed on $[x, 1]$ according to $F_2|_{\{\theta_2 \geq x\}}$.^{16,17} For stating Theorem 1, we first need to define the following strategy profile, σ^* :

- Player 1 offers bribes according to $b_1(\theta_1) = \theta_1 - \pi_1^*(\theta_1)$. When player 2 sees an offer of the form $x - \pi_1^*(x)$ for some $x \in [0, 1]$, he rejects it if and only if $\theta_2 > x$, in which case he counters with $\pi_1^*(x)$; offers greater than $1 - \pi_1^*(1)$ are also accepted by player 2. Player 1's acceptance policy of counterbribes b is as follows: If player 1's initial offer was of the form $x - \pi_1^*(x)$ for some $x \in [0, 1]$ and player 2 countered with $b = \pi_1^*(x)$, then

¹⁶Given a cdf F and an event A , $F|_A$ denotes the associated conditional distribution.

¹⁷Specifically, $\tilde{\pi}_1(\theta_1, x) = 0$ for $\theta_1 \leq x$ and $\tilde{\pi}_1(\theta_1, x) = \int_x^{\theta_1} (\theta_1 - t) \frac{f_2(t)}{1 - F_2(x)} dt$ for $\theta_1 > x$. Note that $\tilde{\pi}(\theta_1, 0) = \pi_1^*(\theta_1)$.

player 1 accepts b if and only if b is weakly greater than $\tilde{\pi}_1(\theta_1, x)$; in any other case, player 1 accepts a counterbribe b if and only if $\theta_1 < 2b$.

Theorem 1. (*Characterization of efficient monotonic equilibria*) (1) *There exist beliefs μ^* such that (σ^*, μ^*) is an equilibrium if and only if $\mathbb{E}(\theta_2) \geq \frac{1}{2}$.* (2) *If (σ, μ) is an efficient monotonic equilibrium and the density of F_1 , f_1 , satisfies $f_1'(0) > 0$, then (i) $\mathbb{E}(\theta_2) \geq \frac{1}{2}$, and (ii) σ induces the same path of play as σ^* .*

Remark 3: Suppose that we restrict attention to priors (F_1, F_2) such that $f_1'(0) > 0$. Under this restriction, the condition $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ is sufficient and necessary for the existence of an efficient and monotonic equilibrium. Moreover, such equilibrium, when exists, is unique (up to off-path behavior); it is bribery-involving.

Remark 4: Without the condition $f_1'(0) > 0$, the fact that an equilibrium is efficient and monotonic does not imply that it is bribery-involving (see the Example below).

Remark 5: Even without the condition $f_1'(0) > 0$, the inequality $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ is sufficient for σ^* to be an equilibrium profile. However, it is an open question whether $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ is necessary for the existence of an efficient monotonic equilibrium (i.e., it is an open question whether there exists such an equilibrium in a case where $f_1'(0) \leq 0$ and $\mathbb{E}(\theta_2) < \frac{1}{2}$).

Remark 6: One can generalize the model as to accommodate a general reserve price $r \in (0, 1)$. Then a counterpart of Theorem 1 obtains, where the collusive strategy is the “ r -counterpart” of the aforementioned σ^* : on its path,

player 1 of type $\theta_1 \geq r$ offers the bribe $\theta_1 - \pi_1^*(\theta_1, r)$, where $\pi_1^*(\cdot, r)$ is the expected payoff function in the noncooperative auction with a reserve price r , and player 2 employs an efficient decision rule and counters with $\pi_1^*(\cdot, r)$ following rejections. For uniqueness, a condition analogous to the local convexity of F_1 at the origin is needed. Proving the result for a general r adds no further insight but entails substantially heavier notation.

In the equilibria described in Theorem 1, player 1 reveals his type by offering his surplus (the difference between what he gets in a competition-free world and what he gets under competition) and player 2 employs an efficiency acceptance rule, which is possible because player 1's surplus is a strictly increasing function of his type. When player 2 rejects θ_1 's offer, he reacts to it by making the counteroffer $\pi_1^*(\theta_1)$, which player 1 accepts.¹⁸

The reason that $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ is necessary for σ^* to be supported in equilibrium is that it is equivalent to the following condition, which is necessary for this purpose¹⁹

$$\pi_1^*(\theta_1) \leq \frac{\theta_1}{2} \quad \forall \theta_1 \in [0, 1]$$

¹⁸The reader may suspect that once player 1 reveals his type and player 2 rejects his offer—which, in an efficient equilibrium, is a signal that $\theta_2 \geq \theta_1$ —it must be the case that player 2 counters with an arbitrarily small bribe, because player 1 would agree to any offer from a stronger opponent. This cannot be a part of an equilibrium, however, because then low types of player 2 would have an incentive to mimic higher types. In addition, note that an unexpected bribe indicates that player 2 is not following the equilibrium and therefore player 1's beliefs are unrestricted in such an information set. In particular, player 1 need not infer that player 2 is of a higher type, and therefore need not be willing to accept any bribe.

¹⁹See Lemma 1 in Appendix A.

To see this necessity, assume by contradiction that $\pi_1^*(\theta_1) > \frac{\theta_1}{2}$ for some θ_1 , suppose that this θ_1 is the type that was truthfully revealed through the first offer, and suppose further that $\theta_2 > \theta_1$. In this case player 2 has profitable deviation from the equilibrium action: to offer $\pi_1^*(\theta_1) - \epsilon$ instead of $\pi_1^*(\theta_1)$, for some small $\epsilon > 0$. This offer is necessarily accepted by player 1, because he assigns probability 1 to $\{\theta_2 \geq \pi_1^*(\theta_1) - \epsilon\}$ and will therefore accept the bribe if $\pi_1^*(\theta_1) - \epsilon \geq \theta_1 - (\pi_1^*(\theta_1) - \epsilon)$, which is clearly satisfied for a sufficiently small $\epsilon > 0$.

The condition $f_1'(0) > 0$ rules out equilibria in which the event “no bribing” occurs with a strictly positive probability. Without it, equilibria in which the auction is played noncooperatively with a strictly positive probability are possible. Here is an example of one such equilibrium; in this equilibrium, the auction is played noncooperatively with probability 1. Namely, the bribing stage is effectively skipped.

Example of a “no bribing” equilibrium: Suppose that F_i is uniform for each $i = 1, 2$. Consider the strategy where player 1 offers zero independent of his type, player 2 accepts a bribe b if and only if $b > \theta_2$, player 1 accepts a counterbribe b if and only if $2b > \theta_1$, and the rejections of player 2 are followed by the counterbribe zero. Call this profile σ^N (the superscript N stands for “no bribing”).

Proposition 1. *If F_i is uniform for each i , then there exists a system of beliefs μ such that (σ^N, μ) is an equilibrium.*

To see the rule of the condition $f_1'(0) > 0$, consider the above-mentioned profile σ^N . Suppose that player 1 adhered to this strategy and offered nothing, and it is now player 2’s turn to move. If he deviates and offers some small $\epsilon \in (0, \frac{\theta_2}{2})$,

then all types $\theta_1 \leq 2\epsilon$ would accept it, because player 1 assigns probability 1 to $\{\theta_2 \geq \epsilon\}$. Suppose that player 1 is “as optimistic as possible,” and he assigns probability 1 to $\{\theta_2 = \epsilon\}$ so only $\theta_1 \leq 2\epsilon$ accept the counterbribe ϵ . Let $E = \{\theta_1 > 2\epsilon\}$. Conditional on E player 2’s payoff from the deviation equals his payoff from the equilibrium strategy, because the auction is played noncooperatively in either case. Conditional on the complement of E , these payoffs are the same if F_1 is uniform, because player 2 wins with certainty and pays ϵ in expectation. However, if $f_1'(0) > 0$, then in a neighborhood of 0 strong types of player 1 are relatively more likely than weak types, so by excluding them player 2 gains relatively “a lot.” Thus, he strictly prefers to deviate to some small $\epsilon > 0$, and the equilibrium unravels.

3.1 Payoffs

In the equilibria described in Theorem 1, player 1’s ex post payoff equals his expected noncooperative payoff, $\pi_1^*(\theta_1)$. Player 2’s expected payoff can be computed directly from the strategy σ^* : it is $\int_0^{\theta_2} [\theta_2 - \pi_1^*(t)]f_1(t)dt + \int_{\theta_2}^1 [t - \pi_1^*(t)]f_1(t)dt \equiv \Pi_2(\theta_2)$. Note that $\Pi_2(\theta_2) = \pi_2^*(\theta_2) + C$, where $C = \int_0^1 [t - \pi_1^*(t)]f_1(t)dt$ is the expected surplus extracted from player 1. The fact that this is the expected payoff can also be seen in the following alternative way: we know, due to the classic result of Myerson (1981), that since σ^* —like the dominant strategies in the noncooperative second-price auction—implements the efficient allocation, the expected payoff of player 2 of type θ_2 under σ^* equals $\pi_2^*(\theta_2) + C_0$, where C_0 is the expected payoff of player 2’s minimal type under σ^* . Since under σ^* the minimal type of player 2 accepts the bribe of every θ_1 , $C_0 = \int_0^1 [t - \pi_1^*(t)]f_1(t)dt = C$.

3.2 Comparison with the inefficiency of the ES equilibrium

As was shown by ES, efficiency (with probability 1) is impossible in a bribery-involving equilibrium when there is only one round of offers before a second-price auction, because perfect signaling of the briber's type is not incentive-compatible for him. To see this, assume by contradiction that there is an equilibrium in which the first mover reveals his type through a strictly increasing bribing function, and consider type $\theta_1 = 1$; if he mimics type $1 - \epsilon$, for some small $\epsilon > 0$, then the lowered bribe will surely be accepted, because player 2 cannot hope to obtain more than ϵ in the auction. This shows that there must be pooling among all the types of player 1 above a certain threshold.

This is no longer true with two rounds of bribing, because the second round makes it possible for player 2 to reject an offer without triggering the noncooperative auction. Thus, it is not enough for player 1 to signal that his type is "sufficiently high" in order to secure acceptance. One can think of the following message being implied by player 2 when he rejects a relatively high offer: "I know that your type is high, but mine is even higher, so you better accept my counteroffer and not compete against me in the auction."

Furthermore, the two-round game looks fundamentally different from the TIOLI game also for the *low types* of player 1. In the TIOLI game, player 1 will never offer an amount which exceeds his valuation, but he may very well do so in the two-round game, hoping that his bribe will be rejected and trigger a generous counteroffer. Indeed, this consideration gives rise to equilibria in which there is complete pooling in the first round, where all the types of player 1 offer the same strictly positive offer. Low types follow the equilibrium, de-

spite the fact that they may end up with a negative ex post payoff, because there is a positive probability that player 2's type is high, in which case he rejects the common bribe and counters it with a positive offer of his own. Such equilibria are described formally below.

4 Equilibria with complete first-round pooling

Consider a cdf-pair (F_1, F_2) for which there exist $y, c \in [0, 1]$ such that the following conditions hold:

- (I) $y - c \geq F_1(2\beta)(y - \beta) + \int_{2\beta}^y (y - t)f_1(t)dt$ for all $\beta \geq 0$,
- (II) $c > \max\{1 - \mathbb{E}(\theta_2|\theta_2 \geq y), F_2(y)y\}$, and
- (III) $F_2(y)(1 - y) + c \geq 1 - \mathbb{E}(\theta_2)$.

When this is the case, say that (F_1, F_2) *satisfies (I)-(III) with respect to the parameters y and c* .

Given these numbers $y, c \in [0, 1]$, define the strategy profile $\sigma^*(y, c)$ as follows. Player 1 offers $b = y - c$ independent of his type, player 2 accepts b if and only if $\theta_2 \geq y$, and if he rejects it he counters with c . The counterbribe c is accepted by all θ_1 . Any off-path offer β made by player 1 is accepted by player 2 if and only if $\theta_2 \leq \beta$. Any off-path offer β made by player 2 is accepted by player 1 if and only if $\theta_1 \leq 2\beta$.

When the profile $\sigma^*(y, c)$ is followed, there is complete pooling among all the types of player 1: they all offer b , to which player 2 responds with a simple cut-off strategy: he accepts b if his type is below the threshold y , and rejects it and counters it with c otherwise. If (F_1, F_2) satisfies (I)-(III) with respect to the parameters y and c , then this behavior can be sustained in equilibrium.

Proposition 2. *If (F_1, F_2) satisfies (I)-(III) with respect to the parameters y and c then there exist beliefs μ such that $(\sigma^*(y, c), \mu)$ is an equilibrium.*

Proposition 3. *There exist $y, c \in [0, 1]$, $y > c$, and cdf-pairs (F_1, F_2) that satisfy (I)-(III) with respect to the parameters y and c .*

5 Embedding the ES equilibrium in the two-round game

The equilibria in Section 3 and 4 were obtained under certain distributional assumptions. One may wonder whether the game has an equilibrium in the absence of these assumptions, because equilibrium existence is a nontrivial issue in games such as the one studied here. For example, in Rachmilevitch (2011) I showed that in the game that consists of a first-price auction and a TIOLI bribing protocol, there is a large class of (well-behaved) type distributions given which that game does not have an equilibrium. This section is dedicated to establishing the existence of an equilibrium in the two-round second-price game under general conditions.

The following modification of the ES equilibrium comes to mind as a plausible equilibrium candidate for the two-round game: let player 1 offer zero independent of his type and let the ES equilibrium be played starting at the second round, with player 2 as the briber and player 1 as the respondent. Establishing that this path of play is sustainable as an equilibrium is an easy task, because the only thing that requires verification is that player 1 does not have an incentive to deviate to a positive bribe; the rest of the incentive constraints follow from the fact that the continuation game that starts at the second round *is* the ES game.

Proposition 4. *Suppose that for each i the distribution F_i is log-concave. Then the following behavior can be sustained in equilibrium: player 1 offers zero independent of his valuation, and the ES equilibrium is played in the continuation game that starts at the second round.*

Log-concavity is only needed in the continuation game that starts at the second round, because, for technical reasons, it is needed in the equilibrium-characterization result of ES. I view this technical assumption as a mild one; most well-behaved distributions considered in the literature are log-concave.

In the equilibrium described in Proposition 4 the first round is skipped. Recall that in Proposition 1 we saw a “no bribing” equilibrium, in which both bribing rounds are skipped. As was explained, if $f'_1(0) >$ then this equilibrium unravels. Note, that $f'_1(0) > 0$ is allowed in Proposition 4. The reason for the difference is that in the “no bribing” equilibrium player 2’s payoff equals his noncooperative payoff, $\pi_2^*(\theta_2)$, and a small positive bribe can improve on it, provided that $f'_1(0) >$. In the equilibrium from Proposition 4, by contrast, player 2’s payoff is *not* $\pi_2^*(\theta_2)$ —it is the first mover’s payoff from the ES equilibrium, which is greater than the competitive payoff. In this case, a deviation to a positive bribe is nonprofitable.

6 Interdependent valuations

In this section I depart from the IPV assumption by assuming that a player’s valuation does not coincide with his type; instead, valuations are functions of both types, hence are interdependent. Player i ’s valuation from obtaining the good is $v(\theta_i, \theta_j)$, and overall preferences, as before, are quasi-linear. The function v satisfies $v(0, 0) = 0$ and it is strictly increasing in θ_i , given any

fixed $\theta_j > 0$. I assume that assigning the good to the maximal type is Pareto efficient,²⁰ and that the function $V(x) \equiv v(x, x)$ is differentiable with $V' > 0$. Call these valuations *general indeterdependent valuations*.

With these general preferences, the second-price auction does not have weakly dominant strategies and bidding behavior needs to be specified after every leading-to-the-auction history. In particular, the game is, formally speaking, a different extensive form from the one that was described in Section 2. I do not spell out the formalism of this enriched model, because it will not be needed (also, it is obvious).

In this model, the equilibria described in Theorem 1 continue to be equilibria. More precisely, there exist equilibria in which player 1 employs a strictly increasing bribing function, player 2 employs an efficient acceptance rule, and rejections by player 2 never lead to the auction—they result in an acceptance by player 1 of a common counterbribe from player 2.

To define the equilibrium strategy, I first define the bribing and counterbribing functions, b and β , respectively.

$$b(\theta) = \int_0^\theta (1 - F_2(t))V'(t)dt,$$

and

$$\beta(\theta) = V(\theta) - b(\theta).$$

Note that for $V(x) = x$ one obtain the bribing and counterbribing functions from Theorem 1.

In order to support a path of play analogous to the one from Theorem 1, one needs to check various deviations and therefore, in particular, to deal with two kinds of off-path auctions. The two kinds correspond to whether player

²⁰i.e., that $a > b$ implies $v(a, b) \geq v(b, a)$.

2 deviated after having seen player 1's offer. If he did not, then the auction to be addressed is a one in which player 1 deviates, by rejecting player 2's equilibrium counterbribe. This is an auction which is played under the following information: player 1's type is the one he revealed, say x , and player 2's type is distributed according to $F_2|_{\{\theta_2 \geq x\}}$. Similarly to Section 3, I denote player 1's expected payoff in such an auction by $\tilde{\pi}_1(\theta_1, x)$. The other kind of auction that needs to be addressed is a one in which player 2 responded to $b(x)$ with a deviation—he reciprocated with a counter different than $\beta(x)$. In such an auction in which, in addition, player i believes with probability one that the opponent's type is θ_j , his expected payoff is denoted by $\hat{\pi}_i(\theta_i, \theta_j)$.²¹ I assume that $\hat{\pi}_1(x, 0) \geq \beta(x)$, that $\tilde{\pi}_1(x, x) = \hat{\pi}_2(x, x) = 0$, and that $\frac{\partial}{\partial \theta_2} v(\theta_2, x) \geq \frac{\partial}{\partial \theta_2} \hat{\pi}_2(x, x)$ for all $x, \theta_2 \in [0, 1]$.²²

Let σ^{**} be the following strategy:

- Player 1 offers bribes according to $b(\theta_1)$. When player 2 sees an offer of the form $b(x)$ for some $x \in [0, 1]$, he rejects it if and only if $\theta_2 > x$, in which case he counters with $\beta(x)$; offers greater than $b(1)$ are also accepted by player 2. Player 1's acceptance policy of counterbribes b is as follows: If player 1's initial offer was of the form $b(x)$ for some $x \in [0, 1]$ and player 2 countered with $\beta(x)$, then player 1 accepts $\beta(x)$ if

²¹To be more precise, I assume that there is an equilibrium-selection mapping that generate the aforementioned functions $\tilde{\pi}_1$ and $(\hat{\pi}_i)_{i=1,2}$; in the IPV model this equilibrium-selection mapping is trivial: it always prescribes the dominant-strategy equilibrium.

²²For example, one can check that these assumptions are satisfied in the case where $v(\theta_i, \theta_j) = \theta_i + \lambda\theta_j$, where $0 < |\lambda| < 1$. In this case the equilibrium-selection mapping assigns type θ_i the bid $(1 + \lambda)\theta_i$. It is easily verified that this behavior is supported by appropriate beliefs as a Bayesian Nash equilibrium in the aforementioned off-path auctions.

and only if it is weakly greater than $\tilde{\pi}_1(\theta_1, x)$; in any other case, player 1 accepts a counterbribe b if and only if $\hat{\pi}_1(\theta_1, 0) < b$.

Theorem 2. (*Existence of efficient collusive equilibrium under general interdependent valuations*) *Assume general interdependent valuations. Then there exists beliefs μ^{**} such that (σ^{**}, μ^{**}) is a PBE.*

Remark 7: Theorem 2, as opposed to Theorem 1, only addresses equilibrium existence, not uniqueness. Establishing uniqueness requires elaborate arguments, even in the IPV setting. Many of these arguments refer to off-path auctions, and since they build on the existence of dominant strategies, they do not have counterparts in the general interdependent case.

Remark 8: The solution concept in Theorem 2 is simply PBE, not an (A)-satisfying PBE like in Theorem 1. The reason for not assuming refinement (A) is that in order to prove that the minimal bribe that type θ_1 is willing to accept is $\beta(\theta_1)$, one needs to prove that any smaller bribe will be rejected. In Theorem 1, this is done by combining refinement (A) with the special (and simple) payoff structure of the dominant strategy equilibrium of the IPV second-price auction. This is no longer the case where player 1’s off-path payoffs are described by the function $\hat{\pi}_1$, so in order to guarantee the optimality of such rejections, I allow player 1 to adopt the optimistic belief that his deviating opponent is of the minimal type, $\theta_2 = 0$. Note that the reason for the unattractiveness of such off-path beliefs in the IPV model—the last mover has no reason to offer more than his valuation—no longer have a bite in the general interdependent case. This is because of two reasons: first, “valuation” no longer equals “type”; second, deviations may be interpreted as signals of intended bidding behavior—i.e., as an equilibrium-selection device.

The (general interdependent) valuations are **separable** if $v(\theta_i, \theta_j) = \theta_i + \phi(\theta_j)$ for some differentiable function ϕ . It turns out that under separability, the full surplus extraction result from the IPV model continues to hold: the positional advantage of the second mover enables him to keep the first mover’s payoff at the competitive level, and extract the entire collusive gain.

Theorem 3. *(Full surplus extraction under separable interdependent valuations) If valuations are interdependent and separable, then in the equilibrium described in Theorem 2, player 1’s expected payoff coincides with his expected competitive payoff—namely, with his payoff in the symmetric Bayesian Nash equilibrium of the noncooperative auction.*

Theorem 3 refers to player 1’s expected payoff in the symmetric equilibrium of the noncooperative auction. Interestingly, it turns out that this payoff is independent of ϕ ; in particular, it equals the one from the dominant strategy equilibrium of the IPV model, the one corresponding to $\phi \equiv 0$.

Proposition 5. *Assume separable interdependent valuations. Then the non-cooperative second-price auction has a unique symmetric Bayesian Nash equilibrium; in it, the expected payoff of type θ_i of player i is $\pi_i^*(\theta_i)$.*

7 Conclusion

This paper has presented a sequential model of collusion via bribes: two players are about to attend a second-price IPV auction, prior to which each can try, in his turn, to bribe his rival to drop out. This is a natural extension of the “take-it-or-leave-it” (TIOLI) model.

The two models differ along several dimensions. First, whereas efficiency is

possible with two rounds, it is impossible with one round. The reason is that in the TIOLI game it is enough for the first mover to signal that he is “sufficiently strong” in order to secure acceptance, hence perfect signaling is impossible in equilibrium. Perfect signaling, however, is necessary for implementing the efficient allocation. With two rounds the aforementioned reasoning no longer applies. It is no longer true that player 2 will accept player 1’s bribe whenever he infers that the latter is “sufficiently strong,” because it is feasible for him to decline the bribe without triggering the noncooperative auction. In particular, perfect signaling can be sustained in equilibrium, and efficiency can be achieved.

The possibility of efficiency may seem surprising when contrasted with the inefficiency result of Myerson and Satterthwaite (1983), who showed that (under certain mild conditions) one cannot achieve ex post efficiency in a bilateral trade problem by employing an incentive compatible, budget balanced, individually rational mechanism. Stated a little differently, this impossibility is about transforming the initial ownership structure, where the seller’s share is one and the buyer’s share is zero, to a structure where the former is zero and the latter is one. Cramton, Gibbons, and Klemperer (1987) extended the Myerson-Satterthwaite setting and considered N agents with ownership shares $(\alpha_1, \dots, \alpha_N)$ ($\alpha_i \geq 0$ for all i and $\sum_i \alpha_i = 1$), who face the task of transforming it to some other ownership structure $(\alpha'_1, \dots, \alpha'_N)$. They showed that this transformation is achievable via a mechanism with all the aforementioned properties, provided that the initial shares are sufficiently egalitarian—i.e., sufficiently close to $(\frac{1}{N}, \dots, \frac{1}{N})$. The efficiency result of the current paper can be thought of as corresponding to the case where both “ownership shares” are zero.

The key point in moving from a TIOLI protocol to a two-round protocol is

that enlarging the set of possible responses of the second mover enlarges the set of the game’s equilibria. Not only separating—pooling equilibria also come into existence when a round of offers is added. By contrast, the equilibria of the TIOLI game are “in between” these extremes—every TIOLI equilibrium involves separation among low types and pooling among high types.²³

An analysis analogous to the one presented in the current paper can be carried out for other auction formats. The main feature that distinguishes the second-price format is that the behavior in the auction—truthful bidding—is independent of the pre-auction activity. This separation no longer exists with other formats. For example, in Rachmilevitch (2011) I showed that when an IPV first-price auction is preceded by a TIOLI stage, the information which is inferred from the bribing stage necessarily affects the bidding behavior in the auction; moreover, the link between the two unravels any bribery-involving continuous equilibrium: the only equilibrium of the first-price TIOLI game in which the bribing function is continuous is such that the bribing function is identically zero. As for the two-round game, under a slight modification of the solution concept, the path of the equilibria from Theorem 1 can also be sustained in equilibrium under the first-price format.²⁴

²³As mentioned in subsection 1.1, ES derive the unique equilibrium of the TIOLI game in which the bribing function is continuous; this function is strictly increasing up to a certain threshold, and is constant from this threshold onwards. In addition, ES derive all the D1-satisfying equilibria of the TIOLI game; the bribing function in each such equilibrium is strictly increasing up to a threshold $x < 1$ and is constant on $(x, 1]$. Moreover, it coincides with the bribing function of the unique continuous equilibrium on $[0, x)$; the threshold x is its unique discontinuity point.

²⁴Instead of PBE, one needs to consider *essentially perfect Bayesian equilibrium* (EPBE) when the format is first-price; EPBE is due to Blume and Heidhues (2006); see Rachmilevitch (2010) for details.

The following question remains, at present, open: I do not know whether the two-round (IPV) game has fully-revealing equilibria other than the efficient one. This turns out to be a difficult problem even in the special case where both type distributions are uniform.

Acknowledgments This paper is based on my job market paper, which was written under the supervision of Wojciech Olszewski, Ehud Kalai, and Jim Schummer; I am grateful for their continuous invaluable support. I would also like to thank Guy Arie, Eddie Dekel, Jeff Ely, Sivan Frenkel, Hanna Halaburda, Aviad Heifetz, Toomas Hinnosaar, Tai-Wei Hu, Bernardo Moreno, Alessandro Pavan, Robert Porter, Ron Siegel, Marciano Siniscalchi, Yannay Spitzer, and Asher Wolinsky for stimulating conversations. Finally, I would like to thank Philippe Jehiel and an anonymous referee for helpful comments.

8 Appendix A: Proof of Theorem 1

The following lemmas will be needed for the proof of Theorem 1.

Lemma 1. $\mathbb{E}(\theta_2) \geq \frac{1}{2}$ if and only if $\pi_1^*(\theta_1) \leq \frac{\theta_1}{2} \quad \forall \theta_1 \in [0, 1]$.

Proof. Suppose that $\pi_1^*(\theta_1) \leq \frac{\theta_1}{2} \quad \forall \theta_1 \in [0, 1]$. Taking $\theta_1 = 1$ gives $1 - \mathbb{E}(\theta_2) \leq \frac{1}{2}$, or $\mathbb{E}(\theta_2) \geq \frac{1}{2}$. Conversely, suppose that $\mathbb{E}(\theta_2) \geq \frac{1}{2}$. We need to prove that $\gamma(\theta_1) \geq 0$, where $\gamma(\theta_1) \equiv \frac{\theta_1}{2} - \pi_1^*(\theta_1)$. Note that $\gamma'(\theta_1) = \frac{1}{2} - F_2(\theta_1)$, hence $\gamma''(\theta_1) = -f_2(\theta_1) \leq 0$, so γ is concave. Also, $\gamma(0) = 0$ and $\gamma(1) = \frac{1}{2} - (1 - \mathbb{E}(\theta_2)) = \mathbb{E}(\theta_2) - \frac{1}{2} \geq 0$, where the inequality is by assumption. Therefore, $\gamma(\theta_1) \geq \theta_1\gamma(1) + (1 - \theta_1)\gamma(0) = \theta_1\gamma(1) \geq 0$. \square

Let:

$$Z(x, \theta_1) \equiv F_2(x)(\theta_1 - x + \pi_1^*(x)) + (1 - F_2(x))\max\{\pi_1^*(x), \tilde{\pi}_1(\theta_1, x)\}$$

where $\tilde{\pi}_1(\theta_1, x)$ denotes θ_1 's expected payoff in the noncooperative auction against player 2 whose type is distributed on $[x, 1]$ according to the conditional distribution $F_2|_{\{\theta_2 \geq x\}}$.

Definition 1. *The distribution function F_2 , defined on $[0, 1]$, satisfies condition C1 if $x = 1$ is a maximizer of $Z(\cdot, 1)$.*

Condition C1 is simply the incentive compatibility constraint of the maximal type of player 1 in the first information set in the game. As the following lemma shows, if this constraint is satisfied for the maximal type (i.e., if C1 holds), then it is satisfied for all types.

Lemma 2. *If C1 is satisfied, then $x = \theta_1$ is a maximizer of $Z(\cdot, \theta_1)$ for all $\theta_1 \in [0, 1]$.*

Proof. Suppose that C1 is satisfied. Assume by contradiction that $x = \theta_1$ does not maximize $Z(\cdot, \theta_1)$ over $[0, 1]$, for some $\theta_1 \in [0, 1]$. Clearly we can assume that $\theta_1 > 0$. Since the objective is continuous and $[0, 1]$ is compact, a maximizer $x^* \in [0, 1]$ exists. First, I argue that $x^* < \theta_1$. To see this, assume by contradiction that $x^* \geq \theta_1$. Then, x^* is a maximizer of $Z(\cdot, \theta_1)$ over $[\theta_1, 1]$. On this sub-domain, the objective is

$$Z(x, \theta_1) = F_2(x)(\theta_1 - x + \pi_1^*(x)) + (1 - F_2(x))\pi_1^*(x) = F_2(x)(\theta_1 - x) + \pi_1^*(x). \quad (1)$$

The FOC is $f_2(x)(\theta_1 - x)$, which is negative on $(\theta_1, 1]$, hence on this range the maximum is at $x^* = \theta_1$, a contradiction. Hence, as was argued, $x^* < \theta_1$.

It is easy to see that there is a unique $m = m(\theta_1) \in (0, \theta_1)$ such that the objective satisfies (1) on $[m, 1]$. Since on that sub-domain θ_1 is the unique maximizer of $Z(\cdot, \theta_1)$, we only need to consider $[0, m]$, where the objective takes the form

$$Z(x, \theta_1) = F_2(x)(\theta_1 - x + \pi_1^*(x)) + \int_x^{\theta_1} (\theta_1 - t)f_2(t)dt.$$

Since x^* gives the objective a higher value than θ_1 does, it follows that

$$Z(x^*, \theta_1) = F_2(x^*)(\theta_1 - x^* + \pi_1^*(x^*)) + \int_{x^*}^{\theta_1} (\theta_1 - t)f_2(t)dt > \int_0^{\theta_1} (\theta_1 - t)f_2(t)dt = Z(\theta_1, \theta_1),$$

and therefore

$$F_2(x^*)(\theta_1 - x^* + \pi_1^*(x^*)) > \int_0^{x^*} (\theta_1 - t)f_2(t)dt.$$

Adding $F_2(x^*)(1 - \theta_1) = \int_0^{x^*} (1 - \theta_1)f_2(t)dt$ to both sides gives

$$F_2(x^*)(1 - x^* + \pi_1^*(x^*)) > \int_0^{x^*} (1 - t)f_2(t)dt,$$

therefore

$$Z(x^*, 1) = F_2(x^*)(1 - x^* + \pi_1^*(x^*)) + \int_{x^*}^1 (1 - t)f_2(t)dt > \int_0^1 (1 - t)f_2(t)dt = Z(1, 1),$$

in contradiction to *C1*. □

The following lemma guarantees that the incentive constraint of the maximal type, *C1*, is indeed satisfied.

Lemma 3. *Every distribution F_2 satisfies condition C1.*

Proof. Assume by contradiction that there exists an F_2 such that there exists an $x < 1$ such that $Z(x, 1) > Z(1, 1) = \pi_1^*(1)$. By Lemma 2, it must be that this x is sufficiently small, so that $\tilde{\pi}_1(1, x) \geq \pi_1^*(x)$, and consequently the value of the objective satisfies

$$Z(x, 1) = F_2(x)(1 - x + \pi_1^*(x)) + \int_x^1 (1 - t)f_2(t)dt > \int_0^1 (1 - t)f_2(t)dt,$$

or

$$F_2(x)(1 - x + \pi_1^*(x)) > \int_0^x (1 - t)f_2(t)dt.$$

Since $\int_0^x (1 - t)f_2(t)dt = \int_0^x [x - t + (1 - x)]f_2(t)dt = \pi_1^*(x) + \int_0^x (1 - x)f_2(t)dt = \pi_1^*(x) + F_2(x)(1 - x)$, we obtain

$$F_2(x)(1 - x + \pi_1^*(x)) > \pi_1^*(x) + F_2(x)(1 - x),$$

and therefore

$$F_2(x)\pi_1^*(x) > \pi_1^*(x),$$

a contradiction. □

With the lemmas at hand, we can turn to the proof of the theorem.

Proof. (1). If $\mathbb{E}(\theta_2) \geq \frac{1}{2}$, then there exist beliefs μ such that (σ^*, μ) is an efficient monotonic equilibrium.²⁵ Let μ be a system of beliefs such that (i) when player 2 sees an offer of the form $x - \pi_1^*(x)$ for some $x \in [0, 1]$, he assigns probability 1 to the event $\{\theta_1 = x\}$, while when he sees an offer strictly greater than $1 - \pi_1^*(1)$ he assigns probability 1 to the event $\{\theta_1 = 1\}$, and (ii) in the information set following the rejection of the first offer b_1 where player 2 offered the counterbribe b_2 , player 1's beliefs are the following: if $b_1 = x - \pi_1^*(x)$ for some $x \in [0, 1]$ and $b_2 = \pi_1^*(x)$, then 1 believes that θ_2 is distributed according

²⁵The “only if” statement follows (easily) from the combination of (i) Lemma 1, and (ii) the explanation in the text that follows Remark 6.

to $F_2|_{\{\theta_2 \geq x\}}$; otherwise, he believes that $\{\theta_2 = \min\{b_2, 1\}\}$.

It is easy to see that σ^* is efficient and monotonic. We need to prove that σ_i^* prescribes a best-response for i against σ_j^* , in each of i 's information sets (given his belief μ_i).

Start with player 1. Suppose that he is of type $\theta_1 \in [0, 1]$. First, I argue that he does not have a strictly profitable deviation in the first node in the game tree. To see this, note that if there exists such a deviation to a different bribe, then there exists such a deviation where he mimics a different type (because all offers above $1 - \pi_1^*(1)$ are accepted by player 2). But this contradicts the fact that $x = \theta_1$ is a maximizer of $Z(\cdot, \theta_1)$, which, in turn, is guaranteed by Lemma 2 and Lemma 3.

Next, consider the information set where player 1 responds to player 2's counteroffer b_2 . Suppose first that player 1's initial offer was of the form $x - \pi_1^*(x)$ for some $x \in [0, 1]$ and player 2 countered with $b_2 = \pi_1^*(x)$. Then clearly following σ_1^* —by definition—is a best-response. In any other case player 1 believes that $\{\theta_2 = \min\{b_2, 1\}\}$. If $\min\{b_2, 1\} = 1$ the obviously acceptance of b_2 is a best-response, and this is the response instructed by σ_1^* . If, on the other hand, $\min\{b_2, 1\} = b_2$, then since player 1's belief assigns probability 1 to the event $\{\theta_2 = b_2\}$, it follows that accepting the bribe if and only if $b_2 \geq \theta_1 - b_2$ —as instructed by σ_1^* —is a best-response.

Consider now player 2 of an arbitrary type $\theta_2 \in [0, 1]$. In his first information set in the game he responds to player 1's offer. If this offer takes the form $x - \pi_1^*(x)$ for some $x \in [0, 1]$, he infers that $\theta_1 = x$. First, I argue that any counter-offer $b < \pi_1^*(x)$ is rejected. To see this, assume by contradiction that there exists a counter-offer $b < \pi_1^*(x)$ which is accepted by player 1. This offer is accepted if and only if $\theta_1 \leq 2b$, or $\frac{x}{2} \leq b$. Combining these inequalities we obtain $\frac{x}{2} < \pi_1^*(x)$; however, since $\mathbb{E}(\theta_2) \geq \frac{1}{2}$, Lemma 1 guarantees that

$$\frac{x}{2} \geq \pi_1^*(x).$$

Hence, player 2 has effectively three options: (a) to accept the bribe $x - \pi_1^*(x)$, (b) to reject and counter with $\pi_1^*(x)$, which is accepted by player 1,²⁶ and (c) to compete in the noncooperative auction against type $\theta_1 = x$.

Consider first $\theta_2 < x$. Here, we need to prove that acceptance is a best-response. If player 2 rejects the bribe then the best he can obtain is $\max\{\theta_2 - \pi_1^*(x), 0\} < x - \pi_1^*(x)$, and therefore adhering to σ_2^* is a best-response. Consider now $\theta_2 \geq x$. Here, we need to prove that countering with $\pi_1^*(x)$ is a best-response. First, note that doing so is weakly better than accepting player 1's offer, and is strictly better if $\theta_2 > x$. This follows from $\theta_2 - \pi_1^*(x) \geq x - \pi_1^*(x)$. Secondly, this is also better than competing the auction, because $\theta_2 - \pi_1^*(x) \geq \theta_2 - x$.

Next, consider the case where player 2 sees an offer strictly greater than $1 - \pi_1^*(1)$. Since he believes that $\theta_1 = 1$, acceptance of this offer is a best-response.

Therefore, (σ^*, μ) is an equilibrium.

(2). Suppose that $f_1'(0) > 0$ and let (σ, μ) be an efficient monotonic equilibrium. Then, $\mathbb{E}(\theta_2) \geq \frac{1}{2}$, and σ has the same path of play as σ^* .

Let b_1 denote player 1's bribing function.

Claim 1: b_1 is strictly increasing.

Proof of Claim 1: Assume by contradiction that there exists some nondegen-

²⁶The counter-bribe $\pi_1^*(x)$ is accepted by player 1 of type $\theta_1 = x$ because $\pi_1^*(x) \geq \tilde{\pi}_1(x, x) = 0$.

erate interval I on which b_1 is constant, taking the value b^* .

Case 1: There exists such an I with $b^ > 0$.*

Let $\underline{\theta}_1 \equiv \inf I$.

Case 1.1: $\underline{\theta}_1 > 0$. Let $(\underline{\theta}_1, \underline{\theta}_1 + \epsilon) \equiv J$, where $\epsilon \in (0, b^*)$. By efficiency, all $\theta_2 \in J$ reject b^* . Let $\beta: J \rightarrow \mathbb{R}_+$ denotes player 2's counterbribing function on this domain.

Case 1.1.1: β is constant on a nondegenerate subinterval of its domain, $J' \subset J$. Let b' denote its level on J' . Note that $b' > 0$. To see this, assume by contradiction that $b' = 0$ and consider a type $\theta_2 \in J'$. When he rejects b^* and counters with $b' = 0$, the offer $b' = 0$ is rejected by player 1 with probability 1;²⁷ hence, this behavior gives θ_2 a payoff which is bounded from above by $\epsilon < b^*$, and it therefore cannot be a part of an equilibrium. Thus, $b' > 0$. Since all θ_1 arbitrarily close to $\underline{\theta}_1$ reject b' , it follows that there exists a $\theta_2^* < \underline{\theta}_1$ who rejects b^* and counters with b' (otherwise, if only $\theta_2 \geq \underline{\theta}_1$ offered the counter $b' > 0$, it would not be incentive compatible for types θ_1 near $\underline{\theta}_1$ to reject it). Therefore b' is rejected with probability 1 by player 1; therefore, rejecting b^* and countering with b' is not a best-response for such a $\theta_2^* < \underline{\theta}_1$.

Case 1.1.2: There does not exist a subinterval $J' \subset J$ on which β is constant. Then, by monotonicity, every type $\theta_2 \in J$ reveals himself via his counter bribe $\beta(\theta_2)$; namely, the following holds for every $\epsilon > 0$: $\beta(\theta_2 - \epsilon) < \beta(\theta_2) < \beta(\theta_2 + \epsilon)$ (if one of these inequalities was an equality, then an interval J' would have existed). Consider types $\theta_1 \in I$ and $\theta_2 \in J$ such that $\beta(\theta_2) > \max\{\theta_1 - \theta_2, 0\}$. Conditional on this realization of types, play goes as follows: player 1 offers b^* , which is rejected by player 2, who counters it with

²⁷In equilibrium player 1 can accept the counter offer $b' = 0$ only if he is certain that $\theta_2 \geq \theta_1$. Thus, only type $\theta_1 = \underline{\theta}_1$ can accept it in equilibrium.

the revealing offer $\beta(\theta_2)$; it is the unique best-response of player 1 to accept this counteroffer, in contradiction to efficiency.

Case 1.2: $\underline{\theta}_1 = 0$. Here, b_1 is constant at the level $b^* > 0$ on an interval of the form $[0, x)$ for some $x > 0$. All $\theta_2 < b^*$ accept this bribe, in contradiction to efficiency.

Case 2: $b^ = 0$ for every interval I on which b_1 is constant.*

In this case $b_1 \equiv 0$ on an interval of the form $[0, x)$ for some $x > 0$. Wlog, suppose that x is the supremum number for which this is true. Since all $\theta_2 > 0$ reject the offer 0, there is a probability $F_1(x) > 0$ that play will go to the second round, where the continuation game is a TIOLI game, played with the following beliefs: the distribution of θ_2 is given by F_2 and the distribution of θ_1 is given by $G \equiv F_1|_{\{\theta_1 < x\}}$. Let $g \equiv G'$. Let $b_2 \equiv b_2(\cdot|0)$ denote 2's bribing function in this continuation game.

I argue that $b_2 \equiv 0$ on $[0, x)$. To see this, assume by contradiction that $b_2(\theta_2) > 0$ for some $\theta_2 < x$. If there exists a $\theta'_2 \in (\theta_2, x)$ such that $b_2(\theta'_2) = b_2(\theta_2) \equiv b^*$, then b_2 is constant at the level $b^* > 0$ on some interval J . Wlog, suppose that J is the maximal such interval (i.e. the union of all such intervals). All types θ_1 sufficiently close to $\inf J$ accept the bribe b^* , and therefore the probability of inefficiency is positive. If, on the other hand, there does not exist such a θ'_2 , then b_2 is strictly increasing on (θ_2, x) , and again we conclude that the probability of inefficiency is strictly positive (given each revealed type $\theta'_2 \in (\theta_2, x)$, all types above and close to him accept his bribe, because bidding in the auction is truthful). Thus, $b_2 \equiv 0$ on $[0, x)$.

Let $0 < \theta_2 < x$ and let $\epsilon \in (0, \frac{\theta_2}{2})$. When player 1 is contemplating acceptance vs. rejection of ϵ (in the continuation game which follows the rejection of 0), he compares the two associated payoffs. The payoff from acceptance

is obvious, ϵ . The payoff from rejection is the expected payoff in the non-cooperative auction, which is bounded from above by $\max\{\theta_1 - \epsilon, 0\}$. To see why this is an upper bound, note that there are exactly two possibilities: (i) ϵ constitutes a detectable deviation (i.e., there does not exist a $\theta_2 \geq x$ such that $b_2(\theta_2) = \epsilon$), or (ii) there exists some $\theta_2 \geq x$ such that $b_2(\theta_2) = \epsilon$. In case (i) the fact that $\max\{\theta_1 - \epsilon, 0\}$ is an upper bound on the payoff player 1 expects follows from the fact that player 2 will bid his valuation truthfully in the auction and player 1 believes that $\theta_2 \geq \epsilon$. As for case (ii), note that if θ_2 offers ϵ in equilibrium then $\theta_2 \geq \epsilon$; otherwise, he would be better off offering zero.

Therefore, a sufficient condition for acceptance of ϵ is $\theta_1 - \epsilon < \epsilon$, or $\theta_1 < 2\epsilon$. Therefore, θ_2 's expected payoff from the deviation ϵ is bounded from below by

$$G(2\epsilon)(\theta_2 - \epsilon) + (1 - G(2\epsilon)) \int_{2\epsilon}^x M(\epsilon, t) \frac{g(t)}{1 - G(2\epsilon)} dt \quad (2)$$

where $M(\epsilon, t)$ is defined as the minimum utility that player 1 of type $\theta_1 = t$ can impose on player 2 of type θ_2 following the offer ϵ , given that the players bid truthfully in the auction. Specifically, $M(\epsilon, t)$ is given by

$$M(\epsilon, t) = \begin{cases} \min\{\theta_2 - \epsilon, \theta_2 - t\} & \text{if } t \leq \theta_2 \\ 0 & \text{if } t > \theta_2 \end{cases}$$

Since $M(\epsilon, t) \geq 0$ and $x > \theta_2$, the expression in (2) is bounded from below by

$$G(2\epsilon)(\theta_2 - \epsilon) + (1 - G(2\epsilon)) \int_{2\epsilon}^{\theta_2} M(\epsilon, t) \frac{g(t)}{1 - G(2\epsilon)} dt = G(2\epsilon)(\theta_2 - \epsilon) + \int_{2\epsilon}^{\theta_2} (\theta_2 - t)g(t)dt \quad (3)$$

The equality in (3) follows from the fact that in the range $2\epsilon \leq t \leq \theta_2$ we have $M(\epsilon, t) = \min\{\theta_2 - t, \theta_2 - \epsilon\} = \theta_2 - t$, where the latter equality follows from $t \geq 2\epsilon > \epsilon$. Therefore, θ_2 's expected payoff if he deviates to ϵ is bounded from

below by the RHS of (3). Setting $z \equiv 2\epsilon$, we see that his payoff is bounded from below by

$$\psi(z) \equiv G(z)\left(\theta_2 - \frac{z}{2}\right) + \int_z^{\theta_2} (\theta_2 - t)g(t)dt.$$

Note that $\psi(0)$ equals θ_2 's equilibrium payoff.²⁸ Therefore, it must be that ψ is locally non-increasing in a neighborhood of 0. I will now prove that this is not the case. Indeed, the first two derivatives of ψ are zero and the third derivative is positive (all evaluated at 0).

$$\psi'(z) = g(z)\frac{z}{2} - \frac{G(z)}{2} \Rightarrow \psi'(0) = 0$$

$$\psi''(z) = \frac{1}{2}g'(z)z \Rightarrow \psi''(0) = 0$$

$$\psi'''(z) = \frac{1}{2}(g''(z)z + g'(z)) \Rightarrow \psi'''(0) = \frac{g'(0)}{2} = \frac{f_1'(0)}{2F(x)} > 0$$

Thus, Claim 1 is proved— b_1 is strictly increasing on $[0, 1]$.

When θ_1 offers $b_1(\theta_1)$ he reveals his identity. By efficiency, all $\theta_2 > \theta_1$ reject this offer. Consider $\theta_2 < \theta_1$. If he rejects the bribe then the auction results (because of efficiency), which gives him a zero payoff; we conclude then, that on the equilibrium's path, player 2 accepts $b_1(\theta_1)$ if and only if $\theta_2 < \theta_1$. Moreover, types $\theta_2 > \theta_1$ who reject $b_1(\theta_1)$ counter with an offer of their own. To see this, note that such θ_2 's who are sufficiently close to θ_1 prefer $b_1(\theta_1)$ to the non-cooperative auction and therefore they counter with a bribe offer that player

²⁸In this equilibrium player 2 of type θ_2 offers 0 (in this continuation game), which is rejected. This leads to a noncooperative play of the auction against player 1 whose type is distributed according to G , giving the expected payoff $\int_0^{\theta_2} (\theta_2 - t)g(t)dt = \psi(0)$.

1 accepts. Moreover, they all counter with the same bribe. This is the case because they players play pure strategies; two types $\theta_2, \theta'_2 > \theta_1$ cannot offer different counterbribes in equilibrium, because such bribes—if offered on the path—would both be accepted by player 1; this, in turn, means that offering the higher counterbribe is not incentive compatible. Therefore, when θ_1 reveals himself through the bribe $b_1(\theta_1)$, all $\theta_2 < \theta_1$ accept it and all $\theta_2 > \theta_1$ reject it and counter it with a common offer, $\beta(\theta_1)$, which player 1 accepts. Since the threshold type $\theta_2 = \theta_1$ is indifferent between acceptance and rejection,

$$b_1(\theta_1) = \theta_1 - \beta(\theta_1). \quad (4)$$

When player 1 of type θ_1 makes the first move in the game, he can nondetectably mimic any other type x ; therefore, his objective is to maximize the following

$$\Pi(x|\theta_1) = F_2(x)(\theta_1 - b_1(x)) + (1 - F_2(x))\max\{\beta(x), \tilde{\pi}_1(\theta_1, x)\}.$$

By Myerson's result (see subsection 3.1 in the text above), $\Pi(\theta_1|\theta_1) = \pi_1^*(\theta_1) + C^*$, where C^* is some constant. I argue that $C^* = 0$. To see this, note that $\Pi(0|0) = \max\{\beta(0), \tilde{\pi}_1(0, 0)\} = \beta(0)$, and therefore, since $\pi_1^*(0) = 0$, $C^* = \beta(0)$. Then, $C^* = 0$ follows from (4). Combining $\Pi(\theta_1|\theta_1) = \pi_1^*(\theta_1)$ with (4) and $\tilde{\pi}_1(\theta_1, \theta_1) = 0$ we obtain $\beta(\theta_1) = \pi_1^*(\theta_1)$.

Finally, I argue that $\mathbb{E}(\theta_2) \geq \frac{1}{2}$. By Lemma 1, this is equivalent to $\pi_1^*(\theta_1) = \beta(\theta_1) \leq \frac{\theta_1}{2}$ for every θ_1 . To see this, assume by contradiction that there exists some θ_1 for which $\beta(\theta_1) > \frac{\theta_1}{2}$. Consider the case where player 1 of type θ_1 revealed himself through the offer $b_1(\theta_1)$ and it is now player 2's turn to respond, and $\theta_2 > \theta_1$. I argue that 2 has a profitable deviation: to counter with $\frac{\theta_1}{2} + \epsilon$, where $\epsilon > 0$ is sufficiently small, so that $\frac{\theta_1}{2} + \epsilon < \beta(\theta_1)$. To prove that this

deviation is profitable, it is sufficient to prove that it will be accepted by player 1 with certainty. Note that this deviation by player 2 is a detectable deviation: player 1 expects to see the offer $\beta(\theta_1)$, but instead he sees a different offer. Thus, he needs to revise his beliefs, and by assumption he assigns probability 1 to the event $\{\theta_2 \geq \frac{\theta_1}{2} + \epsilon\}$. Thus, if he rejects player 2's offer he expects his payoff in the auction to be no greater than $\theta_1 - (\frac{\theta_1}{2} + \epsilon) = \frac{\theta_1}{2} - \epsilon < \frac{\theta_1}{2} + \epsilon$. Thus, player 2's offer is accepted and hence constitutes a profitable deviation. \square

9 Appendix B: Proof of Theorem 2

The following lemma will be needed for the proof.

Lemma 4. *Under σ^{**} , player 1 does not have an incentive to deviate when he makes the first move in the game; i.e., the bribe $b(\theta_1)$ is ex ante optimal.*

Proof. Clearly there is no reason to offer more than $b(1)$. When type θ_1 mimics a type x , his corresponding expected payoff is

$$\begin{aligned} F_2(x) \left(\int_0^x v(\theta_1, t) \frac{f_2(t)}{F_2(x)} dt - b(x) \right) + (1 - F_2(x))\beta(x) &= \\ &= \int_0^x v(\theta_1, t) f_2(t) dt + (1 - F_2(x))V(x) - b(x). \end{aligned}$$

The derivative of this expression is $f_2(x)(v(\theta_1, x) - v(x, x)) + [(1 - F_2(x))V'(x) - b'(x)]$. Since the bracketed term is zero, $x = \theta_1$ is an optimum. \square

We can now turn to the proof of the theorem.

Proof. Let μ^{**} be a system of beliefs such that (i) when player 2 sees an offer of the form $b(x)$ for some $x \in [0, 1]$, he assigns probability 1 to the event $\{\theta_1 = x\}$, while when he sees an offer strictly greater than $b(1)$ he assigns probability 1

to the event $\{\theta_1 = 1\}$, and (ii) in the information set following the rejection of the first offer b_1 where player 2 offered the counterbribe b_2 , player 1's beliefs are the following: if $b_1 = b(x)$ for some $x \in [0, 1]$ and $b_2 = \beta(x)$, then 1 believes that θ_2 is distributed according to $F_2|_{\{\theta_2 \geq x\}}$; otherwise, he believes that $\theta_2 = 0$.

We need to prove that σ_i^{**} prescribes a best-response for i against σ_j^{**} , in each of i 's information sets (given his beliefs).

Start with player 1. Suppose that he is of type $\theta_1 \in [0, 1]$. The fact that he does not have a profitable deviation when he makes the first move in the game is established in Lemma 4. Next, consider the information set where player 1 responds to player 2's counteroffer b_2 . Suppose first that player 1's initial offer was of the form $b(x)$ for some $x \in [0, 1]$ and player 2 countered with $b_2 = \beta(x)$. Then clearly following σ_1^{**} —by definition—is a best-response. It is easy to see that the response that σ_1^{**} assigns to him is optimal also in the complement of this event.

Consider now player 2 of type $\theta_2 \in [0, 1]$. In his first information set in the game he responds to player 1's offer. If this offer takes the form $b(x)$ for some $x \in [0, 1]$, he infers that $\theta_1 = x$. First, I argue that any counteroffer $b_2 < \beta(x)$ is rejected. To see this, assume by contradiction that there exists a counteroffer $b_2 < \beta(x)$ which is accepted by player 1. This offer is accepted if and only if $\hat{\pi}_1(x, 0) < b_2$, in contradiction to $\hat{\pi}_1(x, 0) \geq \beta(x)$.

Therefore, player 2 has effectively three options: (a) to accept the bribe $b(x)$, (b) to reject and counter with $\beta(x)$, which is accepted by player 1,²⁹ and (c) to compete in the noncooperative auction against type $\theta_1 = x$.

Consider first $\theta_2 < x$. Here, we need to prove that acceptance is a best-response. If player 2 rejects the bribe then the best he can obtain is

²⁹The counterbribe $\beta(x)$ is accepted by player 1 of type $\theta_1 = x$ because $\beta(x) \geq \tilde{\pi}_1(x, x) = 0$.

$\max\{v(\theta_2, x) - \beta(x), 0\} < v(x, x) - \beta(x) = b(x)$, and therefore adhering to σ_2^{**} is a best-response. Consider now $\theta_2 \geq x$. Here, we need to prove that countering with $\beta(x)$ is a best-response. First, note that doing so is weakly better than accepting player 1's offer, and is strictly better if $\theta_2 > x$. This follows from $v(\theta_2, x) - \beta(x) \geq v(x, x) - \beta(x) = b(x)$. Secondly, this is also better than competing the auction, because $v(\theta_2, x) - \beta(x) \geq \tilde{\pi}_2(\theta_2, x)$.³⁰ Next, consider the case where player 2 sees an offer strictly greater than $b(1)$. Since he believes that $\theta_1 = 1$, acceptance of this offer is a best-response. Therefore, (σ^{**}, μ^{**}) is a PBE.

□

10 Appendix C: Deletion of weakly dominated strategies

Condition (A) is satisfied by PBE that survive *iterated deletion of weakly dominated strategies*. Unfortunately, however, this refinement is problematic because of the following reasons. First of all, when deleting weakly dominated strategies “order matters.” A natural remedy to this difficulty is to assume that in every round all the weakly dominated strategies are deleted.³¹ Even if we adopt this maximal deletion approach, one serious problem still remains: it is not known how the entire deletion process looks like, and it is therefore unknown whether there exist strategy profiles that survive this deletion. I

³⁰To see that $v(\theta_2, x) \geq \tilde{\pi}_2(\theta_2, x) + \beta(x)$, note (i) this inequality holds at $\theta_2 = x$, because $\tilde{\pi}_2(x, x) = 0$ and $v(x, x) = V(x) = \int_0^x V'(t)dt \geq \int_0^x F_2(t)V'(t)dt = \beta(x)$, and (ii) $\frac{\partial}{\partial \theta_2} v(\theta_2, x) \geq \frac{\partial}{\partial \theta_2} \tilde{\pi}_2(\theta_2, x)$.

³¹This is the common approach to deletion of weakly dominated strategies. See, for example, Fudenberg and Tirole (1991, p. 461).

therefore consider the following weaker version of this refinement—I will consider a single round of deletion. That is, with i 's set of strategies in the game denoted by Σ_i , keeping the entire Σ_i fixed, I delete all the weakly dominant strategies for j ; the resulting set is denoted Σ_j^1 . Let G denote the original game—the one with strategy sets Σ_i , and let G' denote the game which is the same as G , except that the strategy sets are Σ_i^1 .

Lemma 5. *If (σ, μ) is a PBE of G' then it is an (A)-satisfying PBE of G .*

Proof. First, I will prove that any strategy that instructs player 2 to counter an offer with a bribe which is greater than his valuation is deleted, and is therefore outside Σ_2^1 . Consider a strategy σ_2 that instructs player 2 to counter b^* with some $x > \theta_2$ for some b^* and θ_2 . Let σ'_2 be identical to σ_2 except that it instructs player 2 to counter b^* with zero whenever σ_2 instructs him to counter with an offer which exceeds his valuation. Clearly σ'_2 does at least as good as σ_2 against any strategy that player 1 may play. Consider the strategy for player 1 where all θ_1 offer b^* and accept any counterbribe (including zero) if b^* is rejected. Clearly, σ'_2 does strictly better than σ_2 against this strategy. Therefore, player 2 does not have strategies in G' that instruct him to counter with offers which exceed his type. In particular, if (σ, μ) is a PBE of G' then μ assigns zero probability to such behavior (in any information set). Therefore, if this is a PBE in G , then it is an (A)-satisfying PBE of G . I argue that indeed (σ, μ) is a PBE of G . To see this, assume by contradiction that some player i has a profitable deviation in some information set. Clearly we can assume that the deviation is to an undominated strategy and is therefore feasible in G' . But this contradicts the assumption that (σ, μ) is a PBE in G' . \square

Next, I prove that the efficient monotonic equilibria which were studied in the paper survive one round of deletion.

Lemma 6. *There exist beliefs such that (σ^*, μ) is a PBE of G' .*

Proof. Since there exist beliefs μ such that (σ^*, μ) is a PBE of G , it is enough to prove that $\sigma_i^* \in \Sigma_i^1$ for each player i .

Start with $i = 1$. Suppose that σ_1^* is weakly dominated by some other strategy σ_1 . First, I argue that σ_1 prescribes the same bribing function as σ_1^* . To see this, let b_1 be the bribing function under σ_1^* and let B_1 be the one under σ_1 . Let $\theta_1 \in [0, 1]$. Let $x \equiv b_1(\theta_1)$ and $y \equiv B_1(\theta_1)$. I argue that $x = y$. Consider the following strategy of player 2: the bribe x is rejected (by all θ_2) and followed by the counteroffer 1, and any other bribe is rejected and countered with zero. Since σ_1 needs to do at least as good as σ_1^* against any possible strategy of player 2, $x = y$. Next, I argue that σ_1 instructs the same acceptance policy of counterbribes, and therefore it coincides with σ_1^* . To see this, suppose now that player 1 employs the bribing function b_1 , offers $\theta_1 - \pi_1^*(\theta_1)$, and his offer is rejected and countered with b . If $b = \pi_1^*(\theta_1)$ then σ_1 must prescribe acceptance, because player 2 may be playing σ_2^* . If $b \neq \pi_1^*(\theta_1)$ then σ_1 must assign acceptance if and only if $2b > \theta_1$, because it may be the case that player 2 plays a strategy that instructs him to offer his valuation following a rejection. Therefore, there does not exist a strategy that weakly dominates σ_1^* .

Consider now $i = 2$. Bribes greater than $1 - \pi_1^*(1)$ must be accepted because it may be the case that player 1 plays the following strategy: all $\theta_1 < 1$ offer 0 and $\theta_1 = 1$ offers the aforementioned bribe, and he rejects all counterbribes. Since player 1 maybe playing σ_1^* , it follows from the proof of Theorem 1 that if σ_2 does at least as good against σ_1^* as σ_2^* , then the responses it prescribes to offers in $[0, 1 - \pi_1^*(1)]$ coincide with the ones prescribed by σ_2^* . \square

11 Appendix D: Additional proofs

Proof of Proposition 1: Consider player 1. The expected payoff corresponding to a deviation $x \leq \theta_1$ is $x(\theta_1 - x) + (1 - x)\left(\frac{\theta_1 - x}{1 - x}\right)(\theta_1 - \frac{\theta_1 + x}{2}) = \frac{\theta_1^2}{2} - \frac{x^2}{2}$, which is clearly maximized at $x = 0$. It is straightforward that a deviation to $x > \theta_1$ cannot be profitable. Next, consider player 2 who needs to respond to a bribe $b \leq \theta_2$. Assume by contradiction that he reacts to it with a deviation to a positive counteroffer, $x \in (0, 1]$. Let player 1's beliefs be such that he assigns probability 1 to $\{\theta_2 = x\}$, hence accepting x if and only if $\theta_1 \leq 2x$. Thus, we can restrict our attention to deviations $x \in (0, \frac{1}{2}]$; each such x gives an expected payoff of $2x(\theta_2 - x) + (1 - 2x)\left(\frac{\theta_2 - 2x}{1 - 2x}\right)(\theta_2 - \frac{\theta_2 + 2x}{2}) = \frac{\theta_2^2}{2}$, so adhering to the equilibrium strategy is optimal. Therefore, σ^N can be supported in equilibrium. \square

Proof of Proposition 2: Make the assumptions of the proposition. The refinement (A) does not impose any restriction on player 2's beliefs, so we can choose any belief we like in order to support his acceptance rule as optimal when he sees an unexpected offer $\beta \neq b$. Thus, when player 1 makes a detectable deviation β , he obtains the expected payoff

$$F_2(\beta)(\theta_2 - \beta) + \int_{\beta}^{\theta_1} (\theta_2 - t)f_2(t)dt.$$

The derivative of this expression is $-F_2(\beta)$ and the corresponding value at $\beta = 0$ is the noncooperative payoff, $\pi_1^*(\theta_1)$. Thus, it is sufficient to prove that following the strategy gives a payoff weakly greater than the noncooperative payoff. The expected payoff from following the strategy is $F_2(y)[\theta_1 - b] + (1 - F_2(y))c$. Substituting $b = y - c$ gives that this expected payoff is $h(\theta_1) \equiv F_2(y)(\theta_1 - y) + c$. Thus, it suffices to prove that $h(\theta_1) \geq \pi_1^*(\theta_1)$ for all $\theta_1 \in$

$[0, 1]$. Since both are increasing, h is linear and π_1^* is convex, it is sufficient to show that $h(x) \geq \pi_1^*(x)$ for $x \in \{0, 1\}$. The inequality $h(0) \geq \pi_1^*(0)$, or $-F_2(y)y + c \geq 0$, is guaranteed by (II); the inequality $h(1) \geq \pi_1^*(1)$ is guaranteed by (III). Next, consider player 1 who faces the offer c . By (II) it is optimal for type $\theta_1 = 1$ to accept c , and therefore it is optimal for every $\theta_1 \in [0, 1]$.

Next, consider player 2. It will be convenient to define

$$\Phi(\beta, \theta_2) \equiv F_1(2\beta)(\theta_2 - \beta) + \int_{2\beta}^{\theta_2} (\theta_2 - t)f_1(t)dt.$$

We need to establish that his response to the on-path offer b is optimal. Suppose first that $\theta_2 < y$. Assume by contradiction that he has a strictly profitable deviation. This deviation involves rejection and a counteroffer β . It must be that $\beta \neq c$, because the payoff corresponding to $\beta = c$ is $\theta_2 - c < y - c = b$. Thus, the expected payoff corresponding to the deviation is $\Phi(\beta, \theta_2) \leq \Phi(\beta, y) \leq b$, where the first inequality follows from $\frac{\partial}{\partial \theta_2} \Phi(\beta, \theta_2) = F_1(\theta_2) \geq 0$. Suppose, on the other hand, that $\theta_2 > y$. Clearly acceptance is sub-optimal. Suppose then that there exists a profitable deviation to a counteroffer $\beta \neq c$. But this gives the payoff $\Phi(\beta, \theta_2)$, which obeys

$$\Phi(\beta, \theta_2) = \Phi(\beta, y) + \int_y^{\theta_2} \frac{\partial}{\partial t} \Phi(\beta, t) dt = \Phi(\beta, y) + \int_y^{\theta_2} F_1(t) dt \leq (y-c) + (\theta_2 - y) = \theta_2 - c,$$

because $\Phi(\beta, y) \leq y - c$ by (I), in contradiction to the profitability of the deviation. \square

Proof of Proposition 3: I will prove that there exist an $\underline{\alpha} \in (0, 1)$ such that the following is true: If $F_1(t) = F_2(t) = t^\alpha$ for some $\alpha \in [\underline{\alpha}, 1]$, then (F_1, F_2) satisfies (I)-(III) with respect to y and c , for some $y, c \in [0, 1]$. Recall $\Phi(\beta, \theta_2)$ from the proof of Proposition 2. For such F_i , $\Phi(\beta, y) = \frac{y^{1+\alpha}}{1+\alpha} + 2^\alpha \beta^{1+\alpha} \left[\frac{2\alpha}{1+\alpha} - 1 \right]$. For $\alpha \in (0, 1]$ the bracketed term is nonpositive, so (I) is implied by

$$y - c \geq \frac{y^{1+\alpha}}{1+\alpha}. \quad (5)$$

The conditional expectation $\mathbb{E}(\theta_2|\theta_2 \geq y)$ is given by $\mathbb{E}(\theta_2|\theta_2 \geq y) = \frac{\alpha}{1+\alpha} \cdot \frac{1-y^{1+\alpha}}{1-y^\alpha}$. Thus, the first inequality that (II) requires is

$$c > 1 - \frac{\alpha}{1+\alpha} \cdot \frac{1-y^{1+\alpha}}{1-y^\alpha} = \frac{1-(1+\alpha)y^\alpha + \alpha y^{1+\alpha}}{(1+\alpha)(1-y^\alpha)} \quad (6)$$

The second inequality that (II) requires is simply

$$c > y^{1+\alpha}. \quad (7)$$

Finally, (III) is

$$y^\alpha(1-y) + c \geq \frac{1}{1+\alpha}. \quad (8)$$

Our task is to find $y, c \in [0, 1]$ that satisfy (5)-(8) for all $\alpha > 0$ sufficiently close to 1.

Take $y = (\frac{1}{2})^{\frac{1}{\alpha}}$ and $c = \frac{2}{3}y$.

First, note that $y^{1+\alpha} = y \cdot [(\frac{1}{2})^{\frac{1}{\alpha}}]^\alpha = \frac{1}{2}y < \frac{2}{3}y = c$, so (7) holds. Next, note that plugging this choice of y and c into (5) gives $\frac{1}{3}y \geq \frac{y^{1+\alpha}}{1+\alpha} \Leftrightarrow \frac{1}{3} \geq \frac{\frac{1}{2}}{1+\alpha}$, and the latter is satisfied as a strict inequality at $\alpha = 1$. Similarly, (6) also boils down to $\frac{1}{3} \geq \frac{1}{4}$ at $\alpha = 1$. Finally, (8) becomes $\frac{1}{2} + \frac{y}{6} \geq \frac{1}{1+\alpha}$, which obviously holds for all α near 1. \square

Proof of Proposition 4: In view of the argument in the beginning of the proof of Proposition 2, it is enough to show that under the equilibrium which is described in Proposition 4 player 1's payoff is weakly greater than $\pi_1^*(\theta_1)$.

Given a type θ_2 let $\pi_1^*(\theta_1|\theta_2)$ be player 1's payoff in the noncooperative auction against θ_2 ; similarly, given a set of types of player 2, A , let $\pi_1^*(\theta_1|A)$ be player 1's noncooperative payoff conditional on $\theta_2 \in A$. Let b^{ES} be the bribing function in the ES equilibrium and let $\bar{\theta}_2$ be the unique point from which onwards it is constant. Player 1's expected payoff from the behavior described in Proposition 4 is

$$F_2(\bar{\theta}_2) \underbrace{\int_0^{\bar{\theta}_2} \max\{b^{ES}(t), \pi_1^*(\theta_1|t)\} \frac{f_2(t)}{F_2(\bar{\theta}_2)} dt}_X + (1 - F_2(\bar{\theta}_2))b^{ES}(\bar{\theta}_2)$$

Obviously, $X \geq \pi_1^*(\theta_1|\theta_2 \leq \bar{\theta}_2)$; thus, it is enough to show that $b^{ES}(\bar{\theta}_2) \geq \pi_1^*(\theta_1|\theta_2 \geq \bar{\theta}_2)$. Note that since type $\theta_1 = 1$ accepts $b^{ES}(\bar{\theta}_2)$ in the ES equilibrium, $b^{ES}(\bar{\theta}_2) \geq \pi_1^*(1|\theta_2 \geq \bar{\theta}_2)$. The proof is now completed by the fact that $\pi_1^*(1|\theta_2 \geq \bar{\theta}_2) \geq \pi_1^*(\theta_1|\theta_2 \geq \bar{\theta}_2)$. \square

Proof of Proposition 5: Suppose that player 2 employs the (increasing) bidding function b . When player 1 of type θ_1 mimics submit the bid of type $\theta_2 = x$, his expected payoff is

$$F_2(x)(\theta_1 + \int_0^x \phi(t) \frac{f_2(t)}{F_2(x)} dt) - \int_0^x b(t) \frac{f_2(t)}{F_2(x)} dt.$$

If $x = \theta_1$ is an optimum then $b(t) = t + \phi(t)$. The payoff for type θ_1 from adhering to this bidding function is therefore

$$F_2(\theta_1)\theta_1 - \int_0^{\theta_1} [\phi(t) - b(t)]f_2(t)dt = F_2(\theta_1)\theta_1 - \int_0^{\theta_1} t f_2(t)dt = \pi_1^*(\theta_1). \quad \square$$

Proof of Theorem 3: To show that type θ_1 's equilibrium payoff equals his noncooperative payoff we need to show

$$F_2(\theta_1)\theta_1 + \int_0^{\theta_1} \phi(t)f_2(t)dt + (1 - F_2(\theta_1))(\theta_1 + \phi(\theta_1)) - b(\theta_1) = \int_0^{\theta_1} F_2(t)dt.$$

Note that both sides equal zero at $\theta_1 = 0$. The derivative of the LHS (wrt θ_1) is $f_2(\theta_1)\theta_1 + F_2(\theta_1) + \phi(\theta_1)f_2(\theta_1) - f_2(\theta_1)(\theta_1 + \phi(\theta_1)) + (1 - F_2(\theta_1))(1 + \phi'(\theta_1)) - b'(\theta_1) = F_2(\theta_1)$, which equals that of the RHS. \square

12 References

- [1] Aoyagi, M. (2002), “Bid rotation and collusion in repeated auctions,” *Journal of Economic Theory*, **112** 79-105.
- [2] Aoyagi, M. (2007), “Efficient collusion in repeated auctions with communication,” *Journal of Economic Theory*, **134**, 61-92.
- [3] Athey, S., and Bagwell, K. (2001), “Optimal collusion with private information,” *RAND Journal of Economics*, **32**, 428-465.
- [4] Athey, S., and Bagwell, K. (2008), “Collusion with persistent cost shocks,” *Econometrica*, **76**, 493-540.
- [5] Athey, S., and Bagwell, K. (2001), “Optimal collusion with private information,” *RAND Journal of Economics*, **32**, 428-465.
- [6] Athey, S., and Bagwell, K. (2008), “Collusion with persistent cost shocks,” *Econometrica*, **76**, 493-540.

- [7] Baldwin, L.H., Marshall, R.C., and Richard, J.-F. (1997), “Bidder collusion at forest service timber sales,” *Journal of Political Economy*, **105**, 657-699.
- [8] Blume, A., and Heidhues, P. (2006), “Private monitoring in auctions,” *Journal of Economic Theory*, **131**, 179-211.
- [9] Cassady, R. (1967), “Auction and auctioneering,” University of California Press.
- [10] Che, Y-K., and Kim, J. (2009), “Optimal collusion-proof auctions,” *Journal of Economic Theory*, **144**, 565-603.
- [11] Chen, C-L., and Tauman, Y. (2006), “Collusion in one-shot second price auctions,” *Economic Theory*, **28**, 145-172.
- [12] Cramton, P., Gibbons, R., and Klemperer, P. (1987), “Dissolving a partnership efficiently,” *Econometrica*, **55**, 615-632.
- [13] Dequiedt, V. (2007), “Efficient collusion and optimal auctions,” *Journal of Economic Theory*, **136**, 302-323.
- [14] Esö, P., and Schummer, J. (2004a), “Bribing and signaling in second price auctions,” *Games and Economic Behavior*, **47**, 299-324.
- [15] Fudenberg, D., and Tirole, J. (1991), *Game Theory*, MIT Press, Cambridge, MA.

- [16] Graham, D.A., and Marshall, R.C. (1987), “Collusive bidder behavior at single-object second-price and english auctions,” *Journal of Political Economy*, **95**, 1217-1239.
- [17] Hörner, J., and Jamison, J. (2007), “Collusion with (almost) no information,” *RAND Journal of Economics*, **38**, 804-822.
- [18] Kivetz, G., and Tauman, Y. (2010), “Simple collusive agreements in one-shot first-price auctions,” *Games and Economic Behavior*, **69**, 138-149.
- [19] Mailath, G.J., and Zemsky, P. (1991), “Collusion in second price auctions with heterogeneous bidders,” *Games and Economic Behavior*, **3**, 467-486.
- [20] Marshall, R.C., and Marx, L.M. (2007), “Bidder collusion,” *Journal of Economic Theory*, **133**, 374-402.
- [21] Maskin, E., and Tirole, J. (1990), “The principal-agent relationship with an uninformed principal: the case of private values,” *Econometrica*, **58**, 379-409.
- [22] McAfee, P.R., and McMillan, J. (1992), “Bidding rings,” *American Economic Review*, **82**, 579-599.
- [23] Myerson, R.B. (1981), “Optimal auction design,” *Mathematics of Operations Research*, **6**, 58-73.
- [24] Myerson, R.B. (1983), “Mechanism design by an informed principal,”

Econometrica, **51**, 1767-1798.

[25] Myerson, R.B., and Satterthwaite, M.A. (1983), "Efficient mechanism for bilateral trading," *Journal of Economic Theory*, **29**, 265-281.

[26] Pavlov, G. (2008), "Auction design in the presence of collusion," *Theoretical Economics*, **3**, 383-429.

[27] Porter, R., and Zone, D. (1993), "Detection of bid rigging in procurement auctions," *Journal of Political Economy*, **101**, 518-538.

[28] Rachmilevitch, S. (2009), "Endogenous bid-rotation in repeated auctions," Working paper.

[29] Rachmilevitch, S. (2010), "Bribing and signaling in auctions," job-market paper.

[30] Rachmilevitch, S. (2011), "Bribing in first-price auctions," Working paper.

[31] Severinov, S. (2008), "An efficient solution to the informed principal problem," *Journal of Economic Theory*, **141**, 114-133.

[32] Skrzypacz, A., and Hopenhayn, H. (2004), "Tacit collusion in repeated auctions," *Journal of Economic Theory*, **114**, 153-169.