Columbia University

Department of Economics
Discussion Paper Series

Optimal Collusion-Proof Auctions

Yeon-Koo Che
Jinwoo Kim

Discussion Paper No.: 0708-05

Department of Economics
Columbia University
New York, NY 10027

September 2007

Optimal Collusion-Proof Auctions

Yeon-Koo Che* Jinwoo Kim†

September 6, 2007

Abstract: We study an optimal collusion-proof auction in an environment where subsets of bidders may collude not just on their bids but also on their participation. Despite their ability to collude on participation, informational asymmetry facing the potential colluders can be exploited significantly to weaken their collusive power. The second-best auction — i.e., the optimal auction in a collusion-free environment — can be made collusion-proof, if at least one bidder is not collusive, or there are multiple bidding cartels, or the second-best outcome involves a nontrivial probability of the object not being sold. In case the second-best outcome is not weak collusion-proof implementable, we characterize an optimal collusion-proof auction. This auction involves nontrivial exclusion of collusive bidders — i.e., the object is not sold to any collusive bidder with positive probability.

Keywords: Collusion on participation, subgroup collusion, multiple bidding cartels, an exclusion principle.

*Columbia University, Economics Department. Email: yk.che@columbia.edu.
†Yonsei University, School of Economics, Email: jikim@yonsei.ac.kr.

Both authors are grateful to Gorkem Celik and Larry Samuelson for helpful comments. Yeon-Koo Che acknowledges financial support from the Shoemaker Faculty Fellowship at the University of Wisconsin, and the hospitality from the Yonsei University. Jinwoo Kim acknowledges the hospitality from the Columbia University and University of Wisconsin during his visits.

1 Introduction

Collusion by participants often poses a serious threat to markets and organizations. Nowhere is such a threat more intimately felt than in auctions where bidders can manipulate or simply withdraw their bids to limit competition. Not surprisingly, auctions have provided the volume and prominence to the study of collusion, with its evidence found in highway construction contracts (Porter and Zona (1993)), timber sales (Baldwin et al. (1997)), and in school milk delivery contracts (Pesendorfer (2000), and Porter and Zona (1999)).

In keeping with these evidences are the theoretical findings that "standard" auctions are vulnerable to bidder collusion, even when the cartel members face mutual asymmetric information. McAfee and McMillan (1992) demonstrate in their seminal article that asymmetrically informed cartel members can structure a knock-out auction that enables them to (re)allocate the good among themselves efficiently while limiting the seller's revenue to at most her reserve price. The ability by the cartel members to exchange side payments (without getting detected) is crucial for this result, but they can achieve the same effect via adjusting their bid rotation or market shares, if auctions are repeated.[1]

If standard auctions are vulnerable to collusion, can one find an auction rule that is not? This is the question we pursue in this paper. What makes this question nontrivial is the informational asymmetry facing potential colluders. If colluders have complete information about one another, then they can effectively act like a single agent and maximize their joint payoffs. Then, there would be little room for auction design, for an optimal scheme would simply reduce to textbook monopoly pricing. If bidders face mutual asymmetric information, however, they may not effectively coordinate their behavior, and the seller may exploit this to undermine collusion. Although standard auctions are not capable of this (as has been shown by extant literature), other auction rule may enable the seller to exploit the bidders' mutual asymmetric information more effectively. We seek to identify such an auction rule.

In order to study an optimal response to collusion, one must understand a bidder's incentive to participate in collusion. In particular, one must deal with the question of what happens after a bidder refuses to participate in collusion. What belief would they form about the subsequent competition and about the types of bidders they face? Can the remaining cartel members punish the defecting bidder, and if so, to what degree? How one models the (out-of-equilibrium) belief and the cartel members' ability to punish a defector determines a bidder's

[1] Graham and Marshall (1987), McAfee and McMillan (1992), Mailath and Zemsky (1991), Eso and Schummer (2003), and Marshall and Marx (2007) study collusion in one-shot auctions of various formats, while Aoyagi (2003), Athey, Bagwell and Sanchirico (2004), Blume and Heidhues (2002), and Skrzypacz and Hopenhayn (2004) study collusion in repeated auctions.

incentive to participate in a collusive arrangement, which in turn determines the scope and the nature of the seller's response. In this regard, we take an eclectic approach by considering both weak and strong notions of collusion-proof auctions.

The weak notion postulates that collusion arises *only* when it benefits all types of bidders relative to a non-cooperative play *without any updating of beliefs*. An auction rule is said to be *weak collusion-proof* if it admits no such collusion. This notion is reasonable to the extent that cartel members will often find it difficult to punish a defector more severely than bidding competitively. At the same time, the weak notion restricts the bidders' out-of-equilibrium beliefs. Hence, we also consider a "strong" notion, which imposes no restriction on the cartel members' out-of-equilibrium beliefs or their ability to punish a defector: an auction rule is defined to be *strong collusion-proof* if, under that rule, the seller enjoys in *every* Bayesian Nash equilibrium the same expected revenue as she would absent any collusion. Further, we require the strong collusion-proof auction to be robust to the specifics of the cartel operation.

While the alternative notions matter to some extent, they do not affect the main thrust of our results. We find, largely irrespective of the particular notion used, that a seller can overcome her vulnerability to collusion in a surprisingly broad range of circumstances. Specifically, a seller can attain the highest revenue she can without any collusion, either if a cartel is not all-inclusive or if the object is not sold to any bidder with some probability. This result holds with respect to the weak notion of collusion-proofness but also with respect to the strong notion given an additional condition (which is satisfied for the case of the all-inclusive cartel). When neither condition is met, we identify an optimal collusion-proof auction in some restricted class of auctions. In the process, we establish *an exclusion principle* which states that an optimal collusion-proof auction involves a positive probability of not selling to any collusive bidder. The exclusion principle holds quite generally, regardless of the buyers' support of valuations, thus exhibiting a qualitative departure from the collusion-free auction design.[2] In sum, the present paper suggests that a seller can overcome collusion completely in many cases and do generally much better than she could if she would resort to simple monopoly pricing.

The current paper is related to several recent papers.[3] First, it is related to our companion

[2]As is well known, the standard optimal auction allocates the good to a buyer with probability one, if there is at least one buyer whose infimum valuation is sufficiently high. See Myerson (1981).

[3]Other authors have studied optimal collusion-proof mechanisms in different contexts. Quesada (2004) finds an optimal collusion-proof mechanism in the LM setting where an (informed) agent proposes a side contract. In fact, she treats collusion on participation, but adopts the strong notion similar to Dequiedt (2007) where the side contract can impose maximum punishment on refusing agents. Jeon and Menicucci (2005) shows that second-best is achievable in the weak collusion-proof sense, much like Che and Kim (2007), in the nonlinear

paper, Che and Kim (2007) (henceforth, CK), as well as Laffont and Martimort (1999, 2000) (henceforth, LM), which studies a collusion-proof contract when, unlike the current setting, agents can collude only after they participate in the contract. This latter assumption may not be appropriate in many auctions where bidders are intimately familiar with their opponents even before participating. In fact, an allegedly predominant form of collusion involves bidders coordinating on their participation decisions: Colluders either refuse to participate or withdraw their bids to let a designated cartel member win without facing competition. Further, the idea of "selling the firm" to potential colluders, featured in Che and Kim (2007), relies on the agents' inability to collude on their participation decision. The current paper relaxes the assumption by allowing the bidders to collude on their participation decision.

In this latter respect, the current paper is related to Dequiedt (2007) and Pavlov (2006), who also study collusion-proof auctions when bidders can collude on their participation. Dequiedt considers two bidders with binary types (i.e., of either low or high valuation). He shows that, if a cartel can commit to punish a defector to his reservation utility, then the seller can at most collect her reserve price when a bidder's valuation exceeds that price. This seemingly pessimistic result, however, has more to do with the binary type of the model than say with the bidders' ability to collude on participation or their ability to punish a defector (which in fact corresponds to the strong notion in the current paper). With binary types, exclusion of a low valuation type is equivalent to setting a reserve price equal to the high valuation type. Hence, the optimality of this latter behavior, although completely explainable by our exclusion principle, may appear to suggest that the seller can't do any better than adopting a text-book monopoly pricing. Our paper will show that the optimal collusion-proof auction does not generally reduce to monopoly pricing, and will typically generate strictly higher revenue.

Like us, Pavlov (2006) considers a model with a continuous bidder types, and his main result parallels some of the current paper, particularly Theorem 5. There are several differences, however. His analysis concerns only the case of the all-inclusive cartel and focuses only on ex ante symmetric bidders.[4] The current paper goes much beyond that environment. First of all, we consider the general case in which subsets of bidders are collusive. In fact, the most important result concerns the case in which a proper subset of bidders is collusive — i.e., at least one bidder is noncollusive or there are multiple bidding cartels — in which case the second-best outcome is shown to be collusion-proof implementable. Second, we can

pricing context. Laffont and Martimort (1998), Celik (2004) and Faure-Grimaud et al. (2004) study the optimality of delegation in the presence of collusive agents.

[4]Our results for the all-inclusive cartel case are obtained independently. As will be apparent, the methods of analysis are quite different. Of course, the results on other subjects, particularly the subgroup collusion, are completely new here.

handle the case of ex ante asymmetric bidders, at least with the all-inclusive cartel: We show that the second-best outcome is collusion-proof implementable, given a somewhat stronger condition than is needed for the symmetric bidders case. Third, our model of collusion differs from his in that we allow members of collusion to reallocate the good once it is sold to one of the members.

Above all, the ability to handle collusion by a subset(s) of bidders is practically important and useful. In many circumstances, not all bidders are in a position to collude. Government auctions used in defense procurement, mineral extraction, or spectrum licenses often have incumbents with long history of operation competing against relative new comers. Long term interaction and shared experiences among the incumbents will put them in a better position to collude than the new comers. Likewise, in auctions for construction repairs or food service procurement, competition may involve both local and non-local providers, and the former group may be able to collude more effectively, based on their regular contacts and their interaction through trade associations. The problem of only a subset of bidders being collusive introduces a new challenge, since the cartel may prey on noncollusive bidders as much as on the seller. Hence, a collusion-proof design must eliminate incentives for the cartel to engage in such behavior.

The rest of the paper is organized as follows. Section 2 introduces an auction model and describes the second-best outcome in a collusion-free environment. Section 3 introduces a model of collusion and the notion of weak collusion-proof auctions. Section 4 identifies the properties of weak collusion-proof auctions and uses them to derive a condition that is necessary and sufficient for implementing the second-best outcome in a collusion-proof fashion. In Section 5, we characterize the optimal collusion-proof auction when the second-best is not weak collusion-proof implementable. Section 6 characterizes strong collusion-proof implementation. Section 7 concludes.

2 Primitives

A risk-neutral seller has an object for sale. The seller's valuation of the object is normalized to zero. There are $n \geq 2$ risk-neutral buyers who each independently draw a value, θ_i, on the object from an interval $\Theta_i := [\underline{\theta}_i, \overline{\theta}_i] \subset \mathbb{R}_+$ according to distribution F_i, which has strictly positive density f_i on the support. We assume that both

$$J_i(\theta_i) := \theta_i - \frac{1 - F_i(\theta_i)}{f_i(\theta_i)} \text{ and } K_i(\theta_i) := \theta_i + \frac{F_i(\theta_i)}{f_i(\theta_i)}.$$

are strictly increasing in θ_i for all $i \in N$. Throughout, we let $\mathbb{E}[\cdot] := \int_{\Theta}[\cdot]d(\prod_{i\in N} F_i(\theta_i))$. and $\mathbb{E}_{\tilde{\theta}_{-i}}[\cdot] := \int_{\Theta_{-i}}[\cdot]d(\prod_{j\neq i} F_j(\theta_j))$ denote expectation operators based on the prior distribution, where $\Theta := \prod_{i\in N}\Theta_i$ and $\Theta_{-i} := \prod_{j\neq i}\Theta_j$.

For a later analysis, it is convenient to augment each bidder's type space to include the "participation decision" as part of his possible type. Specifically, we let $\theta_\emptyset$ denote "non-participation" or "exit" option available to each bidder with the convention that $\theta_\emptyset < \theta_i, \forall i \in N, \forall \theta_i \in \Theta_i$, and define $\overline{\Theta}_i := \{\theta_\emptyset\} \cup \Theta_i$. We then let $\theta := (\theta_1, ..., \theta_n) \in \overline{\Theta} := \prod_{i\in N}\overline{\Theta}_i$ denote a possible profile of types in these enriched type spaces. Since we shall consider randomization in cartel members' reports over their augmented type spaces, it is convenient to consider arbitrary probability distribution, μ^C, over $\overline{\Theta}_C := \prod_{i\in C}\overline{\Theta}_i$ for any $C \subset N$ and to use $\mathbf{E}_{\mu^C}[\cdot] := \int_{\overline{\Theta}_C}[\cdot]d(\mu^C(\theta_C))$ as an expectation operator relative to μ^C.

We now describe arbitrary auction rules, and we do so in direct mechanisms. An *auction rule*, $M = (q,t)$, consists of *an allocation rule*, $q = (q_1, \cdots, q_n) : \overline{\Theta} \to \mathcal{Q}$, where $\mathcal{Q} := \{x \in [0,1]^n | \sum_{i\in N} x_i \leq 1\}$ and *a payment rule*, $t = (t_1, ..., t_n) : \overline{\Theta} \to \mathbb{R}^n$, such that $q_i(\theta_\emptyset, \theta_{-i}) = t_i(\theta_\emptyset, \theta_{-i}) = 0, \forall i, \theta_{-i} \in \overline{\Theta}_{-i}$. An auction rule determines, for each profile of bidders' reports in $\overline{\Theta}$, a vector of probabilities for the bidders to obtain the object and a vector of expected payments they must pay, subject to the constraint that, if a bidder does not participate, he does not receive the good and collects his reservation utility, normalized to zero. Any equilibrium arising in any auction game can be described as an auction rule in this framework, so we sometimes use an "outcome" interchangeably with an auction rule.

Fix an auction rule, $M = (q,t)$. Buyer i's interim payoff when his valuation is $\theta_i \in \Theta_i$ but reports $\tilde{\theta}_i \in \overline{\Theta}_i$ is

$$u_i^M(\tilde{\theta}_i, \theta_i) := \theta_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i),$$

where $Q_i(\theta_i) := \mathbb{E}_{\tilde{\theta}_{-i}}[q_i(\theta)]$ and $T_i(\theta_i) := \mathbb{E}_{\tilde{\theta}_{-i}}[t_i(\theta)]$. Given hidden information and the availability of the non-participation option, an auction rule must be incentive compatible and individually rational to be consistent with equilibrium. We say an auction rule M is *feasible* if

$$U_i^M(\theta_i) := u_i^M(\theta_i, \theta_i) \geq u_i^M(\tilde{\theta}_i, \theta_i), \quad \forall i, \theta_i \in \Theta_i, \tilde{\theta}_i \in \overline{\Theta}_i. \qquad (IC^*)$$

Note (IC^*) subsumes both incentive compatibility and individual rationality, since it requires

$$U_i^M(\theta_i) \geq u_i^M(\theta_\emptyset, \theta_i) = 0. \qquad (IR)$$

Let $\mathcal{M}$ denote the set of all feasible auction rules. For later analysis, the following characterization of feasible auction rules proves useful. Its proof, along with most of the others, are relegated to the Appendix.

LEMMA 0. *If $M = (q, t) \in \mathcal{M}$, then, for each $\theta_i \in \Theta_i$,*

$$U_i^M(\theta_i) = \mathbb{E}\left[K_i(\tilde{\theta}_i)q_i(\tilde{\theta})1_{\{\tilde{\theta}_i \leq \theta_i\}} + J_i(\tilde{\theta}_i)q_i(\tilde{\theta})1_{\{\tilde{\theta}_i \geq \theta_i\}} - t_i(\tilde{\theta})\right]. \tag{1}$$

Before proceeding, it is useful to consider a collusion-free environment. It is by now well known that, in such an environment, an optimal auction rule, called *second-best* or *noncollusive optimal outcome*, solves

$$[NC] \qquad \max_{M \in \mathcal{M}} \mathbb{E}\left[\sum_{i \in N} t_i(\theta)\right],$$

and its associated outcome is characterized as follows:

THEOREM 0. (MYERSON) *An optimal mechanism that solves $[NC]$ involves the allocation rule given by $\forall \theta \in \Theta$,*

$$q_i^*(\theta) = \begin{cases} 1 & \text{if } J_i(\theta_i) > \max\{0, \max_{k \neq i} J_k(\theta_k)\}, \\ 0 & \text{otherwise}, \end{cases}$$

and yields revenue of

$$V^* := \mathbb{E}\left[\sum_{i \in N} J_i(\theta_i)q_i^*(\theta)\right]$$

to the seller.

We assume that $\mathbb{E}[q_i^*(\theta)] < 1$ for each $i \in N$, or else the optimal mechanism reduces to bargaining with a single buyer, so there would be no problem of collusion. Letting $\hat{\theta}_i := \min\{\theta_i \in [\underline{\theta}_i, \bar{\theta}_i] | J_i(\theta) \geq 0\}$, the optimal mechanism allocates the good to the bidder with the highest virtual valuation $J_i(\theta_i)$ as long as $\theta_i \geq \hat{\theta}_i$. In particular, the object is sold with probability one if there is a bidder i such that $J_i(\underline{\theta}_i) > 0$.

3 A Model of Collusion

We here develop our model of collusion geared to the weak notion of collusion-proof auctions, and later we introduce the strong notion by relaxing some of the assumptions. To this end, we follow LM and CK and suppose that there are subsets of bidders, called *cartels*, that enforce side contracts via uninformed representatives to influence the outcome of the auction game being played. Formally, *a cartel structure* is an arbitrary partition $\mathcal{C}$ on N whose element $C \in \mathcal{C}$ represents a cartel of bidders who "may" collude with one another. This framework encompasses a range of possibilities that include the all-inclusive cartel (i.e., $\mathcal{C} = \{N\}$), that allows for the presence of noncollusive bidders (i.e., some elements of $\mathcal{C}$ may be singleton) and/or for multiple bidding cartels (i.e., $\mathcal{C}$ may include C_j, $j = 1, ..., k$ with $|C_j| \geq 2$). The cartel structure $\mathcal{C}$ is a common knowledge for all bidders in N and for the seller. The assumption that the seller knows the cartel structure, albeit not innocuous, may not be as restrictive as it may appear. For instance, our analysis would still apply if some cartel may not collude effectively. Also, the structure of potential bidding cartels (who is likely to collude with whom) can be sometimes discerned from prior auction experiences and other industry observables. Of course, none of these issues arise if there is only one cartel, as has been assumed in all existing papers. In this sense, the current model generalizes all existing models of collusion.

The time line is similar to that of LM and CK, except for one important difference: Cartels are formed prior to the bidders' participation into the mechanism.

□ Time line:

- At date 0, each bidder learns his type, θ_i, drawn from Θ_i. The realized type is private information of the bidder, unobservable to the seller as well as to other bidders.
- At date 1, the seller proposes an auction rule $M \in \mathcal{M}$.
- At date 2, the (uninformed) representative of each cartel $C \in \mathcal{C}$ simultaneously proposes a collusive side contract (to be described in detail later). Each member of $C \in \mathcal{C}$ then accepts or rejects the contract. If all bidders of C accept, then that cartel's side contract is enforced; or else, the members of C play the subsequent game non-cooperatively. Neither the side contract proposed for a cartel C nor its members' decision on accepting that contract is observed by the bidders outside C (and by the seller).
- At date 3, each bidder, $i \in N$, chooses $\tilde{\theta}_i \in \overline{\Theta}_i$; i.e., he accepts or rejects M, and reports from Θ_i if he accepts. (If the side contract of a cartel was accepted, then its members report according to the side contract.)
- At date 4, if collusion by a cartel is active, then the outcome of their side contract arises. If no collusion is active, then M results.

□ Collusion Technology:

We assume that each cartel has at its disposal four instruments: (a) its members' participation decisions, (b) participating members' communication with the seller (e.g., bids), (c) reallocation of the good within the cartel, in case a member of that cartel receives the good, and (d) side payments that the cartel members can exchange in a budget-balanced fashion. These four instruments together encompass all possible ways in which a cartel can coordinate their members' behavior.

To formally describe possible manipulations utilizing all these instruments, fix a possible cartel $C \in \mathcal{C}$, and an auction rule $M = (q, t) \in \mathcal{M}$ the seller may propose. We then suppose that an uninformed representative of each cartel C, $|C| \geq 2$ proposes a side contract to its members, given that bidders outside C behave non-collusively (or equivalently their representatives offer null contracts). The latter presumption is made since later we shall focus on how non-collusive behavior can be supported as an equilibrium, which requires a *unilateral* deviation by each cartel to be prevented. Instead of considering a possible side contract by each cartel, it is convenient to think of a *manipulation*, the outcome that will emerge when that side contract is enforced *and* all others, including noncollusive bidders and members of different cartels, report truthfully.

Formally, an outcome, $\tilde{M} = (\tilde{q}, \tilde{t})$ is a *manipulation of M by cartel C*, if there exists a function, $\mu^C : \Theta_C \to \Delta\overline{\Theta}_C$, that maps from their types in Θ_C into a probability distribution over $\overline{\Theta}_C$ such that, $\forall \theta \in \Theta$,

$$\sum_{i \in C} \tilde{q}_i(\theta) = \mathbf{E}_{\mu^C(\theta_C)}[\sum_{i \in C} q_i(\tilde{\theta}_C, \theta_{N \backslash C})], \qquad (RC_C^M)$$

$$\tilde{q}_i(\theta) = \mathbf{E}_{\mu^C(\theta_C)}[q_i(\tilde{\theta}_C, \theta_{N \backslash C})], \forall i \in N \backslash C, \qquad (RC_{N \backslash C}^M)$$

$$\mathbb{E}\left[\sum_{i \in C} \tilde{t}_i(\theta)\right] = \mathbb{E}\left[\sum_{i \in C} \mathbf{E}_{\mu^C(\theta_C)}[t_i(\tilde{\theta}_C, \theta_{N \backslash C})]\right], \qquad (BB_C^M)$$

$$\tilde{t}_i(\theta) = \mathbf{E}_{\mu^C(\theta_C)}[t_i(\tilde{\theta}_C, \theta_{N \backslash C})]. \forall i \in N \backslash C. \qquad (BB_{N \backslash C}^M)$$

These conditions are explained as follows. First, condition (RC_C^M) requires the final assignment of the good to be "reallocationally consistent" in the sense that the good is allocated to any cartel member only if some member of that cartel obtains the good from the seller under some manipulation of reports/participation decision. Condition (BB_C^M) allows the cartel members to exchange side transfers in a budget-balanced fashion. Since budget balancing is required at the ex ante level, we are allowing for the cartel to finance (from a competitive capital market) across different realizations of its members' type profiles.[5] Conditions $(RC_{N \backslash C}^M)$ and $(BB_{N \backslash C}^M)$ simply assume that bidders outside C are not colluding: there is no reallocation and no exchange of side payments among all bidders outside C and between C

[5]Our results do not change, if budget balancing is required at the *ex post* level. Clearly, our collusion-proof implementation result would be stronger with the ex ante version of budget balancing, explaining our choice.

and $N\backslash C$. This presumption would be without any loss if $N\backslash C$ were all noncollusive. Even if $N\backslash C$ may involve some bidding cartels, the above conditions are still sufficient for there to be an equilibrium with no collusion, since they ensure that the representative of each cartel offers a null side contract *when other cartels adopt null side contracts.*

□ Incentive Feasibility of Collusion

For collusive manipulation to work, the members of the cartel must have the incentive to carry it out. We say that $\tilde{M}$ is *feasible* if it satisfies

$$U_i^{\tilde{M}}(\theta_i) \geq u_i^{\tilde{M}}(\tilde{\theta}_i, \theta_i), \quad \forall i \in C, \theta_i \in \Theta_i, \tilde{\theta}_i \in \overline{\Theta}_i, \qquad (IC_C^*)$$

and

$$U_i^{\tilde{M}}(\theta_i) \geq U_i^M(\theta_i), \quad \forall i \in C, \theta_i \in \Theta_i. \qquad (IR_C^M)$$

These conditions are explained as follows. First of all, (IC_C^*) requires the outcome resulting from collusion to be incentive compatible to all members of cartel. Since the cartel members face asymmetric information about one another, this condition must hold, regardless of the specifics of how the cartel is formed and how the members bargain over their collusive arrangement. Next, (IR_C^M) requires that each member of the cartel must do as well with the proposed manipulation as they would by vetoing that manipulation and acting noncooperatively. Clearly, what each member will get in the latter event depends on the inferences made by the members of the cartel about that member. Condition (IR_C^M) assumes that no new inferences about the members' types are made in such an event. This "passivity" of out-of-equilibrium beliefs is an important element of LM's weak collusion-proofness notion. Although (IR_C^M) assumes passive out-of-equilibrium beliefs, it in fact accommodates all non-pessimistic beliefs for our purpose. If a collusive proposal is made unattractive to a bidder with a passive belief about what will happen when he refuses the proposal, it will be unattractive to him if his beliefs were more optimistic about that event. In this sense, the real restriction arising from (IR_C^M) is for out-of-equilibrium beliefs to be non-pessimistic. This restriction serves as a reasonable discipline over belief formation.[6]

Lastly, note these conditions are imposed only for the members of the cartel, since the manipulation constitutes its deviation unobserved by outsiders of that cartel. We turn next

[6]In fact, it is not too difficult to construct *a* non-collusive equilibrium, supported by an arbitrarily *optimistic* belief. The seller can simply make available an option which would pay an arbitrarily large amount to a bidder (say paid by a different bidder) if the bidders were to coordinate in the *right* way; the very optimistic belief that such a coordination would occur in the event of rejecting a collusive offer can sustain a non-collusive equilibrium. Clearly, such an equilibrium is not believable, and the "passivity" restriction can be seen to place a discipline against such an equilibrium by limiting the degree of optimism entertainable by the potential colluders when rejecting a collusive offer.

to the weak notion of collusion-proof auctions.

Definition WCP. *An auction rule $M \in \mathcal{M}$ is weak collusion-proof (henceforth, WCP), if, for each cartel $C \in \mathcal{C}$ with $|C| \geq 2$, any feasible manipulation of M by C makes no member of C strictly better off.*

To explain this notion, suppose the seller offers an auction rule M. If M is WCP, then, for each cartel C, there exists no feasible manipulation that would make some members of C strictly better (without making the other members of C worse off), given that all other cartels are inactive. Our WCP notion is the same as the WCP of LM, except that we allow for randomization and reallocation possibilities in the collusive bidders and that we allow for proper subsets of bidders to be collusive. It is in fact a natural generalization of their notion to allow for these new features. It follows from LM that, if an auction rule is WCP, it will admit an equilibrium in which the representatives of cartels propose no collusive manipulations — or equivalently, they all propose null side contracts.[7]

Despite the restrictions, the WCP auctions are worth studying for several reasons. First, WCP auctions offer a reasonable protection against collusion since it is often unrealistic for cartel members to punish more severely than bidding noncooperatively. Second, the weak notion provides a conservative test of when collusion imposes a real cost to the seller: If for instance there is no WCP auction that would allow a seller to earn the second-best revenue, then one can safely conclude that collusion matters, for the seller would not fare any better if the bidders can collude more effectively. Finally, the restrictions involved in the WCP notion are not crucial for the results obtained. We will later show that under some condition, the main result can be strengthened to the strong notion of collusion-proofness.

4 WCP Implementation of the Second-Best Outcome

In this section, we characterize the properties of WCP auction rules (Lemma 1 and 2) and use these properties to obtain a necessary condition for the second-best outcome to be WCP implementable (Theorem 1). We then show that, for symmetric bidders, the necessary condition is also sufficient for the WCP implementability of the second-best outcome.

[7] Our WCP requirement is equivalent to that the auction rule be "interim incentive efficient" according to the terminology of Holmstrom and Myerson (1983). Crawford (1985) shows that this latter condition is sufficient for a mechanism to be "attainable" in the sense of withstanding a collusive proposal when the proposal is negotiated according to some modified Nash demand game. The WCP requirement is weaker than "durability" required by Holmstrom and Myerson (1983), however.

4.1 Properties of WCP Auction Rules

Fix a cartel $C \in \mathcal{C}$, and an auction rule $M = (q,t)$ that the seller proposes. It is useful to have a few definitions. Let $q_i^C(\theta_C) := \mathbb{E}_{\tilde{\theta}_{N\setminus C}}[q_i(\theta_C, \tilde{\theta}_{N\setminus C})]$. Let $q_C(\theta) := \sum_{i\in C} q_i(\theta)$ and $Q_C(\theta_C) := \mathbb{E}_{\tilde{\theta}_{N\setminus C}}[q_C(\theta_C, \tilde{\theta}_{N\setminus C})]$ denote the probability that the auction rule allocates to good to a member of the cartel given the value profile of all bidders and that of the cartel members, respectively. Let $\mathcal{Q}_C := [0, \sup_{\theta_C \in \overline{\Theta}_C} Q_C(\theta_C)]$ be the set of all probabilities with which the cartel can secure the good to its members under M. This set contains zero since all its members can boycott the auction i.e., $Q_C(\theta_\emptyset, \cdots, \theta_\emptyset) = 0$. This set is convex since the cartel members can randomize between boycotting and reporting some profile $\theta_C \in \Theta_C$. We then obtain our first property of WCP auction rules.

LEMMA 1. *If $M = (q,t) \in \mathcal{M}$ is WCP, then for each $C \in \mathcal{C}$ there exists a convex function, $r : \mathcal{Q}_C \to \mathbb{R}_+$ with $r(0) = 0$, such that*

$$\mathbb{E}_{\tilde{\theta}_{N\setminus C}}\left[\sum_{i\in C} t_i(\theta_C, \tilde{\theta}_{N\setminus C})\right] = r(Q_C(\theta_C)), \forall \theta_C \in \overline{\Theta}_C.$$

To see how this property restricts the auction rules, suppose all bidders belong to one cartel, and suppose the seller wishes to implement a deterministic allocation (i.e., $q(\cdot) \in \{0,1\}$). Lemma 1 implies that, for the auction to be weak collusion proof, it must charge a single price if and only if the good is sold. More generally, the seller cannot collect any fee from a cartel whenever its members do not obtain the good. This feature arises from the abilities of the bidders to collude on their participation decisions; were they charged positive entry fees, they could all simply refuse to participate. Similarly, the collusive bidders can never be charged different prices for the same probability of obtaining the good; or else, they could manipulate their reports (or bids) to pick the lowest price for a given probability of obtaining the good. The surplus generated from such manipulation can be shared among all cartel members via appropriate side transfers and reallocations so that (IC^*) and (IR_C^M) conditions are satisfied. Finally, the sale price is (weakly) convex in the probability of the object being allocated to any cartel member, since the cartel members can at least randomize between non-participation and any probability of allocation attainable by some reports.

The next property is obtained for a class of allocation rules satisfying *monotonicity*: for all $C \in \mathcal{C}$, $q_C(\cdot)$ is nondecreasing in θ_C and, for all $C \in \mathcal{C}$ and for all $i \in C$, $q_i(\cdot)$ is nondecreasing in θ_i and nonincreasing in θ_{-i}. Let $\mathcal{M}_0 \subset \mathcal{M}$ denote the set of auction rules satisfying this monotonicity. The monotonic allocation rules are reasonable, and would naturally arise from the standard auctions such as first- and second-price auctions. Next, we define the average price charged to the cartel per unit probability:

$$p(\theta_C) := \begin{cases} \dfrac{r(Q_C(\theta_C))}{Q_C(\theta_C)} & \text{if } Q_C(\theta_C) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

LEMMA 2. *If $M = (q,t) \in \mathcal{M}_0$ is WCP, then $\forall\, C \in \mathcal{C}$, $\forall\, i \in C$ and for almost every $\theta_C \in \Theta_C$,*

$$(K_i(\theta_i) - p(\theta_C))\, q_i^C(\theta_C) \geq \max_{\theta_i' \in \{\theta_\emptyset\} \cup [\underline{\theta}_i, \theta_i]} (K_i(\theta_i) - p(\theta_i', \theta_{C-i}))\, q_i^C(\theta_i', \theta_{C-i}). \tag{2}$$

This lemma characterizes the extent to which each cartel can "behave like a single agent." Specifically, condition (2) resembles an incentive compatibility constraint for a "single" agent who may consume one of $|C|$ alternative values. But this resemblance is not perfect. First, that agent realizes "pseudo" value $K_i(\theta_i)$ rather than true value θ_i. Second, the bidder's constraint is required only in one direction, i.e., not to under-report or withdraw from the auction. Third, the agent may not be able to shift his consumption among the alternative uses. All together, these features serve to limit the extent to which the cartel can coordinate their members' behavior. For instance, the fact that pseudo values, rather than true values, matter means that the cartel can be forced to sustain some ex post loss. Since $K_i(\theta_i) > \theta_i$, an average price of $p(\theta_i) > \theta_i$ need not violate the above constraint. The cartel's limited ability to coordinate their behavior arises from the fact that any collusive defection requires a consensus from all types of bidders. Different types of bidders may have conflicting interests, say about consumption of the good by any particular type θ_i. For instance, if $\theta_i < p(\theta_C)$, then the cartel wishes to cancel such consumption, but the highest type of bidder i would not agree as long as $K_i(\theta_i) > p(\theta_C)$.

Naturally, the necessary conditions for WCP implementation (given by Lemma 1 and 2) constrain the set of circumstances in which the second-best outcome is WCP implementable. We next characterize these circumstances. To this end, fix any bidder $i \in C$ for some $C \in \mathcal{C}$ with $|C| \geq 2$. For each profile $\theta_{N \setminus C} \in \Theta_{N \setminus C}$, let

$$\phi_i(\theta_{N\setminus C}) := \inf\{\theta_i \in \Theta_i \mid J_i(\theta_i) \geq \max\{\max_{j \in N \setminus C} J_j(\theta_j), 0\}\}$$

denote the lowest type of bidder i that can obtain the good with positive probability in the second-best allocation.

CONDITION (SB): (i) If $\mathcal{C} = \{N\}$, then

$$K_i(\hat{\theta}_i)\left(\mathbb{E}\left[\sum_{i \in N} q_i^*(\theta)\right]\right) \geq \mathbb{E}\left[\sum_{i \in N} J_i(\theta_i) q_i^*(\theta)\right], \forall i \in N.$$

(ii) If $\mathcal{C} \neq \{N\}$, then, for each $C \in \mathcal{C}$ with $|C| \geq 2$,

$$\mathbb{E}\left[\sum_{i\in C} K_i(\phi_i(\theta_{N\setminus C}))q_i^*(\theta)\right] \geq \mathbb{E}\left[\sum_{i\in C} J_i(\theta_i)q_i^*(\theta)\right].$$

This condition is explained as follows. The RHS of the inequalities represent the amounts of surplus that *should* be extracted from the cartel to implement the second-best payoff for the seller. As will be proven next, the LHS of the inequalities represent the highest payments that are collusion-proof collectable from the cartel, given the second-best allocation q^*. Thus, the inequalities are necessary for the second-best outcome to be WCP implementable.

THEOREM 1. (NECESSITY) CONDITION (SB) *is necessary for the second-best outcome to be WCP implementable.*

This theorem immediately identifies a class of situations for which the second-best outcome is not WCP implementable.

COROLLARY 1. *A second-best outcome is not WCP implementable if it assigns the good to members of a cartel with probability one.*

PROOF: Without loss $C = N$, or else we can simply redefine N to coincide with C. We show that Condition (SB)-(i) fails if the members of N receives the good with probability one. Since $\sum_{i\in N} q_i^*(\theta) = 1$ for all θ, there must exist i such that $J_k(\underline{\theta}_i) \geq 0$, so $\hat{\theta}_k = \underline{\theta}_k$. Hence,

$$K_k(\hat{\theta}_k) = \underline{\theta}_k = \mathbb{E}[J_k(\theta_k)] < \mathbb{E}[\sum_{i\in N} J_i(\theta_j)q_i^*(\theta)],$$

where the strict inequality follows from the assumption (made in Section 2) that $\mathbb{E}[q_k^*(\theta)] < 1$. ∎

This shows that excluding some types of collusive bidders is crucial for WCP implementation of the second-best outcome, a theme that will be generalized later.

4.2 WCP Implementation of the Second-Best Outcome: Symmetric Bidders

Here we show that Condition (SB) is also sufficient for the second-best outcome to be WCP implementable when bidders are symmetric. Specifically, we construct an auction rule that will WCP implement the second-best outcome, given (SB). Further, this sufficient condition will be seen to hold if either at least one bidder is noncollusive or the second-best mechanism involves a nontrivial probability of no sale.

We begin with the symmetry assumption: $F_i(\cdot) =: F(\cdot)$ for all $i \in N$, for some common cdf $F(\cdot)$ which has a positive density f. The associated virtual valuations J and K are defined analogously, and their monotonicity properties are maintained. Likewise, we let $\hat{\theta} := \inf\{\theta | J(\theta) \geq 0\}$. Condition (SB) is now more succinctly described in this environment. Define first $\theta_C^{(1)} := \max_{i \in C} \theta_i$ and $\theta_{N \backslash C}^{(1)} := \max\{\max_{i \in N \backslash C} \theta_i, \hat{\theta}\}$. (We adopt a convention that $\theta_{N \backslash C}^{(1)} := \hat{\theta}$ when $\mathcal{C} = \{N\}$.) Then, Condition (SB) simplifies to:

Condition (SB′): For each C with $|C| \geq 2$, $\mathbb{E}\left[K(\theta_{N \backslash C}^{(1)}) | \theta_C^{(1)} > \theta_{N \backslash C}^{(1)}\right] \geq \mathbb{E}\left[J(\theta_C^{(1)}) | \theta_C^{(1)} > \theta_{N \backslash C}^{(1)}\right]$.

As pointed out earlier, this condition requires a collusive bidder's valuation to be sufficiently high whenever the good is allocated to him. It turns out that this requirement is not very onerous to satisfy. The condition holds if at least one buyer is noncollusive or there are more than one bidding cartel, or if the cutoff threshold $\hat{\theta}$ is sufficiently high.

Lemma 3. Condition (SB′) *holds if* $\mathcal{C} \neq \{N\}$, *or if* $\mathcal{C} = \{N\}$ *and*

$$K(\hat{\theta}) \geq \mathbb{E}\left[J(\theta_N^{(1)}) | \theta_N^{(1)} > \hat{\theta}\right]. \tag{3}$$

In case of the all-inclusive cartel (i.e., $\mathcal{C} = \{N\}$), Condition (SB′), or equivalently (3), is not trivial. For instance, if $\hat{\theta} = \underline{\theta}$, then $K(\hat{\theta}) = \underline{\theta}$, so the condition fails. In other words, a non-trivial exclusion is necessary to prevent collusion. Although not obvious at first glance, a similar exclusion is used in the case $\mathcal{C} \neq \{N\}$ in that the good is sold to bidders outside that cartel with positive probability.

We now construct an auction rule that WCP implements the second-best outcome when Condition (SB′) holds. Suppose that a second price auction is held with a reserve price $\hat{\theta}$, and consider the associated outcome $M^* = (q^*, t^*)$ defined over $\overline{\Theta}$. (Recall that the outcome is well defined even when some bidders do not participate, a situation described by $\theta_\emptyset$ being chosen by these bidders.) We then construct a new auction rule $\hat{M} = (\hat{q}, \hat{t})$ defined over $\overline{\Theta}$. The allocation rule $\hat{q}$ is constructed so that $\hat{q}(\cdot) = q^*(\cdot)$. To construct the payment rule $\hat{t}$, we first determine the sale price against each cartel $C \in \mathcal{C}$ (with $|C| \geq 2$). Let $\alpha_C \in [0, 1]$ satisfy

$$\mathbb{E}\left[\alpha_C K(\theta_{N \backslash C}^{(1)}) + (1 - \alpha_C) J(\theta_{N \backslash C}^{(1)}) | \theta_C^{(1)} > \theta_{N \backslash C}^{(1)}\right] = \mathbb{E}\left[J(\theta_C^{(1)}) | \theta_C^{(1)} > \theta_{N \backslash C}^{(1)}\right]. \tag{4}$$

Condition (SB′) allows such an α_C to be well defined. The sale price against cartel C is then set at $H_C(\theta_{N \backslash C}^{(1)}) := \alpha_C K(\theta_{N \backslash C}^{(1)}) + (1 - \alpha_C) J(\theta_{N \backslash C}^{(1)})$. This sale price is defined in terms of the highest type of bidder outside C and is set above $J(\theta_{N \backslash C}^{(1)})$ just enough to extract $J(\theta_C^{(1)})$ on average from the highest valuation bidder in C. Let $\delta_C(\theta) := H_C(\theta^{(1)}_{N\backslash C})\sum_{i\in C} q_i^*(\theta)$ denote the expected sale price charged against cartel C.

We now describe the payment rule $\hat{t}$. For each noncollusive bidder i (i.e., $\{i\} \in \mathcal{C}$), we set $\hat{t}_i(\theta) := t_i^*(\theta)$, $\forall \theta \in \overline{\Theta}$. For each cartel $C \in \mathcal{C}$, let $C(\theta_C) := \{i \in C | \theta_i \neq \theta_\emptyset\}$ be the set of its members who participate in the auction, given θ_C. For each $i \in C$, if $C(\theta_C) = C$, then we set

$$\begin{aligned}\hat{t}_i(\theta) \quad &:= \quad \frac{1}{|C|}\delta_C(\theta) + \mathbb{E}_{\tilde{\theta}_{-i}}\left[t_i^*(\theta_i, \tilde{\theta}_{-i}) - \frac{1}{|C|}\delta_C(\theta_i, \tilde{\theta}_{-i}))\right] \\ &\quad - \frac{1}{|C|-1}\sum_{k\in C\backslash\{i\}} \mathbb{E}_{\tilde{\theta}_{-k}}\left[t_k^*(\theta_k, \tilde{\theta}_{-k}) - \frac{1}{|C|}\delta_C(\theta_k, \tilde{\theta}_{-k})\right]\end{aligned} \tag{5}$$

and, if $C(\theta_C) \subsetneq C$, then we set

$$\hat{t}_i(\theta) \quad := \begin{cases}\delta_C(\overline{\theta}) & \text{if } i \in C(\theta_C) \\ 0 & \text{if } i \in C\backslash C(\theta_C).\end{cases} \tag{6}$$

Two properties of the current construction are important. First, in case all bidders participate, (5) implies $\mathbb{E}_{\tilde{\theta}_{-i}}[\hat{t}_i(\theta_i, \tilde{\theta}_{-i})] = \mathbb{E}_{\tilde{\theta}_{-i}}[t_i^*(\theta_i, \tilde{\theta}_{-i})]$, $\forall i \in N$, $\forall \theta_i \in \Theta_i$,[8] so that each bidder has the same interim incentives as with M. This property means that the new auction rule $\hat{M}$ inherits the incentive and participation properties of the original auction M^*. Hence, $\hat{M}$ satisfies (IC^*) and implements V^*. Second, it can be checked from (5) and (6) that $\sum_{i\in C}\hat{t}_i(\theta) = \delta_C(\theta) = H_C(\theta^{(1)}_{N\backslash C})\sum_{i\in C} q_i^*(\theta)$ if every member of each cartel participates (or else, each participating cartel member pays a high "punishment price" equal to $\delta_C(\overline{\theta})$). That is, upon participating, each cartel $C \in \mathcal{C}$ is charged a single sale price, $H(\theta^{(1)}_{N\backslash C})$, depending only on non-cartel members' types and payable only when the object is allocated to its member. This property satisfies Lemma 1 and also ensures that the cartel cannot manipulate the sale price charged against its members. These two properties deliver weak collusion-proof implementation of the second-best outcome:

[8] To see this, note first, by (5) and symmetry,

$$\mathbb{E}[\delta_C(\theta)] = \mathbb{E}\left[J(\theta^{(1)}_C)1_{\{\theta^{(1)}_C > \theta^{(1)}_{N\backslash C}\}}\right] = \mathbb{E}\left[\sum_{i\in C} t_i^*(\theta)\right] = |C|\mathbb{E}\left[t_i^*(\theta)\right], \forall i \in C, \tag{7}$$

from which it follows that

$$\mathbb{E}_{\tilde{\theta}_{-i}}[\hat{t}_i(\theta_i, \tilde{\theta}_{-i})] = \mathbb{E}_{\tilde{\theta}_{-i}}[t_i^*(\theta_i, \tilde{\theta}_{-i})] - \frac{1}{|C|-1}\sum_{k\in C\backslash\{i\}} \mathbb{E}_{\tilde{\theta}}\left[t_k^*(\tilde{\theta}) - \frac{1}{|C|}\delta_C(\tilde{\theta})\right] = \mathbb{E}_{\tilde{\theta}_{-i}}[t_i^*(\theta_i, \tilde{\theta}_{-i})],$$

where the last equality follows from (7).

Theorem 2. (Sufficiency) *Given* Condition (SB′), *the auction rule* $\hat{M}$ *is WCP and achieves the second-best revenue.*

Combining Lemma 3 with Theorem 2 produces one of the main results of this paper.

Corollary 2. *The second-best outcome is WCP implementable if* $\mathcal{C} \neq \{N\}$, *or if* $\mathcal{C} = \{N\}$ *but (3) holds.*

In words, the second-best outcome is weak collusion-proof implementable if a cartel is not all-inclusive, which will be the case if either there exists at least one noncollusive bidder or there are multiple bidding cartels. In these latter cases, the seller can leverage the presence of the bidders outside a cartel to extract sufficient rents from the cartel. If entire bidders form a bidding cartel, then no such leverage exists, but the seller can still use the threat of no sale to accomplish the same objective as long as that threat is sufficiently credible in the sense of (3).

The following examples illustrate two different scenarios.

Example 1. ($\mathcal{C} = \{\{1, 2\}\}$) *Suppose there are two bidders each with valuation drawn uniformly from* $[0, 1]$. *According to Theorem 0, the second-best outcome allocates the object efficiently for valuation exceeding* $\hat{\theta} = \frac{1}{2}$, *and yields revenue of* $\frac{5}{12}$. *This also satisfies (3), so the second-best is WCP implementable. The WCP auction rule charges a sale price of* $r^* := \mathbb{E}[J(\theta_N^{(1)})|\theta_N^{(1)} > \hat{\theta}] = 5/9$ *to the bidders, regardless of who win and what their bids are. Without collusion, each bidder receives the interim payoff of*

$$U^{\hat{M}}(\theta) = \begin{cases} 0 & \text{if } \theta \in [0, \frac{1}{2}) \\ \frac{1}{2}\theta^2 - \frac{1}{8} & \text{if } \theta \in [\frac{1}{2}, 1]. \end{cases}$$

Since the cartel is charged a sale price of $5/9$, *it suffers an ex post loss whenever the highest valuation is in the interval* $[\frac{1}{2}, \frac{5}{9}]$. *Why can they not simply boycott the auction in this situation? Indeed, their joint surplus will increase by doing so. The problem, however, is that the increased surplus cannot be allocated to benefit all types; some types will be strictly worse off and thus object to that move. To illustrate, suppose indeed that the bidders boycott auction whenever no bidder has valuation exceeding* $5/9$, *and, otherwise, the high-valuation bidder consumes the object. Under this collusive arrangement, labeled* $\tilde{M}$, *each bidder's interim payoff is*[9]

$$U^{\tilde{M}}(\theta) = \begin{cases} \frac{32}{2187} & \text{if } \theta \in [0, \frac{5}{9}) \\ \frac{1}{2}\theta^2 - \frac{611}{4374} & \text{if } \theta \in [\frac{5}{9}, 1]. \end{cases}$$

[9] This payoff can be obtained by applying the transfer rule in (29) with $n = 2$, $r = 5/9$, $\rho_i = 0$, and $q_i(\theta_i, \theta_{-i})$ being equal to 1 if $\theta_i > \max\{\theta_{-i}, 5/9\}$ and 0 otherwise.

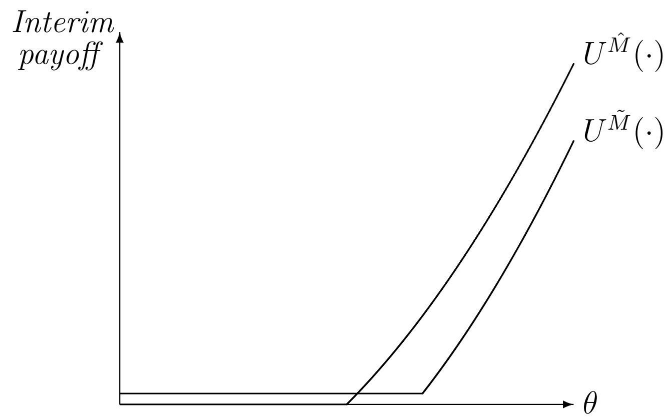

Figure 1: Comparison of interim payoffs from $\hat{M}$ and $\tilde{M}$

As can be seen from Figure 1, a bidder benefits from this collusion when his valuation is less than 0.528 *but is strictly worse off if his valuation is higher. Hence, a collusive arrangement $\tilde{M}$ is not feasible. (The same is true for any other feasible manipulations.) Even though the net expected surplus may rise with some collusive manipulation, incentive compatibility facing the collusive bidders limits the way surplus can be allocated across types to make them uniformly better off. In this sense, our WCP auction exploits the informational asymmetry facing the collusive bidders.*

Example 2. ($\mathcal{C} = \{\{1,2\},\{3\}\}$) *Consider the example from Section 2, where there are two collusive bidders and a noncollusive bidder, each with valuation drawn from* $[1,2]$. *Here, the presence of a noncollusive bidder is crucial for WCP implementability of the second-best outcome. (If all three bidders belong to one grand cartel, the second-best is not WCP implementable, for the distribution fails (3).) Our WCP auction charges the sale price of* $H_{\{1,2\}}(\theta_3) = 2\theta_3 - \frac{5}{4}$ *to the cartel* $\{1,2\}$*, In fact, the non-collusive can be induced to reveal its type as a dominant strategy. The cartel cannot manipulate the sale price, for it is tied to the non-collusive bidder's strategy. Further, their ability to collude on participation does not help. Since the allocation is efficient, when a collusive bidder, say with valuation θ wins, the cartel ends up paying $\theta - \frac{1}{4}$ in expected value, which leaves the expected surplus of $\frac{1}{4}$ to the cartel.*[10] *Hence, the cartel has no incentive to boycott the auction regardless of θ. Other*

[10]This can be seen by the fact that

$$\mathbb{E}[t_1(\tilde{\theta}) + t_2(\tilde{\theta}) | \tilde{\theta}_3 < \max\{\tilde{\theta}_1, \tilde{\theta}_2\} = \theta] = \mathbb{E}[H_{\{1,2\}}(\tilde{\theta}_3) | \tilde{\theta}_3 < \theta] = \int_1^\theta \left(\frac{2\tilde{\theta}_3 - \frac{5}{4}}{\theta - 1} \right) d\tilde{\theta}_3 = \theta - \frac{1}{4}.$$

possible incentive for misrepresenting the cartel's types is also thwarted by the lack of consensus in the interests among the cartel members in a way much like Example 1.

4.3 WCP Implementation of the Second-Best Outcome: Asymmetric Bidders

We now turn to the case of asymmetric bidders. In this case, the optimal noncollusive auction, as characterized in Theorem 0, requires bidders to be treated differently based on their ex ante distribution of types. This presents an extra challenge for the WCP implementation since, as shown in Lemma 1, the same price is charged no matter which member of the cartel receives the good. This does not mean, however, that the collusive bidders cannot be treated differently, for different interests of the heterogeneous types can be exploited to make (IR_N^M) difficult to satisfy. Indeed, we will show that the second-best outcome is WCP implementable at least with respect to the all-inclusive cartel (i.e., $\mathcal{C} = \{N\}$), under a condition that is not much stronger than Condition (SB).

Consider now a strict inequality version of (SB)-(i):

CONDITION (SB*): $K_i(\hat{\theta}_i) > r^* := \frac{\mathbb{E}\left[\sum_{i \in N} J_i(\theta_i) q_i^*(\theta)\right]}{\mathbb{E}\left[\sum_{i \in N} q_i^*(\theta)\right]}, \forall i \in N.$

THEOREM 3. *Assume* $\mathcal{C} = \{N\}$ *and* CONDITION (SB*) *holds. Suppose also that* $(J_i(\cdot) - r^*) f_i(\cdot)$ *is increasing in the interval* $[\underline{\theta}_i, J_i^{-1}(r^*)]$ *for all* $i \in N$*. Then, there exists an auction rule which is WCP and achieves the second-best outcome.*

5 Optimal WCP Auctions

What happens if the second-best outcome is not collusion-proof implementable? What will the optimal WCP auction look like in such a case? While a complete answer to this latter question is unavailable, we provide two observations that will help answer that question. First, we show that any optimal WCP auction in a general monotonic class involves "exclusion" — a positive probability that the object is not sold to any collusive bidder. Second, we characterize the optimal WCP auction completely in the linear class for the case of symmetric bidders.

5.1 An Exclusion Principle

Corollary 1 shows exclusion of collusive bidders is necessary for a second-best outcome to be WCP implementable. We generalize this result to an optimal WCP auction in the monotonic class. That is, we show below that the optimal auction rule in the monotonic class must entail some exclusion of collusive bidders. To gain some idea behind this result, suppose to the contrary that optimal WCP auction sells the object to some members of any given cartel C. Lemma 1 implies that the seller can only charge a sale price to collusive bidders, regardless of their types. Meanwhile, Lemma 2 says that this price cannot be too high relative to the pseudo value, $K_i(\theta_i)$, of the collusive bidder who consumes the good. Combined together, these lemmas imply that the seller must either charge a low sale price, or else she should exclude types with low pseudo values from consuming the good.

THEOREM 4 (Exclusion Principle). *Assume that there are more than one bidder $i \in C$ with $\underline{\theta}_i = \underline{\theta} := \max_{j\in C} \underline{\theta}_j$. Then, the optimal WCP mechanism in $\mathcal{M}_0$ requires that the object not be sold to any member of C with a positive probability.*

PROOF: Let $M = (q, t)$ denote the optimal WCP mechanism. Suppose to the contrary that $\sum_{i\in C} q_i(\theta) = 1$ for all $\theta \in \Theta$, which implies by Lemma 1 that $\sum_{i\in C} t_i(\theta) = r$ for some r. Then, Lemma 2 requires that

$$r \leq \max_{i\in C} K_i(\underline{\theta}_i) = \underline{\theta}.$$

Thus, the revenue cannot exceed $\underline{\theta}$. We now generate a contradiction by constructing a WCP mechanism which raises a higher revenue than $\underline{\theta}$: Sell the object at a fixed price $\tilde{r}$, which is slightly greater than $\underline{\theta}$, if and only if at least one member of C has a value higher than $\tilde{r}$. This take-it-or-leave offer is clearly WCP and generates a revenue equal to $R(\tilde{r}) := \tilde{r}(1 - \prod_{i\in C} F_i(\tilde{r}))$. And $R(\tilde{r}) > R(\underline{\theta}) = \underline{\theta}$ for $\tilde{r}$ slightly above $\underline{\theta}$ since

$$\frac{d}{d\tilde{r}}R(\tilde{r})\Big|_{\tilde{r}=\underline{\theta}} = (1 - \prod_{i\in C} F_i(\underline{\theta})) - \underline{\theta}\sum_{i\in C} f_i(\underline{\theta})\prod_{j\neq i} F_j(\underline{\theta}) = 1 > 0,$$

where the last equality holds because for each $i \in C$, there exists at least one bidder $j \neq i$ for whom $F_j(\underline{\theta}) = 0$. ▌

An optimal auction excludes some low valuation bidders even in the absence of collusion. Yet, exclusion never arises if there is some buyer i whose valuation is always high so that $J_i(\underline{\theta}) \geq 0$. Collusion tilts the tradeoff toward more exclusion, since the seller can only charge a single sale price, whereas absent collusion bidding competition generates higher payment from high valuation types beyond the reserve price. Consequently, an optimal collusion-proof auction *always* excludes some types of collusive bidders.

5.2 Optimal Linear WCP Auctions for Symmetric Bidders

Here, we search for an optimal WCP auction rule when Condition (SB) fails with symmetric bidders. This condition can only fail when all bidders are collusive (recall Lemma 3 or equivalently Corollary 1), so we focus on that case and assume $\mathcal{C} = \{N\}$. Further, we restrict our search to the class of *linear auction rules* where the aggregate payment schedule is linear in the probability of sale (to any bidder): $\sum_{i \in N} t_i(\theta) = r \cdot Q(\theta)$, for some nonnegative constant r.[11]

With a linear auction rule, the seller offers a uniform price against the cartel. This restriction appears to entail very little loss. Recall from Lemma 1 that the seller can only charge a single price against the cartel for each probability of sale, and that the sale price must be convex in the probability of sale. The linearity restriction simply eliminates the strictly convex portion of the sale price. The convex portion would matter only if the cartel is assigned the object with probability between zero and one at *price discount*, but this latter feature seems unlikely to be appealing to the seller (although we have not ruled out this possibility). In fact, most of the plausible auction rules allocate the object deterministically as functions of bidders' types. Any such auction rules can be implemented by linear auction rules.[12]

Given the linearity restriction, the constraint in Lemma 2 implies

$$(K(\theta_i) - r)q_i(\theta) \geq 0, \forall i \in N, \forall \theta \in \Theta. \tag{8}$$

Consider the following revenue maximization program:

$$[C] \qquad \max_{(r,q)} \mathbb{E}\left[r \sum_{i \in N} q_i(\theta)\right]$$

subject to

$$\sum_{i \in N} U_i(\underline{\theta}_i) = \mathbb{E}\left[\sum_{i \in N}(J(\theta_i) - r)q_i(\theta)\right] \geq 0 \tag{IC_1^*}$$

and

$$\mathbb{E}\left[\sum_{i \in N} \min\{K(\theta_i) - r, 0\}q_i(\theta)\right] \geq 0. \tag{K}$$

The equality in (IC_1^*) follows from Lemma 0, which utilizes bidders' incentive constraints. Hence, (IC_1^*) follows from (IC^*), which is necessary for any feasible auction rule. In fact,

[11] Alternatively, we can restrict attention to auction rules that allocate the object efficiently among the collusive bidders. The analysis based on this restriction is available from the authors.

[12] As will become clear from the subsequent analysis, a solution to $[C]$ below can be obtained even with an additional constraint $\sum q_i(\theta) \in \{0, 1\}$.

without (K), the above maximization problem simply yields the second-best outcome. The constraint (K) follows from the coalitional incentive constraint in (8) (which in turn follows from Lemma 2). Since the program $[C]$ only imposes necessary conditions for weak collusion-proofness, its solution gives an upper bound for the revenue attainable by any linear WCP auction. We show next that this upper bound is attainable by a WCP auction.

THEOREM 5. *Assume $\mathcal{C} = \{N\}$ and bidders are symmetric. Then, there exists an optimal linear WCP auction rule, which implements the solution of $[C]$. If Condition (SB) fails, then the optimal linear auction rule involves a sales price, r_0, that solves either*

$$\max_{r \in R_+} r(1 - F^n(K^{-1}(r))) \tag{9}$$

or

$$\mathbb{E}[J(\theta_N^{(1)})|\theta_N^{(1)} > K^{-1}(r)] = r, \tag{10}$$

and an allocation rule $\hat{q}(\cdot)$ that satisfies

$$\hat{q}_i(\theta) = \begin{cases} 1 & \text{if } \theta_i > \max\{\max_{j \in N\setminus\{i\}} \theta_j, \hat{\theta}_0\} \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where $\hat{\theta}_0 := K^{-1}(r_0)$.

The features of optimal WCP auctions are illustrated by the next example.

EXAMPLE 3. *Assume that there are two bidders whose types are uniformly distributed on interval $[m, m+1]$. Then, $J(\theta) = 2\theta - (m+1)$, $K(\theta) = 2\theta - m$, and $\hat{\theta} = \max\{\frac{m+1}{2}, m\}$. The exclusion threshold is then given by*

$$\hat{\theta}_0 = \begin{cases} \frac{m+1}{2} & \text{if } m \le \frac{7-\sqrt{33}}{2} \\ m + \frac{\sqrt{33}-5}{4} & \text{if } \frac{7-\sqrt{33}}{2} < m \le \hat{m}, \\ \frac{5m+\sqrt{m^2+12}}{6} & \text{if } m > \hat{m}, \end{cases}$$

for some $\hat{m} > \frac{7-\sqrt{33}}{2}$.[13] *Observe that the optimal WCP auction rule implements the second-best if and only if $m \le \frac{7-\sqrt{33}}{2}$. For $m > \frac{7-\sqrt{33}}{2}$, $\hat{\theta}_0 > \hat{\theta}$, so it involves more exclusion than the second-best outcome. Regardless of m, $\hat{\theta}_0 > m$, so the optimal WCP auction always involves exclusion. This is in sharp contrasts to the second-best outcome which does not involve any exclusion if $m \ge 1$.*

[13] More precisely, $\hat{m}$ is the level of m such that (9) and (10) are satisfied simultaneously, which is given as a root of the following equation:

$$(5m + \sqrt{m^2+12})(2m^2 - 9m - 48) + 6(-2m^3 + 9m^2 + 46m + 18) = 0.$$

Remark 1. *Pavlov (2006) finds the outcome presented in Theorem 5 to be optimal in a more restricted class of WCP auction rules, namely, those that are linear* and *symmetric. (Recall the symmetry restriction is not invoked in the current paper.) In fact, he shows that there are asymmetric or nonlinear auction rules that are WCP and yield higher revenue than the optimal auction in the restricted class. However, such auction rules violate our Lemma 1 and thus are not WCP in our model. This difference arises from the fact that we allow for a reallocational ability by the cartel, which is not allowed in Pavlov (2006), at least for the main analysis.*

6 Strong Collusion-Proof Implementation

Thus far, we have focused on weak collusion-proof implementation. Weak collusion-proof auctions protect a seller from a wide range of manipulations collusive bidders may employ. At the same time, it involves some restrictions. First, it rules out collusion supported by non-pessimistic beliefs on the part of members of a cartel about what may happen when they refuse to collude. Although such pessimistic beliefs may not be very plausible, the restriction on beliefs is nonetheless unsatisfactory. Second, the concept of weak collusion-proofness presumes that a cartel is organized by a third party who is uninformed of the members' types. Clearly, such an assumption simplifies the modeling of the collusive behavior, but has no clear empirical justification. We show, however, that these restrictions are not crucial for the main result of the paper: Given an additional condition, the second-best outcome can be implemented in a way robust to the specific beliefs entertained off the equilibrium path and to the particular way in which a cartel is formed and its proposal is made.

To begin, we define a strong collusion-proof auction. To that end, fix any arbitrary indirect auction rule, $A = (B, \xi, \tau)$, where $B = (B_1, ..., B_n)$ is the profile of the message spaces, B_i for bidder i, and $(\xi, \tau) : B \to \mathcal{Q} \times \mathbb{R}^n$ is the outcome function mapping from messages to an allocation and payments. As before, we assume that B_i includes the option of non-participation by i and the outcome function maps a null outcome for the bidder invoking that option. Next, consider an extensive form game in which, for a cartel $C \in \mathcal{C}$, each member of C (which may include a third party maximizing joint payoffs of C) may propose a side contract which maps from $\overline{\Theta}_C$ to a probability distribution over $\prod_{i \in C} B_i$ and which is reallocationally consistent (in the sense the side contract allocates the good to a member of a cartel only when some member of the cartel obtains the good from the seller) and budget balanced, among those members of C who do not invoke $\theta_\emptyset$. Notice the messages used in the side contract contain the nonparticipation option $\theta_\emptyset$ by each member, meaning that the terms of side contract can vary depending on who refuse and who accept the contract. In other words, like Dequiedt (2007), we allow the remaining members of a cartel to commit themselves to punish a member when the latter refuses to participate in collusion. If a side contract has been proposed, the members of the cartel then simultaneously decide whether to reject all side contracts or accept one of them, and the messages in B are chosen according to the agreed-upon side contract,[14] the seller determines the outcome based on the message according to the auction rule, and the good is reallocated and side transfers are exchanged according to the agreed-upon side contracts. Let $\mathcal{E}_A$ be the set of all Bayesian Nash equilibria in undominated strategies. The strong collusion-proof auctions are then defined as follows.

Definition SCP. *Expected revenue V is strong collusion-proof (henceforth, SCP) implementable if there exists an indirect auction rule, $A = (B, \xi, \tau)$, such that $\mathcal{E}_A$ is nonempty and that the seller receives expected revenue of V in every equilibrium of $\mathcal{E}_A$.*

It is worth emphasizing that SCP implementation restricts neither cartel members' out-of-equilibrium beliefs nor, as noted above, their ability to punish a defector. In fact, an SCP auction implementing V would guarantee the seller the revenue in every Bayesian Nash equilibrium, let alone every Perfect Bayesian equilibrium. More important, an SCP auction (or precisely an auction SCP implementing a target level of revenue) is robust to who proposes the collusive proposal. In these respects, the current SCP notion is stronger than any existing notions proposed by the existing authors.

Consider the case of ex ante symmetric bidders. Assume further that there exists only one cartel $C \in \mathcal{C}$ with $|C| > 1$. In this case, we show that the second-best revenue V^* is SCP implementable, given (SB$'$) and an additional condition. We construct the auction rule that SCP implements the desired outcome. That auction rule, labeled $A = (B, \xi, \tau)$, builds on the WCP auction, $\hat{M}$. Recall that $\hat{M}$ provides a dominant strategy incentive for each noncollusive bidder to participate and report truthfully, a feature we retain in A.

We augment the WCP auction $\hat{M}$ by adding a message, r_z, for each cartel member, which is interpreted as a statement: "I reject $\hat{M}$ but would like to buy the item with an eagerness of z," where z is a positive integer of his choosing. Let $B_i := \overline{\Theta}_i \cup \{r_z, z = 1, 2, \cdots\}$ if $i \in C$ and $B_i := \overline{\Theta}_i$ otherwise. We then define the outcome function (ξ, τ) such that it coincides with $\hat{M}$ if no bidder in C announces r_z, but if some bidder in C announces r_z, then the outcome function is constructed as follows. Consider the following condition:

Condition (R): For each $C \in \mathcal{C}$, there exists some $\theta' \in [\hat{\theta}, \bar{\theta}]$ satisfying

$$U^{\hat{M}}(\bar{\theta}) \leq F^{n-|C|}(\theta') \left(\bar{\theta} - \mathbb{E}\left[H_C(\theta^{(1)}_{N \setminus C}) \middle| \theta^{(1)}_{N \setminus C} \leq \theta' \right] \right). \tag{12}$$

[14] That is, a bidder's subsequent bidding is not bound by any side contract that he refused.

Given Condition (R), the outcome function (ξ, τ) of A depends on whether (a) the inequality in (12) holds strictly for all $\theta' \in [\hat{\theta}, \bar{\theta}]$ or (b) there exists some $\theta^r \in [\hat{\theta}, \bar{\theta}]$ satisfying

$$U^{\hat{M}}(\bar{\theta}) = F^{n-|C|}(\theta^r) \left(\bar{\theta} - \mathbb{E}\left[H_C(\theta^{(1)}_{N\setminus C}) \middle| \theta^{(1)}_{N\setminus C} \leq \theta^r \right] \right). \tag{13}$$

In both cases, if there is a collusive bidder who announces r_z, then we pick the bidder who announces r_z with the highest z (with a tie broken randomly among bidders with the same z). In case (a), the bidder receives the object with probability 1 at a fixed price $T^*_i(\bar{\theta})$. In case (b), the bidder receives the object at price $H_C(\theta^{(1)}_{N\setminus C})$ if $\theta^{(1)}_{N\setminus C} \leq \theta^r$, or else the bidder is not awarded the object and he pays nothing. Last, no other bidders receive the object or pay any amount to the seller.

In words, the auction rule A gives an option for a collusive bidder with the highest valuation to secure her non-collusive payoff by announcing r_z with the highest integer z. This extra option serves to limit the cartel's ability to punish a defector, which in turn constrains the set of side contracts sustainable in equilibrium. At the same time, the option itself may become an instrument of collusion if it proves too profitable. Condition (R) ensures that this is not the case.

Theorem 6. *Suppose that there is a single cartel $C \subset N$. Then, given Conditions (SB′) and (R), the second-best revenue V^* is SCP implementable.*

How restrictive is Condition (R)? If the cartel is all-inclusive (i.e., $C = N$), then Condition (R) involves no restriction, for it is always satisfied. To see this, observe the condition in (12) with $\theta' = \bar{\theta}$ becomes

$$T^*_i(\bar{\theta}) \geq H_N(\hat{\theta}).$$

This inequality holds since

$$T^*_i(\bar{\theta}) > \mathbb{E}\left[\frac{T^*_i(\theta_i)}{Q^*_i(\theta_i)} \frac{Q^*_i(\theta_i)}{\mathbb{E}[Q^*_i(\theta_i)]} \right] = \frac{\mathbb{E}[T^*_i(\theta_i)]}{\mathbb{E}[Q^*_i(\theta_i)]} = \frac{\mathbb{E}[J(\theta_i)Q^*_i(\theta_i)]}{\mathbb{E}[Q^*_i(\theta_i)]} = \frac{\mathbb{E}\left[\sum_{i\in N} J(\theta_i)q^*_i(\theta)\right]}{\mathbb{E}\left[\sum_{i\in N} q^*_i(\theta)\right]} = H_N(\hat{\theta}),$$

where the inequality holds since $\frac{T^*_i(\cdot)}{Q^*_i(\cdot)}$ is increasing (since $\frac{T^*_i(\cdot)}{Q^*_i(\cdot)}$ corresponds to the equilibrium bidding function in the symmetric first-price auction), and the last equality follows from (4).

If $C \neq N$, then Condition (R) need not always hold. Nevertheless, the condition is satisfied for a reasonable class of distributions, including the uniform distribution.[15]

[15]For instance, in the Example 2, (12) holds since

$$\max_{\theta\in[1,2]} F^{n-|C|}(\theta) \left(\bar{\theta} - \mathbb{E}\left[H_C(\theta^{(1)}_{N\setminus C}) \middle| \theta^{(1)}_{N\setminus C} \leq \theta \right] \right) = \max_{\theta\in[1,2]} \left(-\theta^2 + \frac{13}{4}\theta - \frac{9}{4} \right) = \frac{25}{64} > \frac{1}{3} = U^{\hat{M}}(\bar{\theta}).$$

Remark 2. *That Condition* (R) *is often not very restrictive suggests that the particular notion of collusion-proofness does not play a significant role. In particular, the second-best outcome is SCP implementable whenever it is WCP implementable if the cartel is all-inclusive or if an additional reasonable condition holds. This suggests that the seemingly pessimistic result of Dequiedt (2007) is not attributable to the cartel's ability to punish a defector, but rather to the binary type structure of his model. In fact, with the binary type, our weak collusion-proof notion would have led to the same result as his. By our Exclusion Principle, a WCP auction must always exclude the low type, but this means that the seller can never do better than selling the good at the price equal to the high valuation. Theorem 6 suggests that, given non-binary types, the seller can typically do strictly better than this even when the cartel can commit to punish the defector (as has been assumed in our SCP notion).*

7 Conclusion

We have studied optimal collusion-proof auctions when a group of bidders can collude not only on their messages (e.g., "bids"), but also on their participation decisions. Despite this strong collusive power, we have shown that the asymmetric information facing the collusive bidders can be exploited to significantly weaken their collusive power, by eliminating the scope of collusive arrangements that could make all collusive bidders uniformly better off regardless of their types. We show that the second-best outcome is achievable if a cartel is not all-inclusive (which will be the case either if there is a noncollusive bidder or if there are multiple bidding cartels), or if the outcome involves a nontrivial probability of the object not being allocated to any bidder. More generally, we have shown that the optimal collusion-proof auction rule involves a positive probability of the object not being allocated to a collusive bidder.

Our results have two broad implications. First, unlike the prevailing impression based on the existing literature, the presence of bidder collusion need not mean that the seller can do no better than textbook monopoly pricing. Our seller can do as well as if there is no collusion in a broad set of circumstances, and do generally much better than simple monopoly pricing. Second, an auction rule different from standard may be more desirable when bidder collusion is a serious issue. We have identified an auction with a group-based sale price as being unsusceptible to collusion.

It is legitimate to ask whether our collusion-proof auction rule is used in practice, but the answer is not immediately clear. Our collusion-proof auction may be implemented in different ways, some of which may not even resemble an auction. For instance, our auction may be implemented by a seller who negotiates with a group of organized bidders for a single sale price. Indeed, it is quite common for a procurer of a service or a good to negotiate with a prime contractor acting as a representative of a group of contractors. Whether such a collective negotiation approach serves as a response to possible collusion among contractors is an interesting, yet unresolved, question.

It is of course quite possible that our collusion-proof auction has no real world correspondence. To the extent that this is true, there may be a couple of reasons. One may be simply that collusion is not a serious enough problem in many scenarios, at least serious enough to depart from a standard auction. Alternatively, the reason may be attributable to two restrictive features of our auction design. First, our collusion-proof auction involves Bayesian implementation, which does rely on bidders' common knowledge of priors, i.e. the assumption that bidders know the distribution of other bidders' types, know that they know the distribution, and so on and so forth. It is unlikely for bidders to possess such a knowledge, especially when they lack the opportunity to develop such a knowledge from repeated interactions. Relaxing the common knowledge of priors by strengthening the solution concept, say to dominant strategies, seems to be an important next step in the research of collusion-proof auctions. The implication of such an extension is not *a priori* obvious, however, since it affects the contracting problem at both ends, i.e., for both the auction designer and for the colluding bidders. Collusion may indeed become easier to prevent if dominant-strategy incentives are very difficult to provide in a budget-balanced fashion among the colluders. Second, we have assumed that the seller has accurate information about the cartel structure; i.e., which bidders belong to what cartel. While this assumption is not unrealistic in many situations, it would be better if the auction design need not require specific knowledge about the cartel structure. In a sense, an important lesson from the current paper may be the highlighting of these features as further challenges to overcome in collusion-proof auction design.

Appendix: Proofs

Proof of Lemma 0. It suffices to show that (IC^*) implies the following: for all $i \in N$ and all $\tilde{\theta}_i \in \Theta_i$,

$$U_i^M(\theta_i) = \mathbb{E}\left[K_i(\tilde{\theta}_i)Q_i(\tilde{\theta}_i)1_{\{\tilde{\theta}_i \leq \theta_i\}} + J_i(\tilde{\theta}_i)Q_i(\tilde{\theta}_i)1_{\{\tilde{\theta}_i \geq \theta_i\}} - T_i(\tilde{\theta}_i)\right]. \quad (14)$$

Note first that a well-known necessary condition for (IC^*) is: for all $i \in N$ and all $\tilde{\theta}_i, \theta_i \in \Theta_i$,

$$U_i^M(\theta_i) - U_i^M(\tilde{\theta}_i) = \int_{\tilde{\theta}_i}^{\theta_i} Q_i(a)da. \quad (15)$$

We show that (15) implies (14). Since $U_i^M(\tilde{\theta}_i) = \tilde{\theta}_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i)$, (15) becomes

$$U_i^M(\theta_i) = \tilde{\theta}_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i) + \int_{\tilde{\theta}_i}^{\theta_i} Q_i(a)da.$$

Taking expectation on both sides regarding $\tilde{\theta}_i$ yields

$$\begin{aligned}
U_i^M(\theta_i) &= \mathbb{E}[\tilde{\theta}_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i)] + \int_{\underline{\theta}_i}^{\overline{\theta}_i} \int_{\tilde{\theta}_i}^{\theta_i} Q_i(a)dadF_i(\tilde{\theta}_i) \\
&= \mathbb{E}[\tilde{\theta}_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i)] + \int_{\underline{\theta}_i}^{\theta_i} \int_{\tilde{\theta}_i}^{\theta_i} Q_i(a)dadF_i(\tilde{\theta}_i) + \int_{\theta_i}^{\overline{\theta}} \int_{\tilde{\theta}_i}^{\theta_i} Q_i(a)dadF_i(\tilde{\theta}_i) \\
&= \mathbb{E}[\tilde{\theta}_i Q_i(\tilde{\theta}_i) - T_i(\tilde{\theta}_i)] + \int_{\underline{\theta}_i}^{\theta_i} Q_i(\tilde{\theta}_i)F_i(\tilde{\theta}_i)d\tilde{\theta}_i - \int_{\theta_i}^{\overline{\theta}} Q_i(\tilde{\theta}_i)(1 - F_i(\tilde{\theta}_i))d\tilde{\theta}_i \\
&= \mathbb{E}\left[K_i(\tilde{\theta}_i)Q_i(\tilde{\theta}_i)1_{\{\tilde{\theta}_i \leq \theta_i\}} + J_i(\tilde{\theta}_i)Q_i(\tilde{\theta}_i)1_{\{\tilde{\theta}_i \geq \theta_i\}} - T_i(\tilde{\theta}_i)\right],
\end{aligned}$$

where the third equality follows from the integration by parts. ▌

Proof of Lemma 1. To begin with, define $T_C(\theta_C) := \mathbb{E}_{\tilde{\theta}_{N\setminus C}}[\sum_{i\in C} t_i(\theta_C, \tilde{\theta}_{N\setminus C})]$ and $\mathcal{Q}_C^s := \{Q \in \mathcal{Q}_C | Q = Q_C(\theta_C) \text{ for some } \theta_C \in \overline{\Theta}_C\}$. Then, let us define $r : \mathcal{Q}_C \to \mathbb{R}_+$ as the greatest convex function such that for all $Q \in \mathcal{Q}_C^s$,

$$r(Q) \leq \inf\{T_C(\theta_C) | Q_C(\theta_C) = Q\}.$$

We show that $r(Q_C(\theta_C)) = T_C(\theta_C)$ for almost every θ_C. Suppose not. Then, it must be that for some $\epsilon > 0$,

$$\mathbb{E}\left[r(Q_C(\theta_C))\right] + \epsilon < \mathbb{E}\left[T_C(\theta_C)\right]. \tag{16}$$

Also, by the definition of $r(\cdot)$, it is possible to find $\mu^C : \Theta_C \to \Delta\overline{\Theta}_C$ satisfying that for all θ_C,

$$\mathbf{E}_{\mu^C(\theta_C)}\left[T_C(\tilde{\theta}_C)\right] \leq r(Q(\theta_C)) + \epsilon \text{ and } \mathbf{E}_{\mu^C(\theta_C)}[Q_C(\tilde{\theta}_C)] = Q_C(\theta_C). \tag{17}$$

We now show that M cannot be WCP with respect to C by constructing a weakly feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ of M by cartel C with which some bidder is better off than with M.

Let the cartel manipulate its type reports using $\mu^C(\cdot)$, whereafter, the object is reallocated to bidder i with probability $w_i(\theta_C) := \frac{q_i^C(\theta_C)}{Q_C(\theta_C)}$ so that $\sum_{i\in C} w_i(\theta_C) = 1$, satisfying ($RC_C^M$). Note that the interim allocation for each collusive bidder $i \in C$ is preserved since

$$\begin{aligned}
\tilde{q}_i^C(\theta_C) &= \omega_i(\theta_C)\mathbb{E}_{\tilde{\theta}_{N\setminus C}}\left[\mathbf{E}_{\mu^C(\theta_C)}[q_C(\tilde{\theta}_C, \tilde{\theta}_{N\setminus C})]\right] \\
&= \omega_i(\theta_C)\mathbf{E}_{\mu^C(\theta_C)}[Q_C(\tilde{\theta}_C)] = \omega_i(\theta_C)Q_C(\theta_C) = q_i^C(\theta_C),
\end{aligned} \tag{18}$$

where the second equality follows from changing the order of expectations, the third from (17), and the last from the definition of $\omega_i(\cdot)$.

Next, the cartel manipulates the transfer rule as follows: Letting $t_i^\mu(\theta) := \mathbf{E}_{\mu^C(\theta_C)}[t_i(\tilde{\theta}_C, \theta_{N\backslash C})]$, set $\tilde{t}_j(\theta) = t_j^\mu(\theta)$ for each $j \in N\backslash C$, and for each $i \in C$,

$$\tilde{t}_i(\theta) = t_i^\mu(\theta) + \mathbb{E}_{\tilde{\theta}_{-i}}\left[t_i(\theta_i, \tilde{\theta}_{-i}) - t_i^\mu(\theta_i, \tilde{\theta}_{-i})\right] - \frac{1}{|C|-1}\sum_{j\in C\backslash\{i\}} \mathbb{E}_{\tilde{\theta}_{-j}}\left[t_j(\theta_j, \tilde{\theta}_{-j}) - t_j^\mu(\theta_j, \tilde{\theta}_{-j})\right] + \sigma_i,$$

where $\sum_{i\in C}\sigma_i = 0$. Note that $\sum_{i\in C}\tilde{t}_i(\theta) = \sum_{i\in C} t_i^\mu(\theta)$, which satisfies (BB_C^M) while $(BB_{N\backslash C}^M)$ is obviously satisfied. Also,

$$\mathbb{E}_{\tilde{\theta}_{-i}}\left[\tilde{t}_i(\theta_i, \tilde{\theta}_{-i})\right] = \mathbb{E}_{\tilde{\theta}_i}\left[t(\theta_i, \tilde{\theta}_{-i})\right] - \kappa_i, \tag{19}$$

where

$$\kappa_i := \frac{1}{|C|-1}\sum_{j\in C\backslash\{i\}} \mathbb{E}\left[t_j(\theta) - t_j^\mu(\theta)\right] - \sigma_i.$$

Then, one can choose $\sigma_i's$ so that $\kappa_i \geq 0, \forall i \in C$, since

$$\sum_{i\in C}\kappa_i = \mathbb{E}\left[\sum_{i\in C}(t_i(\theta) - t_i^\mu(\mu))\right] = \mathbb{E}\left[T_C(\theta_C) - \mathbf{E}_{\mu^C(\theta_C)}[T_C(\tilde{\theta}_C)]\right] > 0, \tag{20}$$

where the inequality follows from (16) and (17). So, (IC^*) and (IR_N^C) are satisfied for collusive bidders, due to (18), (19), and $\kappa_i \geq 0, \forall i \in C$, which means that $\tilde{M}$ is a weakly feasible manipulation of M. Also, some collusive bidder is better off than in M since $\kappa_j > 0$ for some $j \in C$.▌

Proof of Lemma 2. To begin, we adopt the convention that $\theta_\emptyset < \underline{\theta}_i$ for all $i \in N$. Observe that $Q_C(\cdot)$ and $q_i^C(\cdot)$ inherit the monotonicity of $q_C(\cdot)$ and $q_i(\cdot)$, respectively, and hence are a.e. continuous. Also, since $r(\cdot)$ is convex with $r(0) = 0$, $p(\cdot)$ is nondecreasing and hence a.e. continuous also. Suppose to the contrary that (2) does not hold for almost every type profile. Then, we can find some bidder $k \in C$ and a positive measure set $\hat{\Theta}_{C-k} \subset \Theta_{C-k}$ such that for each $\theta_{C-k} \in \hat{\Theta}_{C-k}$, there exist $\theta_k \in \Theta_k$ and $\theta_k' \in \overline{\Theta}$ satisfying

$$(K_k(\theta_k) - p(\theta_k, \theta_{C-k}))q_i^C(\theta_k, \theta_{C-k}) < (K_k(\theta_k) - p(\theta_k', \theta_{C-k}))q_i^C(\theta_k', \theta_{C-k}).$$

Then, the a.e. continuity of $q_i^C(\cdot)$ and $p(\cdot)$ guarantees that for each $\theta_{C-k} \in \hat{\Theta}_{C-k}$, we can find two types $\hat{\theta}_k(\theta_{C-k}) \in \overline{\Theta}$ and $\hat{\hat{\theta}}_k(\theta_{C-k}) > \hat{\theta}_k(\theta_{C-k})$ such that for all $\theta_k \in (\hat{\theta}_k(\theta_{C-k}), \hat{\hat{\theta}}_k(\theta_{C-k})]$,

$$(K_k(\theta_k) - p(\theta_k, \theta_{C-k}))q_i^C(\theta_k, \theta_{C-k}) < (K_k(\theta_k) - p(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}))q_i^C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}). \tag{21}$$

We now define

$$\hat{\Theta}_C := \{(\theta_k, \theta_{C-k}) \in \Theta | \theta_{C-k} \in \hat{\Theta}_{C-k} \text{ and } \theta_k \in (\hat{\theta}_k(\theta_{C-k}), \hat{\hat{\theta}}_k(\theta_{C-k})]\},$$

$\hat{q}_k^C(\theta_{C-k}) := q_k^C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})$, and $\hat{p}(\theta_{-k}) := p(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})$. Note that (21) holds for all $\theta_C \in \hat{\Theta}_C$.

In order to draw a contradiction, we construct a weakly feasible manipulation of M, $\tilde{M} = (\tilde{q}, \tilde{t})$, which makes bidder k better off.

Consider the following report manipulation, denoted $\mu^C : \Theta_C \to \Delta\overline{\Theta}_C$, and reallocation scheme by the cartel: if $\theta_C \notin \hat{\Theta}_C$, then report truthfully and do not perform any reallocation while if $\theta_C \in \hat{\Theta}_C$, then (i) report truthfully with probability $\frac{\sum_{i \in C\backslash\{k\}} q_i^C(\theta_C)}{Q_C(\theta_C)}$ and, once the object is assigned, reallocate it to bidder $i \in C\backslash\{k\}$ with probability $\frac{q_i^C(\theta)}{\sum_{i \in C\backslash\{k\}} q_i^C(\theta_C)}$, (ii) report $(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})$ (or $(\theta_\emptyset, \cdots, \theta_\emptyset)$ in case $\hat{\theta}_k(\theta_{C-k}) = \theta_\emptyset$) with probability $\frac{\hat{q}_k^C(\theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k})), \theta_{C-k})}$ and, once the object is assigned, reallocate it to bidder k with probability 1, and (iii) choose $(\theta_\emptyset, \cdots, \theta_\emptyset)$ (or nonparticipation) with the remaining probability.[16] This manipulation will result in the following allocation probabilities: for bidder $i \in C\backslash\{k\}$,

$$\tilde{q}_i^C(\theta_C) = Q_C(\theta_C) \frac{\sum_{i \in C\backslash\{k\}} q_i^C(\theta_C)}{Q_C(\theta_C)} \frac{q_i^C(\theta_C)}{\sum_{i \in C-k} q_i^C(\theta_C)} = q_i^C(\theta_C) \text{ if } \theta_C \in \hat{\Theta}_C,$$

and simply $\tilde{q}_i^C(\theta_C) = q_i^C(\theta_C)$ if $\theta_C \notin \hat{\Theta}_C$. Likewise, for bidder k, if $\theta_C \notin \hat{\Theta}_C$, then $\tilde{q}_k^C(\theta_C) = q_k^C(\theta_C)$, and if $\theta_C \in \hat{\Theta}_C$, then

$$\tilde{q}_k^C(\theta_C) = Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}) \frac{\hat{q}_k^C(\theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})} = \hat{q}_k^C(\theta_{C-k}). \tag{22}$$

It can be easily verified that $\tilde{q}_k^C(\cdot, \theta_{C-k})$ is nondecreasing for every θ_{C-k}.[17] Thus, the interim allocation $\tilde{Q}_i(\theta_i) = \mathbb{E}_{\tilde{\theta}_{C-i}}[\tilde{q}_i^C(\theta_i, \tilde{\theta}_{C-i})]$ is also nondecreasing for each $i \in C$.

---

[16] It is important to make sure that the probability of reporting truthfully or $(\hat{\theta}(\theta_{C-k}), \theta_{C-k})$ does not exceed 1, for which it suffices to verify that $\frac{\hat{q}_k^C(\theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})} \leq \frac{q_k^C(\theta_C)}{Q_C(\theta_C)}$. This holds trivially if $\hat{\theta}_k(\theta_{C-k}) = \theta_\emptyset$. If $\hat{\theta}_k(\theta_{C-k}) \neq \theta_\emptyset$, it holds since

$$\frac{\hat{q}_k^C(\theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})} = 1 - \frac{\sum_{i \in C\backslash\{k\}} q_i^C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k})} \leq 1 - \frac{\sum_{i \in C\backslash\{k\}} q_i^C(\theta_C)}{Q_C(\theta_C)} = \frac{q_k^C(\theta_C)}{Q_C(\theta_C)},$$

where the inequality holds since $Q_C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}) \leq Q_C(\theta_k, \theta_{C-k})$ and $q_i^C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}) \geq q_i^C(\theta_k, \theta_{C-k})$ for all $i \neq k$, by the monotonicity of $Q_C(\cdot)$ and $q_i^C(\cdot)$.

[17] To see this, consider arbitrary θ_k and θ'_k with $\theta'_k > \theta_k$, and θ_{C-k}: (i) if $(\theta_k, \theta_{C-k}), (\theta'_k, \theta_{C-k}) \in \hat{\Theta}_C$, then $\tilde{q}_k^C(\theta_k, \theta_{C-k}) = \tilde{q}_k^C(\theta'_k, \theta_{C-k}) = \hat{q}_k^C(\theta_{C-k})$, (ii) if $(\theta_k, \theta_{C-k}) \in \hat{\Theta}_C$ and $(\theta'_k, \theta_{C-k}) \notin \hat{\Theta}_C$, then $\hat{\theta}_k(\theta_{C-k}) < \theta_k \leq \hat{\hat{\theta}}_k(\theta_{C-k}) < \theta'_k$ and thus $\tilde{q}_k^C(\theta_k, \theta_{C-k}) = q_k^C(\hat{\theta}_k(\theta_{C-k}), \theta_{C-k}) \leq q_k^C(\theta'_k, \theta_{C-k}) = \tilde{q}_k^C(\theta'_k, \theta_{C-k})$. And other cases can be dealt with similarly.

After the manipulation, the cartel's aggregate payment becomes

$$\mathbb{E}_{\tilde{\theta}_{N\backslash C}}\left[\mathbf{E}_{\mu^C(\theta_C)}[\sum_{i\in C} t_i(\tilde{\theta}_C,\tilde{\theta}_{N\backslash C})]\right]$$

$$= \begin{cases} r(Q_C(\theta_C))\dfrac{\sum_{i\in C\backslash\{k\}} q_i^C(\theta_C)}{Q_C(\theta_C)} + r(Q_C(\hat{\theta}_k(\theta_{C-k}),\theta_{C-k}))\dfrac{\hat{q}_k^C(\theta_{C-k})}{Q(\hat{\theta}_k(\theta_{C-k}),\theta_{C-k})} & \text{if } \theta_C\in\hat{\Theta}_C \\ r(Q_C(\theta_C)) & \text{otherwise,} \end{cases}$$

which yields

$$\mathbb{E}\left[\mathbf{E}_{\mu^C(\theta_C)}[\sum_{i\in C} t_i(\tilde{\theta}_C,\theta_{N\backslash C})]\right]$$

$$= \mathbb{E}\left[r(Q_C(\theta_C))\right] + \mathbb{E}_{\theta_C\in\hat{\Theta}_C}\left[r(Q(\hat{\theta}_k(\theta_{C-k}),\theta_{C-k}))\frac{\hat{q}_k^C(\theta_{C-k})}{Q_C(\hat{\theta}_k(\theta_{C-k}),\theta_{C-k})} - r(Q_C(\theta_C))\frac{q_k^C(\theta_C)}{Q_C(\theta_C)}\right]$$

$$= \mathbb{E}\left[\sum_{i\in C} t_i(\theta)\right] + \mathbb{E}_{\theta_C\in\hat{\Theta}_C}\left[\hat{p}(\theta_{C-k})\hat{q}_k^C(\theta_{C-k}) - p(\theta_C)q_k^C(\theta_C)\right] \tag{23}$$

Next, $\tilde{t}(\cdot)$ is constructed as follows. For each $j\in N\backslash C$, set $\tilde{t}_j(\theta)=\mathbf{E}_{\mu^C(\theta_C)}[t_j(\tilde{\theta}_C,\theta_{N\backslash C})]$. For each $i\in C$, we set

$$\tilde{t}_i(\theta)=\mathbf{E}_{\mu^C(\theta_C)}[t_i(\tilde{\theta}_C,\theta_{N\backslash C})]+Y_i(\theta_i)-\frac{1}{|C|-1}\sum_{j\in C\backslash\{i\}}Y_j(\theta_j)+\rho_i,$$

where

$$Y_i(\theta_i):=\theta_i\tilde{Q}_i(\theta_i)-\int_{\underline{\theta}_i}^{\theta_i}\tilde{Q}_i(a)da-\mathbb{E}_{\theta_{-i}}[\mathbf{E}_{\mu^C(\theta_C)}[t_i(\tilde{\theta}_C,\theta_{N\backslash C})]],$$

and

$$\rho_i:=\frac{1}{|C|-1}\mathbb{E}_{\theta_{-i}}\left[\sum_{j\in C\backslash\{i\}}Y_j(\theta_j)\right]-U_i^M(\underline{\theta}_i) \text{ for } i\in C\backslash\{k\}, \text{ and } \rho_k=-\sum_{i\in C\backslash\{k\}}\rho_i.$$

By construction, then $\tilde{t}$ satisfies (BB_C^M) and $(BB_{N\backslash C}^M)$.

We now complete the proof by showing that $\tilde{M}$ is a weakly feasible manipulation and makes bidder k better off. To this end, observe that for an arbitrary $\theta_k\in\Theta_k$,

$$U_k^{\tilde{M}}(\theta_k)+\sum_{i\in C\backslash\{k\}}U_i^{\tilde{M}}(\underline{\theta}_i)$$

$$= \mathbb{E}\left[\left(K_k(\tilde{\theta}_k)\tilde{q}_k(\tilde{\theta})\right)1_{\{\tilde{\theta}_k<\theta_k\}}+\left(J_k(\tilde{\theta}_k)\tilde{q}_k(\tilde{\theta})\right)1_{\{\tilde{\theta}_k>\theta_k\}}+\sum_{i\in N\backslash\{k\}}J_i(\tilde{\theta}_i)q_i(\tilde{\theta})-\sum_{i\in C}\tilde{t}_i(\tilde{\theta})\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left(K_k(\tilde{\theta}_k)q_k(\tilde{\theta})\right)1_{\{\tilde{\theta}_k<\theta_k\}} + \left(J_k(\tilde{\theta}_k)q_k(\tilde{\theta})\right)1_{\{\tilde{\theta}_k>\theta_k\}} + \sum_{i\in N\setminus\{k\}} J_i(\tilde{\theta}_i)q_i(\tilde{\theta}) - \sum_{i\in C} t_i(\tilde{\theta})\right] \\
&\quad +\mathbb{E}_{\tilde{\theta}_C\in\hat{\Theta}_C}\Big[K_k(\tilde{\theta}_k)(\tilde{q}_k^C(\tilde{\theta}_C) - q_k^C(\tilde{\theta}_C))1_{\{\tilde{\theta}_k<\theta_k\}} + J_k(\tilde{\theta}_k)(\tilde{q}_k^C(\tilde{\theta}_C) - q_k^C(\tilde{\theta}_C))1_{\{\tilde{\theta}_k>\theta_k\}} \\
&\qquad - \left(\hat{p}(\tilde{\theta}_{C-k})\hat{q}_k^C(\tilde{\theta}_{C-k}) - p(\tilde{\theta}_C)q_k^C(\tilde{\theta}_C)\right)\Big] \\
&= U_k^M(\theta_k) + \sum_{i\in C\setminus\{k\}} U_i^M(\underline{\theta}_i) \\
&\quad +\mathbb{E}_{\tilde{\theta}_C\in\hat{\Theta}_C}\Big[\left((K_k(\tilde{\theta}_k) - \hat{p}(\tilde{\theta}_{C-k}))\hat{q}_k^C(\tilde{\theta}_{C-k}) - (K_k(\tilde{\theta}_k) - p_k(\tilde{\theta}_C))q_k^C(\tilde{\theta}_C)\right)1_{\{\tilde{\theta}_k<\theta_k\}} \\
&\qquad + \left((J_k(\tilde{\theta}_k) - \hat{p}(\tilde{\theta}_{C-k}))\hat{q}_k^C(\tilde{\theta}_{C-k}) - (J_k(\tilde{\theta}_k) - p_k(\tilde{\theta}_C))q_k^C(\tilde{\theta}_C)\right)1_{\{\tilde{\theta}_k>\theta_k\}}\Big] \\
&> U_k^M(\theta_k) + \sum_{i\in C\setminus\{k\}} U_i^M(\underline{\theta}_i). \qquad (24)
\end{aligned}
$$

The first equality follows from Lemma 0, the second from (23), the third from the rearrangement and (22), and the inequality from (21) and the fact that for all $\tilde{\theta}_C \in \hat{\Theta}_C$,

$$J_k(\tilde{\theta}_k)(\hat{q}_k^C(\tilde{\theta}_{C-k}) - q_k^C(\tilde{\theta}_C)) \geq K_k(\tilde{\theta}_k)(\hat{q}_k^C(\tilde{\theta}_{C-k}) - q_k^C(\tilde{\theta}_C)),$$

since $\hat{q}_k^C(\tilde{\theta}_{C-k}) \leq q_k^C(\tilde{\theta}_C)$ and $J_k(\tilde{\theta}_k) < K_k(\tilde{\theta}_k)$.

From the construction of $\tilde{t}(\cdot)$, one can easily verify that $U_i^{\tilde{M}}(\underline{\theta}_i) = U_i^M(\underline{\theta}_i)$, $\forall i \in C\setminus\{k\}$. Then, (24) implies $U_k^{\tilde{M}}(\theta_k) > U_k^M(\theta_k)$ for all $\theta_k \in \Theta_k$ or bidder k is better off. The construction of $\tilde{t}(\cdot)$ and monotonicity of $\tilde{Q}_i(\cdot), \forall i \in C$ guarantee that $\tilde{M}$ satisfies (IC^*) for all collusive bidders. The proof will be complete if (IR_C^M) holds for all $i \in C\setminus\{k\}$:

$$U_i^{\tilde{M}}(\theta_i) = U_i^{\tilde{M}}(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \tilde{Q}_i(\tilde{\theta})d\tilde{\theta} = U_i^M(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} Q_i(\tilde{\theta})d\tilde{\theta} = U_i^M(\theta_i), \forall \theta_i \in \Theta_i$$

since $U_i^{\tilde{M}}(\underline{\theta}_i) = U_i^M(\underline{\theta}_i)$ and $\tilde{Q}_i(\cdot) = Q_i(\cdot)$, $\forall i \in C\setminus\{k\}$. ▌

Proof of Theorem 1. Suppose that an auction rule $M = (q, t) \in \mathcal{M}$ WCP implements the second-best outcome. Then, $q(\cdot) = q^*(\cdot)$ and $U_i(\underline{\theta}_i) = 0$ for all $i \in N$, which implies by Lemma 0 that for any $C \subset N$,

$$\mathbb{E}\left[\sum_{i\in C} t_i(\theta)\right] = \mathbb{E}\left[\sum_{i\in C} J_i(\theta_i)q_i^*(\theta_i)\right]. \qquad (25)$$

By Lemma 1, there exists a convex function $r(\cdot)$ that represents the total payment for the cartel.

We first consider the case $\mathcal{C} = \{N\}$. Since $q_N^*(\theta) = 0$ or 1 for all $\theta \in \Theta$, Lemma 1 implies that $p(\theta) = r^*$ whenever $q_N^*(\theta) = 1$. We first prove $\hat{\theta}_i > \underline{\theta}_i$ for all $i \in N$. Suppose not. Then, there exists k such that $J_k(\underline{\theta}_k) \geq \max\{\max_{i \in N \setminus \{k\}} J_i(\underline{\theta}_i), 0\}$. It follows that $q_k^*(\underline{\theta}_1, \cdots, \underline{\theta}_n) > 0$, so $p(\underline{\theta}_1, \cdots, \underline{\theta}_n) = r^*$. Since $r^* \geq V^* > \underline{\theta}_i = K_i(\underline{\theta}_i)$, we have a contradiction to (2).

We next consider the case $\mathcal{C} \neq \{N\}$. Fix any C with $|C| \geq 2$. If no such C exists, there is no collusion, so we are done. For each bidder $i \in C$ and his type $\theta_i \in \Theta_i$, let $X_i(\theta_i) := \Pr\{\theta_{C-i} \in \Theta_{C-i} \mid J_i(\theta_i) > \max_{k \in C \setminus \{i\}} J_k(\theta_k)\}$ be the probability that i has the highest virtual value among the collusive bidders, and let $Y_i(\theta_i) := \Pr\{\theta_{N \setminus C} \in \Theta_{N \setminus C} \mid J_i(\theta_i) > \max\{\max_{k \in N \setminus C} J_k(\theta_k), 0\}\}$. Letting $p_i(\theta_i) := \frac{r(Y_i(\theta_i))}{Y_i(\theta_i)}$ for each $i \in C$, Lemma 2 implies that, $\forall \theta_i \geq \hat{\theta}_i$

$$(K_i(\theta_i) - p_i(\theta_i))\, Y_i(\theta_i) \geq \max\{0, \max_{\theta_i' \in [\hat{\theta}_i, \theta_i]} (K_i(\theta_i) - p_i(\theta_i'))\, Y_i(\theta_i')\}.$$

By the envelope theorem argument, $\forall \theta_i \geq \hat{\theta}_i$,

$$(K_i(\theta_i) - p_i(\theta_i))Y_i(\theta_i) \geq (K_i(\hat{\theta}_i) - p_i(\hat{\theta}_i))Y_i(\hat{\theta}_i) + \int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)da \geq \int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)da$$

or

$$p_i(\theta_i)Y_i(\theta_i) \leq K_i(\theta_i)Y_i(\theta_i) - \int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)da.$$

Thus, we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{i \in C} t_i(\theta)\right] &= \mathbb{E}\left[\sum_{i \in C} r(Y_i(\theta_i))X_i(\theta_i)\right] \\
&= \mathbb{E}\left[\sum_{i \in C} p_i(\theta_i)Y_i(\theta_i)X_i(\theta_i)\right] \\
&\leq \mathbb{E}\left[\sum_{i \in C} \left(K_i(\theta_i)Y_i(\theta_i) - \int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)da\right) X_i(\theta_i)\right]. \qquad (26)
\end{aligned}$$

Letting $Z_i(\theta_i) = \int_{\theta_i}^{\bar{\theta}_i} X_i(s)dF_i(s)$,

$$\begin{aligned}
&\mathbb{E}\left[\left(K_i(\theta_i)Y_i(\theta_i) - \int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)da\right) X_i(\theta_i)\right] \\
=& \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i(\theta_i)Y_i(\theta_i)X_i(\theta_i)dF_i(\theta_i) - \int_{\hat{\theta}_i}^{\bar{\theta}_i}\int_{\hat{\theta}_i}^{\theta_i} K_i'(a)Y_i(a)daX_i(\theta_i)dF_i(\theta_i) \\
=& \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i(\theta_i)X_i(\theta_i)Y_i(\theta_i)dF_i(\theta_i) - \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i'(\theta_i)Y_i(\theta_i)Z_i(\theta_i)d\theta_i
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i(\theta_i)X_i(\theta_i)Y_i(\theta_i)dF_i(\theta_i) - K_i(\theta_i)Y_i(\theta_i)Z_i(\theta_i)\Big|_{\hat{\theta}_i}^{\bar{\theta}_i} \\
&\quad + \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i(\theta_i)\left(Y_i'(\theta_i)Z_i(\theta_i) + Y_i(\theta_i)Z_i'(\theta_i)\right)d\theta_i \\
&= K_i(\hat{\theta}_i)Y_i(\hat{\theta}_i)Z_i(\hat{\theta}_i) + \int_{\hat{\theta}_i}^{\bar{\theta}_i} K_i(\theta_i)Y_i'(\theta_i)Z_i(\theta_i)d\theta_i \\
&= \mathbb{E}\left[K_i(\phi_i(\theta_{N\setminus C}))q_i^*(\theta)\right].
\end{aligned}
$$

The second and fourth equalities follow from integration by parts. To verify the fifth equality, note that $Y_i(\hat{\theta}_i) = \Pr\{\phi_i(\theta_{N\setminus C}) = \hat{\theta}_i\}$, $Y_i(s) = \Pr\{\phi_i(\theta_{N\setminus C}) \leq s\}$ for each $s > \hat{\theta}_i$, and $Z_i(s) = \mathbb{E}[q_i^*(\theta)|\phi_i(\theta_{N\setminus C}) = s]$. Combine this derivation with (25) and (26) to obtain (ii) of Condition (SB). ▌

Proof of Lemma 3. First, we prove that Condition (SB') holds for any C with $2 \leq |C| < n$. To this end, observe that

$$
\begin{aligned}
&\mathbb{E}\left[K(\theta_{N\setminus C}^{(1)})1_{\{\theta_C^{(1)} > \theta_{N\setminus C}^{(1)}\}}\right] \\
&= K(\hat{\theta})(1-F^{|C|}(\hat{\theta}))F^{N-|C|}(\hat{\theta}) + \int_{\hat{\theta}}^{\bar{\theta}} \left(\theta + \frac{F(\theta)}{f(\theta)}\right)(1-F^{|C|}(\theta))dF^{N-|C|}(\theta) \\
&= K(\hat{\theta})(1-F^{|C|}(\hat{\theta}))F^{N-|C|}(\hat{\theta}) + \int_{\hat{\theta}}^{\bar{\theta}} \theta(1-F^{|C|}(\theta))dF^{N-|C|}(\theta) \\
&\qquad + \int_{\hat{\theta}}^{\bar{\theta}} (N-|C|)(1-F^{|C|}(\theta))F^{N-|C|}(\theta)d\theta. \qquad (27)
\end{aligned}
$$

Observe also that

$$
\begin{aligned}
&\mathbb{E}\left[J(\theta_C^{(1)})1_{\{\theta_C^{(1)} > \theta_{N\setminus C}^{(1)}\}}\right] = \int_{\hat{\theta}}^{\bar{\theta}} \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) F^{N-|C|}(\theta)dF^{|C|}(\theta) \\
&= -(1-F^{|C|}(\theta))\theta F^{N-|C|}(\theta)\Big|_{\hat{\theta}}^{\bar{\theta}} + \int_{\hat{\theta}}^{\bar{\theta}} (1-F^{|C|}(\theta))d\left(\theta F^{N-|C|}(\theta)\right) - \int_{\hat{\theta}}^{\bar{\theta}} |C|(1-F(\theta))F^{N-1}(\theta)d\theta \\
&= \hat{\theta}(1-F^{|C|}(\hat{\theta}))F^{N-|C|}(\hat{\theta}) + \int_{\hat{\theta}}^{\bar{\theta}} (1-F^{|C|}(\theta))F^{N-|C|}(\theta)d\theta \\
&\qquad + \int_{\hat{\theta}}^{\bar{\theta}} \theta(1-F^{|C|}(\theta))dF^{N-|C|}(\theta) - \int_{\hat{\theta}}^{\bar{\theta}} |C|(1-F(\theta))F^{N-1}(\theta)d\theta,
\end{aligned}
$$

where the second equality follows from integration by parts. Subtracting this expression from (27) yields

$$
\mathbb{E}\left[\sum_{i\in C}\left(K_i(\phi_i(\theta_{N\setminus C})) - J_i(\theta_i)\right)q_i^*(\theta)\right]
$$

$$
\begin{aligned}
&= \mathbb{E}\left[\left(K(\theta^{(1)}_{N\setminus C}) - J(\theta^{(1)}_C)\right) 1_{\{\theta^{(1)}_C > \theta^{(1)}_{N\setminus C}\}}\right] \\
&= (K(\hat{\theta}) - \hat{\theta})(1 - F^{|C|}(\hat{\theta}))F^{N-|C|}(\hat{\theta}) \\
&\qquad + \int_{\hat{\theta}}^{\bar{\theta}} \left[(N - |C| - 1)(1 - F^{|C|}(\theta))F^{N-|C|}(\theta) + |C|(1 - F(\theta))F^{N-1}(\theta)\right] d\theta > 0,
\end{aligned}
$$

satisfying Condition (SB).

In case $C = N$, (3) is just a restatement of Condition (SB) with $\theta^{(1)}_{N\setminus C} = \hat{\theta}$. ▮

Proof of Theorem 2. Since $\hat{M}$ implements V^*, it suffices to prove that $\hat{M}$ is WCP. To this end, consider any $C \in \mathcal{C}$ with $|C| \geq 2$. Suppose all bidders outside C report truthfully, but cartel C contemplates a manipulation of $\hat{M}$, $\tilde{M} = (\tilde{q}, \tilde{t})$, that satisfies (IC^*_C) and $(IR^{\hat{M}}_C)$. Then, there exists a function $\mu^C : \Theta_C \mapsto \Delta\overline{\Theta}_C$ such that $(R^{\hat{M}}_C), (R^{\hat{M}}_{N\setminus C}), (BB^{\hat{M}}_C)$ and $(BB^{\hat{M}}_{N\setminus C})$ hold. Since the same sale price is charged against a cartel no matter how many of its members participate, it cannot gain from non-participation of its members. Hence, without loss, we assume that μ^C places no weight on $\overline{\Theta}\setminus\Theta$.

We first prove that $\tilde{q}(\theta) = q^*(\theta)$ for almost every $\theta \in \Theta$. To this end, suppose this is not the case. Then,

$$
\begin{aligned}
& \alpha_C(\sum_{i\in C} U^{\tilde{M}}_i(\bar{\theta})) + (1 - \alpha_C)(\sum_{i\in C} U^{\tilde{M}}_i(\underline{\theta})) \\
= \; & \mathbb{E}\left[\sum_{i\in C} H_C(\theta_i)\tilde{q}_i(\theta) - \sum_{i\in C} \tilde{t}_i(\theta)\right] \\
= \; & \mathbb{E}\left[\sum_{i\in C} H_C(\theta_i)\tilde{q}_i(\theta) - \sum_{i\in C} \mathbf{E}_{\mu^C(\theta_C)}[\hat{t}_i(\tilde{\theta}_C, \theta_{N\setminus C})]\right] \\
\leq \; & \mathbb{E}\left[\sum_{i\in C} H_C(\theta_i)\tilde{q}_i(\theta) - \mathbf{E}_{\mu^C(\theta_C)}[\delta_C(\tilde{\theta}_C, \theta_{N\setminus C})\right] \\
= \; & \mathbb{E}\left[\sum_{i\in C} H_C(\theta_i)\tilde{q}_i(\theta) - H_C(\theta^{(1)}_{N\setminus C})\mathbf{E}_{\mu^C(\theta_C)}[\sum_{i\in C} q^*_i(\tilde{\theta}_C, \theta_{N\setminus C})]\right] \\
= \; & \mathbb{E}\left[\sum_{i\in C}[H_C(\theta_i) - H_C(\theta^{(1)}_{N\setminus C})]\tilde{q}_i(\theta)\right] \\
< \; & \mathbb{E}\left[\sum_{i\in C}[H_C(\theta_i) - H_C(\theta^{(1)}_{N\setminus C})]q^*_i(\theta)\right] \qquad (28) \\
= \; & \alpha_C(\sum_{i\in C} U^{\hat{M}}_i(\bar{\theta})) + (1 - \alpha_C)(\sum_{i\in C} U^{\hat{M}}_i(\underline{\theta})).
\end{aligned}
$$

The first equality follows from Lemma 0, the second from equation ($BB_C^{\hat{M}}$), the third from the definition of $\delta_C(\cdot)$, the fourth from ($R_C^{\hat{M}}$), and the last equality from the above string of equalities repeated in the reverse order. The weak inequality follows from the construction of $\hat{t}_i(\cdot)$ for $i \in C$ as in (5) and (6). Lastly, the strict inequality follows from the definition of α_C and the strict monotonicity of $H_C(\cdot)$. To see this, we compare the LHS and RHS of the inequality (28) at the ex-post level: (i) if $\theta_k > \max\{\max_{i \in N \setminus \{k\}} \theta_i, \hat{\theta}\}$ for some $k \in C$, then $q_k^*(\theta) = 1 \neq \tilde{q}_k(\theta)$ implies that

$$LHS = \sum_{i \in C} (H_C(\theta_i) - H_C(\theta_{N \setminus C}^{(1)}))\tilde{q}_i(\theta_i) < H_C(\theta_k) - H_C(\theta_{N \setminus C}^{(1)}) = RHS,$$

(ii) if $\theta_k > \max\{\max_{i \in N \setminus \{k\}} \theta_i, \hat{\theta}\}$ for some $k \in N \setminus C$, then any manipulated allocation different from $q^*(\cdot)$ implies $\tilde{q}_k(\theta) < 1$ and $\tilde{q}_{k'}(\theta) > 0$ for some $k' \in C$,[18] and thus

$$LHS = \sum_{i \in C} (H_C(\theta_i) - H_C(\theta_{N \setminus C}^{(1)}))\tilde{q}_i(\theta_i) = \sum_{i \in C} (H_C(\theta_i) - H_C(\theta_k))\tilde{q}_i(\theta_i) < 0 = RHS,$$

(iii) if $\max_{i \in N} \theta_i < \hat{\theta}$, then $\tilde{q}(\theta) \neq q^*(\theta) = \mathbf{0}$ implies that the LHS is negative while the RHS is zero. In sum, the LHS of (28) is always less than the RHS, which means that $\tilde{M}$ worsens the (interim) payoff of either the highest type or the lowest type of at least one collusive bidder. This contradicts that $\tilde{M}$ satisfies ($IR_C^{\hat{M}}$). We have thus proven that $\tilde{q}(\theta) = q^*(\theta)$ for almost every θ.

It follows from this result that the gross surplus realized within C from $\tilde{M}$ is the same as from $\hat{M}$, and, combined with (4), that the cartel pays the same expected payments with $\tilde{M}$ as with $\hat{M}$. Hence, the net total expected payoff accruing to C from $\tilde{M}$ is the same as from $\hat{M}$. Together with ($IR_C^{\hat{M}}$), this implies that no bidder of C is strictly better off from manipulation $\tilde{M}$. Since this is true for all feasible manipulation of $\hat{M}$, we conclude that $\hat{M}$ is WCP. ▌

For the remainder of proofs, we will often use the following transfer rule: Given an allocation rule $q_i(\cdot)$ and a sale price r, if all bidders participate, then for each $i \in N$

$$\begin{aligned} t_i(\theta) \;\; := \;\; & \frac{1}{n} r \sum_{j \in N} q_j(\theta) + \left(T_i(\theta_i) - \frac{1}{n} r \mathbb{E}_{\tilde{\theta}_{-i}} \left[\sum_{j \in N} q_j(\theta_i, \tilde{\theta}_{-i}) \right] \right) \\ & - \frac{1}{n-1} \sum_{k \in N \setminus \{i\}} \left(T_k(\theta_k) - \frac{1}{n} r \mathbb{E}_{\tilde{\theta}_{-k}} \left[\sum_{j \in N} q_j(\theta_k, \tilde{\theta}_{-k}) \right] \right) + \rho_i, \end{aligned} \tag{29}$$

[18] This follows from the fact that noncollusive bidders always report truthfully so collusive bidders can change the allocation only by announcing that one of them has at least $\theta_k > \hat{\theta}$, and getting themselves allocated the object.

where

$$T_i(\theta_i) := \theta_i \mathbb{E}_{\tilde{\theta}_{-i}}\left[q_i(\theta_i, \tilde{\theta}_{-i})\right] - \int_{\underline{\theta}_i}^{\theta_i} \mathbb{E}_{\tilde{\theta}_{-i}}\left[q_i(a, \tilde{\theta}_{-i})\right] da$$

and $\rho_i \in \mathbb{R}$ with $\sum_{i\in N}\rho_i = 0$. If some bidder does not participate, then the payment of r is equally shared among those who participate while others make no payments. Note from this and (29) that $\sum_{i\in N} t_i(\theta) = r\sum_{i\in N} q_i(\theta)$, $\forall\theta\in\overline{\Theta}$, implying that bidders pay a sale price r as long as at least one bidder participates. Note also that if all bidders participate, then

$$\mathbb{E}_{\tilde{\theta}_{-i}}\left[t_i(\theta_i, \tilde{\theta}_{-i})\right] = T_i(\theta_i) + c_i, \forall i, \forall \theta_i,$$

for some constant c_i, implying that the incentive compatibility is satisfied as long as the interim allocation probability is nondecreasing. The (IR) constraint will be checked later wherever required.

Proof of Theorem 3. Construct a transfer rule $\hat{t}(\cdot)$ by substituting $q^*(\cdot)$ and r^* into (29). It is straightforward that one can choose ρ_i's to let $\hat{t}(\cdot)$ satisfy (IR) condition. Thus, $\hat{M} = (q^*, \hat{t})$ satisfies (IC^*) and implements the second-best outcome without collusion.

To prove that $\hat{M}$ is WCP consists of several steps.

STEP 1. *Suppose that a feasible manipulation, $M = (q,t)$, of $\hat{M}$ (by N) satisfies $U_i^M(\underline{\theta}_i) > U_i^{\hat{M}}(\underline{\theta}_i)$ for some $i \in N$. Then, there exists another feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ that satisfies*

$$U_i^{\tilde{M}}(\underline{\theta}_i) = U_i^{\hat{M}}(\underline{\theta}_i), \forall i \in N, \quad and \;\; \mathbb{E}\left[\sum_{i\in N} U_i^{\tilde{M}}(\theta_i)\right] > \mathbb{E}\left[\sum_{i\in N} U_i^{\hat{M}}(\theta_i)\right]. \tag{30}$$

Proof. Available in the Supplementary Material.▌

STEP 2. *For any feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ of $\hat{M}$ that satisfies $U_i^{\tilde{M}}(\underline{\theta}_i) = U_i^{\hat{M}}(\underline{\theta}_i), \forall i \in N$, we have*

$$\int_{\underline{\theta}_i}^{J_i^{-1}(r^*)} (J_i(\theta_i) - r^*)(\tilde{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i \leq 0, \forall i \in N. \tag{31}$$

The inequality holds strictly unless $\tilde{Q}_i(\theta_i) = Q_i^*(\theta_i), \forall \theta_i \leq J_i^{-1}(r^*)$.

Proof. It follows from the assumption on $\tilde{M}$ that for all $i \in N$ and all $\theta_i \in \Theta_i$,

$$X_i(\theta_i) := \int_{\underline{\theta}_i}^{\theta_i} [\tilde{Q}_i(a) - Q_i^*(a)] da \;\; = \;\; U_i^{\tilde{M}}(\theta_i) - U_i^{\tilde{M}}(\underline{\theta}_i) - [U_i^{\hat{M}}(\theta_i) - U_i^{\hat{M}}(\underline{\theta}_i)]$$

$$= \quad U_i^{\tilde{M}}(\theta_i) - U_i^{\hat{M}}(\theta_i) \geq 0.$$

Then, the integration by parts yields

$$\begin{aligned}
& \int_{\underline{\theta}_i}^{J_i^{-1}(r^*)} (J_i(\theta_i) - r^*)(\tilde{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i \\
= \quad & (J_i(\theta_i) - r^*) f_i(\theta_i) X_i(\theta_i) \Big|_{\underline{\theta}_i}^{J_i^{-1}(r^*)} - \int_{\underline{\theta}_i}^{J_i^{-1}(r^*)} X_i(\theta_i) d\left[(J_i(\theta_i) - r^*) f_i(\theta_i)\right] \\
= \quad & -\int_{\underline{\theta}_i}^{J_i^{-1}(r^*)} X_i(\theta_i) d\left[(J_i(\theta_i) - r^*) f_i(\theta_i)\right] \leq 0,
\end{aligned}$$

since $(J_i(\cdot) - r^*) f_i(\cdot)$ is increasing. The inequality is strict unless $X_i(\theta_i) = 0$ for all $\theta_i \leq J_i^{-1}(r^*)$, that is $\tilde{Q}_i(\theta_i) = Q_i^*(\theta_i)$ for all $\theta_i \leq J_i^{-1}(r^*)$. ▮

To state the next step, we define $\Theta^* := \{\theta \in \Theta | \max_{i \in N} J_i(\theta_i) \geq r^*\}$.

Step 3. *For any feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ of $\hat{M}$ by N that satisfies $U_i^{\tilde{M}}(\underline{\theta}_i) = U_i^{\hat{M}}(\underline{\theta}_i), \forall i \in N$, we have $\tilde{Q}_i(\theta_i) = \hat{Q}_i^*(\theta_i)$, $\forall i \in N$, $\forall \theta_i \in \Theta_i$.*

Proof. Consider another allocation rule, $\bar{q}(\cdot)$, with $\bar{q}_i(\theta) = \tilde{q}_i(\theta)$ if $\theta_i \geq J_i^{-1}(r^*)$ and $\bar{q}_i(\theta) = q_i^*(\theta)$ otherwise, and let $\bar{Q}_i(\theta_i) := \mathbb{E}_{\theta_{-i}}[\bar{q}(\theta_i, \theta_{-i})]$, for each $i \in N$. (Whether $\bar{Q}_i(\cdot)$ is monotonic or whether $\bar{q}_i(\cdot)$ is implementable is irrelevant for the subsequent argument.) Then, it holds that

$$\begin{aligned}
& \sum_{i \in N} \int_{J_i^{-1}(r^*)}^{\bar{\theta}_i} (J_i(\theta_i) - r^*)(\tilde{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i \\
= \quad & \sum_{i \in N} \int_{\underline{\theta}_i}^{\bar{\theta}_i} (J_i(\theta_i) - r^*)(\bar{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i \\
= \quad & \mathbb{E}\left[\sum_{i \in N} (J_i(\theta_i) - r^*)(\bar{q}_i(\theta) - q_i^*(\theta))\right] \\
= \quad & \mathbb{E}_{\theta \in \Theta^*}\left[\sum_{i \in N} (J_i(\theta_i) - r^*)(\bar{q}_i(\theta) - q_i^*(\theta))\right] \\
= \quad & \mathbb{E}_{\theta \in \Theta^*}\left[\sum_{i \in N} (J_i(\theta_i) - r^*)\bar{q}_i(\theta)\right] - \mathbb{E}_{\theta \in \Theta^*}\left[\max_{i \in N} J_i(\theta_i) - r^*\right] \leq 0,
\end{aligned} \tag{32}$$

where the inequality follows from the definition of $q^*(\cdot)$ and becomes strict unless $\tilde{q}(\theta) = q^*(\theta)$ for almost all $\theta \in \Theta^*$. Thus, we have

$$0 \quad \leq \quad \sum_{i \in N} [\tilde{U}_i(\underline{\theta}_i) - \hat{U}_i(\underline{\theta}_i)] - \sum_{i \in N} \int_{J_i^{-1}(r^*)}^{\bar{\theta}_i} (J_i(\theta_i) - r^*)(\tilde{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i$$

$$= \sum_{i \in N} \int_{\underline{\theta}_i}^{J_i^{-1}(r^*)} (J_i(\theta_i) - r^*)(\tilde{Q}_i(\theta_i) - Q_i^*(\theta_i)) f_i(\theta_i) d\theta_i.$$

In order not to contradict Step 2, this inequality and the inequality (31) both must hold as equality, which in turn implies that the inequality (32) also must hold as equality. Then, (31) and (32) can hold as equality only if $\tilde{Q}_i(\theta_i) = Q_i^*(\theta_i), \forall \theta_i \leq J_i^{-1}(r^*)$, and $\tilde{q}(\theta) = q^*(\theta)$, for almost all $\theta \in \Theta^*$, which yields the desired result. ▌

STEP 4. *$\hat{M}$ is WCP.*

Proof. Consider any feasible manipulation $M = (q, t)$. We claim that $U_i^M(\underline{\theta}_i) = U_i^{\hat{M}}(\underline{\theta}_i), \forall i \in N$. Suppose not. By Step 1, we can find another feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ satisfying (30). Then, by Step 3,

$$U_i^{\tilde{M}}(\theta_i) = U_i^{\tilde{M}}(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} \tilde{Q}_i(a) da = U_i^{\hat{M}}(\underline{\theta}_i) + \int_{\underline{\theta}_i}^{\theta_i} Q_i^*(a) da = U_i^{\hat{M}}(\theta_i),$$

which contradicts the inequality in (30). Thus, it must be that $U_i^M(\underline{\theta}_i) = U_i^{\hat{M}}(\underline{\theta}_i), \forall i \in N$. Applying Step 3 again, we have $\tilde{Q}_i(\cdot) = Q_i^*(\cdot)$ for all $i \in N$, implying that M yields the same interim payoffs as $\hat{M}$ to the bidders, which means that $\hat{M}$ is WCP. ▌

Proof of Theorem 5. Suppose that a pair, r_0 and $\hat{q}(\cdot)$, solves $[C]$. Construct the transfer rule $\hat{t}(\cdot)$ by substituting r_0 and $\hat{q}(\cdot)$ into (29). We show that an auction rule $\hat{M} = (\hat{q}, \hat{t})$ implements the solution outcome of $[C]$ and is WCP. To ensure that $\hat{M}$ implements the solution of $[C]$ absent collusion, we only need to check (IR) constraint. For this, note that due to the fact $\sum_{i \in N} \hat{t}_i(\theta) = r_0 \sum_{i \in N} \hat{q}_i(\theta)$ and (IC_1^*), we have

$$\sum_{i \in N} U_i^{\hat{M}}(\underline{\theta}_i) = \mathbb{E}\left[\sum_{i \in N} J_i(\theta_i)\hat{q}_i(\theta) - r_0 \sum_{i \in N} \hat{q}_i(\theta)\right] \geq 0.$$

Thus, one can choose ρ_i's so that each bidder's (IR) constraint is satisfied.

We next prove that $\hat{M}$ is WCP. Consider any feasible manipulation $\tilde{M} = (\tilde{q}, \tilde{t})$ of $\hat{M}$. Then, we have

$$\begin{aligned} \sum_{i \in N} U_i^{\tilde{M}}(\bar{\theta}) &= \mathbb{E}\left[\sum_{i \in N} K(\theta_i)\tilde{q}_i(\theta) - \sum_{i \in N} \tilde{t}_i(\theta)\right] \\ &= \mathbb{E}\left[\sum_{i \in N} (K(\theta_i) - r_0)\tilde{q}_i(\theta)]\right] \end{aligned}$$

$$\leq \ \mathbb{E}\left[\sum_{i\in N}(K(\theta_i)-r_0)\hat{q}_i(\theta)\right]=\sum_{i\in N}U_i^{\hat{M}}(\bar{\theta}),$$

where the inequality follows from the definition of $\hat{q}(\cdot)$ and becomes strict unless $\tilde{q}(\cdot)=\hat{q}(\cdot)$. Thus, $(IR_N^{\hat{M}})$ requires $\tilde{q}(\cdot)=\hat{q}(\cdot)$, so

$$\mathbb{E}\left[\sum_{i\in N}\tilde{t}_i(\theta)\right]=\mathbb{E}\left[r_0\sum_{i\in N}\tilde{q}_i(\theta)\right]=\mathbb{E}\left[r_0\sum_{i\in N}\hat{q}_i(\theta)\right]=\mathbb{E}\left[\sum_{i\in N}\hat{t}_i(\theta)\right],$$

which, along with $(IR_N^{\hat{M}})$, implies that $U_i^{\tilde{M}}(\cdot)=U_i^{\hat{M}}(\cdot)$, $\forall i$. We thus conclude that $\hat{M}$ is WCP.

We now show that the solution of $[C]$ is characterized by (9), (10), and (11), assuming that Condition (SB) does not hold. We first show that the solution involves the allocation rule of the form described in (11), whatever the value of r_0 is. Let λ_R and λ_K denote the Lagrangian (nonnegative) multipliers for the constraints (IC_1^*) and (K), respectively. Then, the Lagrangian for the problem $[C]$ is written as

$$\mathbb{E}\left[\sum_{i\in N}\Big(r+\lambda_R(J(\theta_i)-r)+\lambda_K\min\{K(\theta_i)-r,0\}\Big)q_i(\theta)\right].$$

Since the maximand is symmetric across bidders and linearly increasing with each q_i, the optimal allocation should follow the efficient cutoff rule: Namely, there exists a threshold value $\bar{\theta}$ such that the rule allocates the object to a bidder whose type is highest and above $\bar{\theta}$. Next observe the constraint (K) must be binding at the solution; or else, the solution corresponds to the second-best outcome. This yields a contradiction, since the solution is WCP implementable and the second-best outcome cannot be WCP implementable without Condition (SB). Therefore, (K) is binding, from which it follows that $\bar{\theta}=K^{-1}(r_0)$.

The optimal sale price r_0 depends on whether (IC_1^*) is binding or not. If (IC_1^*) is not binding, then given the efficient cutoff rule as in (11), r_0 must satisfy (9). Meanwhile, (IC_1^*) is slack only if

$$\mathbb{E}[J(\theta_N^{(1)})|\theta_N^{(1)}>K^{-1}(r_0)]>r_0.$$

If this inequality does not hold at the level solving (9), then (IC_1^*) must be binding, so r_0 must satisfy (10). ▌

Proof of Theorem 6. Recall first that the original auction M^* is the second-price auction. Since the auction A has the same allocation/payment rule as M^* for bidders outside C, it is weakly dominant for them to participate in A and report their true types. Fix any Bayesian Nash equilibrium in $\mathcal{E}_A$. Letting $\tilde{M}=(\tilde{q},\tilde{t})$ denote the mechanism resulting from the equilibrium play of bidders, we show that $\tilde{M}$ must yield V^* to the seller. First of all, $\tilde{M}$ must satisfy (IC^*) since in the equilibrium, each bidder is forming a correct belief about what types propose or accept/reject a collusive side contract, and thereafter playing sequentially rational strategy.[19] We now establish that both $\hat{M}$ and $\tilde{M}$ must yield the same interim payoffs for all collusive bidders. Let us first consider case (a). Then, since (12) holds strictly for all θ',

$$T_i^*(\bar{\theta}) > \mathbb{E}\left[H_C(\theta^{(1)}_{N\backslash C})\right]. \tag{33}$$

Define

$$\breve{\Theta}_C := \{\theta_C \in \Theta_C \,|\, \exists i \in C \text{ with type } \theta_i \text{ that announces } r_z\}.^{20}$$

Then, for all $\theta_C \in \Theta_C$,

$$\begin{aligned}
\mathbb{E}\left[\sum_{i\in C}\tilde{t}_i(\theta)\right] &= \mathbb{E}\left[H_C(\theta^{(1)}_{N\backslash C})\left(\sum_{i\in C}\tilde{q}_i(\theta)\right)1_{\{\theta_C\notin\breve{\Theta}_C\}} + T_i^*(\bar{\theta})1_{\{\theta_C\in\breve{\Theta}_C\}}\right]\\
&\geq \mathbb{E}\left[H_C(\theta^{(1)}_{N\backslash C})\left(\sum_{i\in C}\tilde{q}_i(\theta)\right)1_{\{\theta_C\notin\breve{\Theta}_C\}} + H_C(\theta^{(1)}_{N\backslash C})1_{\{\theta_C\in\breve{\Theta}_C\}}\right]\\
&\geq \mathbb{E}\left[H_C(\theta^{(1)}_{N\backslash C})\sum_{i\in C}\tilde{q}_i(\theta)\right],
\end{aligned} \tag{34}$$

where the first inequality follows from (33). Thus,

$$\begin{aligned}
\alpha_C(\sum_{i\in C}U_i^{\tilde{M}}(\bar{\theta})) + (1-\alpha_C)(\sum_{i\in C}U_i^{\tilde{M}}(\underline{\theta})) &= \mathbb{E}\left[\sum_{i\in C}H_C(\theta_i)\tilde{q}_i(\theta) - \sum_{i\in C}\tilde{t}_i(\theta)\right]\\
&\leq \mathbb{E}\left[\sum_{i\in C}[H_C(\theta_i) - H_C(\theta^{(1)}_{N\backslash C})]\tilde{q}_i(\theta)\right]\\
&\leq \mathbb{E}\left[\sum_{i\in C}[H_C(\theta_i) - H_C(\theta^{(1)}_{N\backslash C})]q_i^*(\theta)\right]\\
&= \alpha_C(\sum_{i\in C}U_i^{\hat{M}}(\bar{\theta})) + (1-\alpha_C)(\sum_{i\in C}U_i^{\hat{M}}(\underline{\theta})),
\end{aligned}$$

where the first and last equalities follow from Lemma 0, the first inequality follows from (34), and the second follows from the definition of $q^*(\cdot)$. Indeed, both inequalities must hold as

[19] Since $\theta_\emptyset \in B_i$ for each $i \in C$, every type of bidder i can secure at least zero (or individual rational) payoff whenever participating in A so that $\tilde{M}$ must satisfy (IR) condition.

[20] A mixed strategy which randomizes between r_z and some other messages can be accommodated without changing the subsequent result.

equality since (i) $U_i^{\tilde{M}}(\underline{\theta}) \geq 0 = U_i^{\hat{M}}(\underline{\theta})$ and (ii) $U_i^{\tilde{M}}(\bar{\theta}) \geq U_i^{\hat{M}}(\bar{\theta})$ for all $i \in C$. First, (i) is immediate from the fact that $\tilde{M}$ satisfies (IC^*). To show (ii), suppose to the contrary that $U_i^{\tilde{M}}(\bar{\theta}) < U_i^{\hat{M}}(\bar{\theta})$ for some $i \in C$. Then, bidder i has a profitable deviation to announce r_z with a sufficiently high z, since it will yield him an (interim) payoff arbitrarily close to $\bar{\theta} - T_i^*(\bar{\theta}) = U_i^{\hat{M}}(\bar{\theta})$, a contradiction. Now, both inequalities hold with equalities only if $\breve{\Theta}_C$ is a measure zero set and $\tilde{q}_i(\cdot) = q_i^*(\cdot)$, which implies that the interim payoffs in $\hat{M}$ and $\tilde{M}$ can only differ by constants. That (i) and (ii) hold with equalities in turn implies that those constants have to be zero. Consequently, $\hat{M}$ and $\tilde{M}$ must yield the same interim payoffs for all parties, which implies that $\tilde{M}$ must yields the seller her second-best payoff V^*.

The proof is similar for the case (b), upon two observations. First, adding the message r_z does not give the cartel any new opportunity to manipulate $\hat{M}$ since announcing r_z results in the same outcome as each collusive bidder announcing θ^r. Second, since (13) holds, the highest type of any bidder $i \in C$ can announce r_z (with sufficiently large z) to obtain at least its noncollusive payoff $U_i^{\hat{M}}(\bar{\theta})$.

Last, we prove for the case (a) that $\mathcal{E}_A$ is non-empty. (A similar proof follows for the case (b).) To this end, we show that there exists a weak perfect Bayesian, and thus Bayesian Nash, equilibrium in which each cartel member proposes no side contract. If no one proposes a side contract, then each collusive bidder i with type θ_i plays $\hat{M}$ and obtains his equilibrium payoff $U_i^{\hat{M}}(\theta_i)$. If a side contract is proposed, then each collusive bidder i responds as follows: "Report $\theta_\emptyset$ if $\theta_i < T_i^*(\overline{\theta})$ or else report r_{z_i} for some integer $z_i > 1$." This response is supported by the out-of-equilibrium belief of bidder i that each bidder $j \neq i$ in C reports $r_{z'}$ for some $z' < z_i$ if $\theta_j > T_j^*(\overline{\theta})$, and $\theta_\emptyset$ otherwise.

We now show that this strategy profile constitutes a weak perfect Bayesian equilibrium. First of all, a deviation by some collusive bidder or third party to a side contract will trigger the response as above and yield each collusive bidder i with θ_i at most $\max\{\theta_i - T_i^*(\overline{\theta}), 0\}$, which is no greater than $U_i^{\hat{M}}(\theta_i)$, his equilibrium payoff. Second, once a side contract has been proposed (out of equilibrium), it is optimal for a collusive bidder i with type $\theta_i > T_i^*(\overline{\theta})$ to report r_{z_i} and obtain $\theta_i - T_i^*(\overline{\theta}) > 0$, given his belief that every other collusive bidder will report either $r_{z'}$ or $\theta_\emptyset$. Also, bidder i with $\theta_i < T_i^*(\overline{\theta})$ optimally reports $\theta_\emptyset$ to obtain zero payoff since (i) reporting some r_z instead is clearly suboptimal and (ii) reporting some type from Θ_i yields either zero payoff (in case some other collusive bidder reports r_z) or at most $\theta_i - \delta_C(\overline{\theta}) < 0$ (in case every other collusive bidder reports $\theta_\emptyset$), and thus is suboptimal too. ▌

References

[1] Abdulkadiroğlu, A. and Chung, K.-S. (2003): "Auction Design with Tacit Collusion," Working Paper, Columbia and Northwestern Universities.

[2] Aoyagi, M. (2003): "Bid Rotation and Collusion in Repeated Auctions," *Journal of Economic Theory*, 112, 79-105.

[3] Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and Price Rigidity," *Review of Economic Studies*, 71, 317-349.

[4] Baldwin, L., R. C. Marshall, and J. F. Richard (1997): "Bidder Collusion in Forest Service Timber Sales," *Journal of Political Economy*, 105, 657-699.

[5] Blume, A. and P. Heidhues (2002): "Modeling Tacit Collusion in Auctions," mimeo., University of Pittsburgh.

[6] Celik, G. (2004): "Mechanism Design with Collusive Supervision," Unpublished Manuscript, University of British Columbia.

[7] Che, Y.-K. and J. Kim (2007): "Robustly Collusion-Proof Implementation," *Econometrica*, 74, 1063-1107.

[8] Crawford, V. (1985): "Efficient and Durable Decision Rules: A Reformulation," *Econometrica*, 53, 817-835.

[9] Dequiedt, V. (2007): "Efficient Collusion and Optimal Auctions," *Journal of Economic Theory*, 136, 302-323.

[10] Eso, P. and J. Schummer (2003): "Bribing and Signaling in the Second-Price Auction," *Games and Economic Behavior*, 47, 299-324.

[11] Faure-Grimaud, A., J.-J. Laffont, and D. Martimort (2003): "Collusion, Delegation and Supervision with Soft Information," *Review of Economic Studies*, 70, 253-279.

[12] Graham, D. A. and R. C. Marshall (1987): "Collusive Bidder Behavior at Single-Object Second-Price and English Auctions," *Journal of Political Economy*, 95, 1217-1239.

[13] Holmstrom, B. and R. Myerson (1983): "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51, 1799-1819.

[14] Jeon, D. S. and D. Menicucci (2005): "Optimal Second-Degree Price Discrimination and Arbitrage: On the Role of Asymmetric Information Among Buyers," *Rand Journal of Economics*, 36, 337-360.

[15] Laffont, J.-J. and D. Martimort (1997): "Collusion under Asymmetric Information," *Econometrica*, 65, 875-911.

[16] Laffont, J.-J. and D. Martimort (1998): "Collusion and Delegation," *Rand Journal of Economics*, 29, 280-305.

[17] Laffont, J.-J. and D. Martimort (2000): "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309-342.

[18] Mailath, G. and P. Zemsky (1991): "Collusion in Second Price Auctions with Heterogeneous Bidders," *Games and Economic Behavior*, 3, 467-486.

[19] Marshall, R. C. and L. M. Marx (2007): "Bidder Collusion," *Journal of Economic Theory*, 133, 374-402.

[20] McAfee, R. P. and J. McMillan (1992): "Bidding Rings," *American Economic Review*, 82, 579-799.

[21] Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58-73.

[22] Pavlov, G. (2006): "Colluding on Participation Decisions," Unpublished Manuscript, Boston University.

[23] Pesendorfer, M. (2000): "A Study of Collusion in First-Price Auctions," *Review of Economic Studies*, 67, 381-411.

[24] Porter, R. and D. Zona (1993): "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101, 518-538.

[25] Porter, R. and D. Zona (1999): "Ohio School Milk Markets: An Analysis of Bidding," *Rand Journal of Economics*, 30, 263-288.

[26] Quesada, L. (2004): "Modeling Collusion as an Informed Principal Problem," Unpublished Manuscript, University of Wisconsin.

[27] Skrzypacz, A. and H. Hopenhayn (2004): "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory*, 114, 153-169.