# THE LIMITS OF AUCTIONS UNDER *EX-ANTE* COLLUSION

RAVI JAGADEESAN, ILYA SEGAL, ALEX TORDJMAN, AND KAREN WU

ABSTRACT. We study revenue-maximizing auction design when bidders can collude *ex ante*—i.e., before each bidder learns his value or decides whether to participate in the auction. We show that bidders' *ex-post* private information about values does not preclude them from colluding efficiently, and as a result, the seller cannot guarantee a higher expected revenue than from posting an optimal price and rationing the good. However, a posted-price mechanism leaves the seller vulnerable to *inefficient* collusion. We construct mechanisms that guarantee the non-collusive expected revenue of an optimal posted price against a broad range of forms of collusion—including *ex-ante* collusion by rings comprised of subsets of bidders, rings that can commit to punish bidders who refuse to join, and rings of which some bidders may be unaware. We show that achieving such a guarantee requires protecting bidders' *ex-ante* expected utilities against all such forms of collusion—a property that we term *strong collusion-robustness*. With symmetric bidders, our strongly collusion-robust mechanism simply combines a posted price with a first-price knockout auction conducted by the seller on behalf of the bidders.

*Date*: August 27, 2026.

Jagadeesan: Department of Economics, Stanford University. Email: ravi.jagadeesan@gmail.com.
Segal: Department of Economics, Stanford University. Email: ilya.segal@stanford.edu.
Tordjman: Department of Economics, Stanford University. Email: tordjman@stanford.edu.
Wu: Institute for Microeconomics, University of Bonn. Email: kawu@uchicago.edu.

This paper is a revision of the *Econometrica* submission "The Limits of Auctions under Ex-Ante Collusion" by Jagadeesan, Segal, and Tordjman (SSRN Working Paper 5231187). Contemporaneous work by Wu, also submitted to *Econometrica* as "Collusion Robust Auction Design: Posted Price and a Knockout Auction" (SSRN Working Paper 6010915, dated September 16, 2025), developed a version of the analysis in Sections 2 and 3. Sections 4–7 develop analysis not present in Wu's submission. Section 8 develops new results related to a design problem considered in Wu's submission. Wu's contributions to the present paper are Example 6 and part of the proof of Proposition 9 in Section 8.

We thank Yeon-Koo Che, Jason Hartline, Navin Kartik, Scott Kominers, Shengwu Li, Elliot Lipnowski, Paul Milgrom, Mike Ostrovsky, Alessandro Pavan, Alex Teytelboym, June Zhu, several seminar participants, the anonymous referees, and the editor, Marina Halac, for helpful comments.

## 1. Introduction

Standard auctions are vulnerable to collusion. For example, in second-price and ascending auctions, members of a bidding ring may run a knockout auction to determine who bids in the final auction. This possibility lowers revenue by reducing competition among bidders—making sale prices drop to the reserve if all bidders collude.[1] A large theoretical literature, initiated by Laffont and Martimort (1997) and Che and Kim (2006), has analyzed how to design auctions that are robust to collusion schemes in which rings entice bidders who know their own private values to collude.

This paper examines the revenue limits of all auction formats when bidding rings are formed *ex ante*—i.e., before bidders learn their values or decide whether to participate in the auction (but after the seller announces the auction format). One setting in which *ex-ante* collusion arises in practice is collusion by bidding rings that operate for the long term across many auctions with similar formats (Pesendorfer, 2000; Asker, 2010), which are formed before any bidder learns his value for, or decides whether to participate in, a particular auction.[2] *Ex-ante* collusion may also arise if bidders participate in a large number of simultaneous auctions with similar formats.

We consider the revenue-maximization problem of a seller with a good to sell to one of several bidders, who have independent private values and may decide to collude *ex ante*. If a bidding ring is formed, it can dictate the messages that its members send to the seller, and then reallocate the good and implement (budget-balanced) side payments amongst its members. However, once bidders learn their values, the ring is subject to asymmetric information about values analogous to the asymmetric information faced by the seller (Laffont and Martimort, 1997; Che and Kim, 2006).

Our results show that *ex-ante* collusion sharply limits the expected revenue that any selling mechanism can guarantee: the seller's optimal expected-revenue guarantee under collusion is the non-collusive expected revenue of an optimal posted price. Thus, the seller is unable to extract any expected revenue from competition between bidders.

We develop our analysis in three parts. First, we show that the grand coalition of bidders is able to collude efficiently despite its members' private information, and if that happens, then the seller's optimal expected revenue is achieved by posting an optimal price and rationing the good among the bidders willing to pay it. Second, we show that simply posting an optimal price leaves the seller vulnerable to *inefficient* collusion that reduces the probability of sale. Last, we construct selling mechanisms

[1] See Graham and Marshall (1987), McAfee and McMillan (1987), and Mailath and Zemsky (1991).
[2] Jagadeesan et al. (2026) formalizes a connection between our model of *ex-ante* collusion and self-enforcing long-term collusion via a version of the Folk Theorem for infinitely repeated games.

that guarantee the (non-collusive) expected revenue of an optimal posted price against a broad range of forms of collusion—including *ex-ante* collusion by rings comprised of subsets of bidders, rings that can commit to punish bidders who refuse to join, and rings of which some bidders may be unaware. When bidders are symmetric, our robust mechanism simply combines a posted price with a first-price knockout auction that the seller conducts on behalf of the bidders.

For the first part, we suppose that the grand coalition of bidders colludes in a way that is *ex-ante* Pareto-optimal subject to (i) the ring's asymmetric information constraints, and (ii) being an *ex-ante* Pareto improvement over non-collusive play. Our first main result shows that the expected revenue delivered by any selling mechanism under any efficient collusion scheme is at most the (non-collusive) expected revenue of an optimal posted price. Moreover, any posted price achieves its non-collusive expected revenue under all efficient collusion schemes. Thus, a posted price is optimal.

To prove these results, we first show that for any selling mechanism, the grand coalition, acting as a bidding ring, can overcome its asymmetric information and *ex-ante* participation constraints to collude in an efficient way. Indeed, it can do so by using budget-balanced d'Aspremont and Gérard-Varet (1979) side payments to incentivize truthful revelation of values, plus lump-sum side payments to make forming the ring an *ex-ante* Pareto improvement for the bidders.[3] Hence, under any efficient collusion scheme, the grand coalition effectively acts as a single buyer that values the good at the highest of the bidders' values. The optimality of a posted price follows because a posted price is optimal for selling to a single buyer, even among randomized selling mechanisms (Myerson, 1981; Riley and Zeckhauser, 1983).

In practice, if the seller thinks there is a (small) chance that collusion will not occur, then the seller may prefer to run a standard auction with the optimal posted price as a reserve—as proposed by Graham and Marshall (1987). With symmetric bidders, the optimal posted price is increasing in the number of bidders and greater than the non-collusive optimal reserve (McAfee and McMillan, 1987). From this perspective, our results imply that the best that the seller can do to counter efficient collusion using any selling mechanism is to simply raise the reserve.[4]

Next, we turn to the case in which bidders may collude inefficiently. We show that posting an optimal price—or running a standard auction using the optimal posted price as a reserve—is then insufficient to guarantee the non-collusive expected revenue

[3]Our construction builds on Mailath and Zemsky's (1991) construction of efficient collusion schemes for second-price auctions using d'Aspremont and Gérard-Varet (1979) side payments.
[4]This classic approach to countering collusion was documented by Cassady (1967, pp. 228–230).

of an optimal posted price. Intuitively, under each of these mechanisms, there are large gains from collusion: under a posted price, from allocating the good efficiently rather than randomly; and under a standard auction, from colluding to lower payments to the seller. So, a collusion scheme in which bidders capture some of these gains but sometimes inefficiently abstain from buying the good can be *ex-ante* Pareto-improving for the bidders while harming the seller. A practical form of inefficient collusion arises in ad auctions, where bidders often delegate their bidding to a demand-side platform.[5] We show that both under a posted price and under a standard auction with the optimal posted price as a reserve, the platform maximizing its own expected revenue (subject to interim participation constraints) leads to inefficient implicit collusion that involves double marginalization. This lowers the seller's expected revenue below the non-collusive expected revenue of an optimal posted price.

Rather than taking a stance on what forms of collusion are reasonable, we show that it is possible to guarantee the non-collusive revenue of an optimal posted price across a broad class of collusion schemes. In addition to collusion that is *ex-ante* Pareto-improving for the bidders, we ensure robustness to inefficient collusion that may harm some of the bidders—such as collusion by a ring smaller than the grand coalition that may harm non-colluding bidders, and collusion by rings that can commit to punish bidders who refuse to join the ring. We do restrict collusion to be *individually rational* for colluding bidders, in the sense that no such bidder may be left worse off than he would be if the other ring members punished him by sending minmax messages.[6] Non-colluding bidders who are aware of collusion play best responses to it, so also receive at least their (*ex-ante*) minmax utilities. As for bidders who are unaware that collusion is taking place, they play the mechanism as though no one were colluding and so may receive less than their (*ex-ante*) minmax utilities. As a special case, this class of collusion schemes includes collusion by rings that are formed after bidders learn their values, where a type of a bidder can have arbitrary off-path beliefs about how the mechanism would be played if he rejected a proposed collusion scheme (as in Laffont and Martimort (1997, Section 8)).

One idea behind the construction is to use the logic of the Collusion-Proofness Principle (Laffont and Martimort, 1997): to deter collusion, the seller implements efficient collusion on behalf of the bidders. In particular, if the good is sold, it is allocated to the highest bidder. However, a selling mechanism constructed in

[5]The long-term collusion in stamp auctions documented by Asker (2010) was also inefficient.
[6]See also, e.g., Quesada (2005), Dequiedt (2007), and Csóka et al. (2026).

this way may still leave the seller vulnerable to collusion that is not *ex-ante* Pareto-improving for the bidders.[7] For example, we show that with three bidders, in a selling mechanism that combines a posted price with Mailath and Zemsky's (1991) efficient collusion scheme (which is based on an all-pay knockout auction or d'Aspremont and Gérard-Varet (1979) side payments), the seller is vulnerable to collusion by smaller rings consisting of pairs of bidders that harms other bidders.

To construct selling mechanisms that are robust to collusion that may harm some bidders, we combine a posted price with particular alternative efficient collusion schemes. Our first mechanism, which is robust to collusion in the case of symmetric bidders, simply combines the posted price with a first-price knockout auction conducted by the seller whose revenue is remitted back to the bidders.[8] Our second mechanism is more complicated, but robust to collusion beyond the symmetric case. It combines a posted price with an efficient collusion scheme whose side payments are based on a mechanism proposed by Safronov (2018), adapting them to protect the seller from collusion about participation in the selling mechanism.

These mechanisms guarantee the optimal non-collusive posted-price expected revenue under the broad class of collusion schemes that we consider, which generalizes the kinds of collusion considered in the previous literature. Since our first result shows that higher expected revenue cannot be obtained even if collusion is known to be efficient, incentive compatible, and *ex-ante* Pareto-improving for the bidders, our mechanisms are max-min optimal for a range of possible models of *ex-ante* collusion.

To develop our robust mechanisms, we show that any mechanism that guarantees the optimal posted-price expected revenue against the forms of collusion that we consider must also protect bidders' *ex-ante* expected utilities from all such forms of collusion—a property that we term *strong collusion-robustness*. Strong collusion-robustness, in turn, ensures that revenue is robust to the forms of collusion that we consider, even *ex post*. A selection of a new Bayes–Nash equilibrium is a particular form of individually rational collusion, so the outcomes of strongly collusion-robust mechanisms are robust to equilibrium selection—a version of full implementation.

We then characterize which mechanisms are strongly collusion-robust and guarantee the optimal posted-price expected revenue by means of four conditions that are

[7]Laffont and Martimort (1997, 2000) derived Collusion-Proofness Principles for their "weak collusion-proofness" concept, which implicitly restricts attention to collusion that cannot harm any bidders.
[8]McAfee and McMillan (1992) showed that a first-price knockout auction with equal revenue sharing among the losing bidders is an efficient collusion scheme for symmetric bidders, but did not study whether the seller implementing it on behalf of the bidders would protect the seller from collusion—in particular, from inefficient collusion that may harm some of the bidders.

sufficient and generically necessary. The first property is that under non-collusive play, the good be allocated to the highest bidder if that bidder's value is above the optimal posted price. The second is that under non-collusive play, the bidders be charged the optimal posted price in aggregate whenever the good is sold. The third is that the bidders be charged at least as much in aggregate out of equilibrium (e.g., when some bidders do not participate). These three properties together ensure robustness to collusion that is *ex-ante* Pareto-improving for the bidders, by giving the bidders an outcome that is efficient for them (accounting for the posted price).

However, ensuring robustness to broader kinds of collusion that may harm some of the bidders (as described above) requires a fourth property: that each bidder can guarantee his non-collusive *ex-ante* expected utility by playing the mechanism as if no one were colluding, regardless of how other bidders actually behave. This *guaranteed-utility property* ensures that all bidders must receive at least their non-collusive payoffs under all the forms of collusion that we consider—implying that we only need to concern ourselves with collusion that is *ex-ante* Pareto-improving for the bidders.

Our mechanisms satisfy the first three properties because they implement efficient collusion on behalf of the bidders. A rough intuition for why the first-price knockout mechanism satisfies the guaranteed-utility property in the symmetric case is that it protects each bidder from other bidders' behavior in two ways. First, the revenue-sharing element compensates him for losing the auction, which protects him from others bidding high to win often. Second, the first-price element ensures that other bidders can only affect him through their highest bid, which protects him from others colluding to lower their knockout-auction payments.[9] As for our mechanism for the general case, Safronov's (2018) side payments ensure that the guaranteed-utility property holds if all bidders participate. Our adaptation ensures that the guaranteed-utility property holds even when some bidders do not participate by appropriately defining participants' outcomes if some bidders do not participate, as well as adding particular lump-sum side payments.

Optimal strongly collusion-robust mechanisms make the seller compromise on expected revenue even if collusion does not occur: they yield the optimal posted-price expected revenue in all Bayes–Nash equilibria. We show that this is inevitable: no mechanism that guarantees the optimal posted-price expected revenue under *ex-ante* collusion sustained with a threat of playing a Bayes–Nash equilibrium if any bidder

[9]While Robinson (1985) argued that the standard first-price auction is relatively resistant to collusion, without revenue sharing, it is still susceptible to collusion.

refuses to collude can yield a higher revenue in any Bayes–Nash equilibrium. Intuitively, any higher-revenue Bayes–Nash equilibrium can be used as a punishment to sustain a lower-expected-revenue collusion scheme. Moreover, if the expected-revenue guarantee must hold over all individually rational collusion, then the mechanism must yield the same revenue under *all* individually rational collusion schemes. In these senses, the mechanisms we construct are undominated.

One unusual property of our robust mechanisms is that they involve payments to losing bidders. We show that this is also inevitable: unless selling to a single bidder is optimal, guaranteeing the optimal posted-price expected revenue under even collusion that is *ex-ante* Pareto-improving for the bidders requires sometimes making payments to a losing bidder in equilibrium. We also provide an example of a mechanism with the optimal expected-revenue guarantee subject to the constraint of never paying bidders.

*Related Literature.* Our analysis builds on seminal analyses of collusion-proof mechanisms by Laffont and Martimort (1997, 2000) and Che and Kim (2006). Those papers considered collusion where rings are formed *after* each bidder learns his value and decides whether to participate in the auction. Che and Kim (2006) suggested a mechanism that guarantees the non-collusive optimal expected revenue against any collective misbehavior that bidders might engage in *after* agreeing to participate in the mechanism. Subsequent work by Pavlov (2008) and Che and Kim (2009) showed that collusion does sometimes limit expected revenue if bidders can form a ring at the interim stage—i.e., *before* any bidder decides whether to participate in the auction, but *after* each bidder learns his value. Instead, we suppose that rings are formed *ex ante*—i.e., *before* each bidder learns his value or decides whether to participate in the auction. We show that collusion then limits the expected revenue that can be guaranteed to the (non-collusive) expected revenue of an optimal posted price. The optimal mechanisms that we construct are robust to a broad class of *ex-ante* collusion schemes.

Our conclusions are closer to those of Dequiedt (2007), who studied optimal auctions in the presence of collusion with two bidders and two types. He considered a stronger form of collusion at the interim stage in which rings can not only collude on whether to participate in the auction, but can also commit to an arbitrary punishment of defectors who refuse to join the ring.[10] Dequiedt (2007) showed that the seller cannot improve upon an optimal posted price. However, unlike our result, it is unclear if his result generalizes to more bidders or types. Moreover, we construct mechanisms that are robust to the type of collusion considered by Dequiedt (2007).

[10] Quesada (2005) studied auctions for procuring complements under a related form of collusion.

Our construction of collusion-robust mechanisms extends the uses of the guaranteed-utility property by Safronov (2018), Csóka et al. (2026), and Csóka et al. (2025) along several dimensions.[11] First, we show that the guaranteed-utility property ensures robustness to stronger forms of collusion than those considered in this literature—including collusion that may harm some bidders and of which some bidders may be unaware. Second, while the literature focused on implementing efficiency, we show that the guaranteed-utility property is sufficient to make revenue robust to collusion, and necessary to provide an optimal expected-revenue guarantee. Third, we obtain robustness to collusion that can reallocate the good. Fourth, we ensure robustness to collusion on participation and construct mechanisms that satisfy *ex-post* participation constraints, rather than merely *ex-ante* or interim participation constraints.[12] Fifth, we show that in the case of symmetric bidders, a much simpler mechanism, namely one based on a first-price knockout auction, also has the guaranteed-utility property and is therefore robust to collusion. Unlike the other proposed mechanisms that satisfy a guaranteed-utility property (Safronov, 2018; Csóka et al., 2026), this first-price-based mechanism relies on the prior only to set the posted price—the mechanism is indirect, and bidders use the prior to calculate their equilibrium bids.

*Paper Outline.* Section 2 introduces our model. Section 3 shows that a posted price is optimal under efficient collusion. Section 4 illustrates difficulties in guaranteeing the posted-price expected revenue under more general collusion. Section 5 formalizes robustness to collusion by rings of bidders that can commit to punish deviators, and introduces and characterizes our strong-collusion-robustness desideratum. Section 6 constructs our strongly collusion-robust mechanisms. Section 7 shows they are undominated. Section 8 explains the role of payments to losing bidders. Section 9 discusses further implications and concludes. The appendix presents omitted proofs supporting the three theorems. The online appendix presents the remaining omitted proofs, additional examples, and additional discussion of our model of collusion.

## 2. Model

A seller has one unit of a good to sell to one of a finite set $I$ of $N \geq 2$ bidders $i$. Thus, the space of feasible allocations of the good is given by $\mathcal{Q} = \left\{ q \in \mathbb{R}_+^I \,\middle|\, \sum_{i \in I} q_i \leq 1 \right\}$. The

[11] Segal and Whinston (2012, Proposition 1), Safronov (2018, Theorems 2 and 3), Csóka et al. (2026, Section 5), and Csóka et al. (2025, Section 3.4) formalized weaker senses in which the guaranteed-utility property ensures robustness to equilibrium selection and collusion.

[12] Csóka et al. (2026, Section 7.1) also constructed a mechanism that ensures robustness to collusion on participation, but their mechanism does not satisfy *ex-post* participation constraints.

bidders have independent, private values for the good. Bidder $i$'s value $\theta_i$ is drawn from a continuous distribution $F_i$ supported on a compact interval $\Theta_i = [0, \overline{\theta}_i]$.

The game proceeds as follows. First, the seller announces a mechanism, which must include a "nonparticipation message" for each bidder. Second, an intermediary (or any bidder[13]) proposes a collusion scheme, whereby bidders can agree about how to manipulate their reports to the selling mechanism—including whether to send a nonparticipation message—as well as whether or how to reallocate the good and exchange side payments. Bidders commit to the collusion scheme before learning their values.

Formally, for each bidder $i$, the seller can set an arbitrary (Polish) message space $M_i$ that includes a *nonparticipation message* $NP_i$, which must guarantee the bidder an allocation and payment of 0. Thus, letting $M = \times_{i \in I} M_i$ denote the space of message profiles, the seller sets a (bounded, measurable) outcome function $(q, t) : M \to \mathcal{Q} \times \mathbb{R}^I$ such that $q_i(NP_i, m_{-i}) = t_i(NP_i, m_{-i}) = 0$ for all message profiles $m_{-i} \in M_{-i}$.[14] To make bidding rings' problems well-behaved, we impose the technical restriction that the space $\left\{ \left( \sum_{i \in I} q_i(m), \sum_{i \in I} t_i(m) \right) \middle| m \in M \right\}$ of obtainable pairs of aggregate allocations and aggregate transfers be closed, hence compact (Pavlov, 2008, Condition 2).[15] A *selling mechanism* then consists of message spaces $M_i$ for each bidder, an outcome function $(q, t) : M \to \mathcal{Q} \times \mathbb{R}^I$, and (measurable) strategies $\sigma_i : \Theta_i \to M_i$.[16,17]

For collusion schemes, by the Revelation Principle, it is without loss of generality to restrict attention to direct mechanisms (Laffont and Martimort, 1997). Rather than directly specifying the reallocation and side payments by a bidding ring, following Che and Kim (2006), we consider the allocations $\hat{q}(\theta)$ and transfers $\hat{t}(\theta)$ induced by the composite of the reallocation and side payments of the ring with the allocation and transfer of the message profile $\mu(\theta)$ submitted to the selling mechanism.

**Definition 1** (Laffont and Martimort, 1997; Che and Kim, 2006)**.** A *collusion scheme* for a selling mechanism $(M, q, t, \sigma)$ consists of a message function $\mu : \Theta \to \Delta(M)$ and

[13]Unlike in Laffont and Martimort (1997) and Che and Kim (2006), for *ex-ante* collusion, it does not matter whether an intermediary proposes the collusion scheme or one of the bidders proposes it, as no bidder has learned his value at the time at which the collusion scheme is proposed.

[14]Throughout, measurability conditions are with respect to Borel $\sigma$-algebras.

[15]Any outcome function can be approximated arbitrarily well by ones satisfying this restriction.

[16]The restriction to pure strategies is without loss of generality, as the message space can be expanded to include stochastic reports over pure messages.

[17]We specify strategies so a selling mechanism gives a well-defined non-collusive outside option to which bidders can compare a proposed collusion scheme when considering collusion that cannot harm any bidder. As we discuss in Section 5, the strategies also give the behavior of bidders who are unaware of collusion. For interpretation, one can suppose $\sigma$ is a Bayes–Nash equilibrium, but (except when stated otherwise) our results do not require this.

a direct mechanism[18] $(\hat{q}, \hat{t}) : \Theta \to \mathcal{Q} \times \mathbb{R}^I$ such that for all type profiles $\theta \in \Theta$, we have

$$\sum_{i \in I} \hat{q}_i(\theta) = \mathbb{E}_{\tilde{m} \sim \mu(\theta)}\Big[\sum_{i \in I} q_i(\tilde{m})\Big] \quad \text{(feasibility of reallocation of goods)} \tag{1}$$

$$\sum_{i \in I} \hat{t}_i(\theta) \geq \mathbb{E}_{\tilde{m} \sim \mu(\theta)}\Big[\sum_{i \in I} t_i(\tilde{m})\Big] \quad \text{(weak budget-balancedness of side payments).} \tag{2}$$

If bidders collude, they are bound to send the messages recommended by $\mu$ and follow through with the reallocation of the good and side payments, resulting in $(\hat{q}, \hat{t})$.[19] (1) requires that the aggregate allocation in the collusion scheme equal the aggregate allocation resulting from bidders sending messages according to $\mu$ to the selling mechanism—i.e., that the reallocation of the good be feasible.[20] (2) requires that the ring's side payments be weakly budget-balanced, which allows a third-party ringleader to extract money or the ring to burn money. (Laffont and Martimort (1997) and Che and Kim (2006) instead imposed strict budget-balancedness.)

Our analysis of efficient collusion places two additional restrictions on collusion. Imposing these constraints makes our first result on the limits of auctions stronger. On the other hand, the mechanisms we construct in Section 6 are robust to much more general forms of collusion that are introduced in Section 5.

First, we suppose that after bidders learn their values, they (like the seller) may be subject to asymmetric information about others' values (Laffont and Martimort, 1997; Che and Kim, 2006). Thus, bidders can affect the messages sent under collusion and the reallocations by manipulating their type reports to the bidding ring. Hence, we restrict attention to collusion schemes that are (Bayesian) incentive-compatible.

**Definition 2.** A collusion scheme $(\mu, \hat{q}, \hat{t})$ (for any selling mechanism) is *incentive-compatible (IC)* if for all bidders $i$ and $F_i$-almost all types $\theta_i \in \Theta_i$, for all reports $\theta_i' \in \Theta_i$:

$$\mathbb{E}_{\tilde{\theta}_{-i}}\big[\theta_i \hat{q}_i(\theta_i, \tilde{\theta}_{-i}) - \hat{t}_i(\theta_i, \tilde{\theta}_{-i})\big] \geq \mathbb{E}_{\tilde{\theta}_{-i}}\big[\theta_i \hat{q}_i(\theta_i', \tilde{\theta}_{-i}) - \hat{t}_i(\theta_i', \tilde{\theta}_{-i})\big].$$

Second, we restrict to collusion schemes that are (weakly) *ex-ante* Pareto-improving for the bidders, which is justified if any bidder can induce non-collusive play $\sigma$ by rejecting a proposed collusion scheme. If all bidders follow the prescribed strategies, then bidder $i$'s *ex-ante* expected utility is $V_i(M, q, t, \sigma) = \mathbb{E}_{\tilde{\theta}}\big[\tilde{\theta}_i q_i(\sigma(\tilde{\theta})) - t_i(\sigma(\tilde{\theta}))\big]$,

[18] Message functions are required to be measurable, where we equip $\Delta(M)$ with the weak-* topology. The allocation and payment rules of direct mechanisms are required to be bounded and measurable.

[19] In the interpretation of *ex-ante* collusion as a long-term collusion scheme for a sequence of auctions, bidders may be deterred from such deviations with the threat of punishments in future periods. See Jagadeesan et al. (2026) for a justification based on a version of the Folk Theorem for repeated games.

[20] For simplicity, following Che and Kim (2006), (1) and (2) impose feasibility in expectation over the colluders' messages $\tilde{m} \sim \mu(\theta)$. This is without loss since any reallocation satisfying these conditions can be implemented in a way that is feasible for all realizations of messages (see Online Appendix E).

where we write $\sigma(\theta) = (\sigma_i(\theta_i))_{i\in I} \in M$. On the other hand, each bidder's *ex-ante* expected utility from a collusion scheme $(\mu, \hat{q}, \hat{t})$ is $U_i(\hat{q}, \hat{t}) = \mathbb{E}_{\tilde{\theta}}\big[\tilde{\theta}_i \hat{q}_i(\tilde{\theta}) - \hat{t}_i(\tilde{\theta})\big]$.

**Definition 3.** A collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism $(M, q, t, \sigma)$ is *ex-ante Pareto-improving* if we have $U_i(\hat{q}, \hat{t}) \geq V_i(M, q, t, \sigma)$ for all bidders $i$.

## 3. Optimality of a Posted Price under Efficient Collusion

We first focus on collusion that is (*ex-ante*) Pareto-efficient subject to the two constraints from the previous section.

**Definition 4.** A collusion scheme for a selling mechanism is *ex-ante incentive-efficient* if it is IC and *ex-ante* Pareto-improving, and is *ex-ante* Pareto-optimal for the bidders over all such collusion schemes for the selling mechanism.

### 3.1. Efficient Collusion.

To analyze auction design under *ex-ante* incentive-efficient collusion, we first show that the grand coalition can overcome its IC and *ex-ante* Pareto-improvement constraints to achieve efficient collusive outcomes. This is possible even though we require collusive side payments to be budget-balanced *ex post*.

Formally, we say that a collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism is *efficient* if it is *ex-ante* Pareto-optimal over all collusion schemes for the selling mechanism. Such schemes exist due to our compactness assumption on the space of aggregate allocations and transfers. (The formal proof in the appendix ensures that the relevant measurability conditions are satisfied.)

**Lemma 1.** *For each selling mechanism, there exists an efficient collusion scheme.*

We show that every efficient collusion scheme can be made IC by modifying collusive side payments in a way that does not change the bidders' *ex-ante* expected utilities. Hence, the grand coalition can overcome its IC and *ex-ante* Pareto-improvement constraints and collude efficiently.

**Proposition 1.** *(a) Given any efficient collusion scheme $(\mu, \hat{q}, \hat{t})$ for any selling mechanism, there exists a payment rule $\hat{t}'$ such that $(\mu, \hat{q}, \hat{t}')$ is an IC collusion scheme that gives the same profile of bidders' ex-ante expected utilities.*

*(b) In particular, every ex-ante incentive-efficient collusion scheme is efficient.*

To prove Part (a) of the proposition, we extend arguments of Mailath and Zemsky (1991) from second-price auctions to general mechanisms. Given an efficient collusion scheme, we can construct payments that ensure IC while preserving *ex-post* budget

balancedness by using a version of d'Aspremont and Gérard-Varet's (1979) construction; adding lump-sum side payments allows us to leave bidders' *ex-ante* expected utilities unchanged. Therefore, there is no efficiency cost of being IC or being *ex-ante* Pareto-improving, which implies Part (b). The formal argument is in the appendix.

Collusion being arranged *ex ante* is crucial to Proposition 1. If bidders could only agree to collude after knowing their values, they could not generally overcome IC and interim Pareto-improvement constraints to obtain efficient collusive outcomes (Laffont and Martimort, 1997; Che and Kim, 2006, 2009; Pavlov, 2008). For example, under the mechanism suggested by Pavlov (2008), a constraint of providing the highest type at least his interim utility under $\sigma$ binds in the grand coalition's surplus maximization problem for interim collusion, which restricts the surplus achievable by IC collusion schemes and precludes efficient interim collusion. Conceptually, the role of collusion being arranged *ex ante* is analogous to the critical role of participation constraints being *ex-ante* for achieving efficient, budget-balanced implementation in general (compare Arrow's (1979) and d'Aspremont and Gérard-Varet's (1979) possibility results with Myerson and Satterthwaite's (1983) impossibility result).

In light of Proposition 1(b), rather than considering *ex-ante* incentive-efficient collusion schemes, it suffices to simply consider efficient collusion schemes.

3.2. **Optimality of a Posted Price.** Our main result on auction design under (*ex-ante* incentive-)efficient collusion shows the optimality of a posted-price mechanism.

A $p$-posted-price mechanism allocates the good randomly among those who are willing to pay $p$ (if any), charging $p$ to the bidder who receives the good. (Throughout the paper, we restrict attention to nonnegative prices $p \geq 0$.) Each bidder only needs to express whether he would like to buy the good (i.e., send participation message $P_i$) or not (i.e., send nonparticipation message $NP_i$). Formally, the *p-posted-price mechanism* is defined by binary message spaces $M_i = \{P_i, NP_i\}$, allocation rule

$$q_i(m) = \begin{cases} 0 & \text{if } m_i = NP_i \\ 1 \big/ |\{i \,|\, m_i = P_i\}| & \text{if } m_i = P_i \end{cases},$$

payment rule $t_i(m) = pq_i(m)$, and (dominant) strategies

$$\sigma_i(\theta_i) = \begin{cases} P_i & \text{if } \theta_i \geq p \\ NP_i & \text{if } \theta_i < p \end{cases}.$$

The *ex-post* revenue under the prescribed strategies is $r(p;\theta) = p\mathbf{1}(\max_{i\in I} \theta_i \geq p)$, and expected revenue under them is $R(p) = \mathbb{E}_{\tilde{\theta}}[r(p;\tilde{\theta})] = p \Pr_{\tilde{\theta}}[\max_{i\in I} \tilde{\theta}_i \geq p]$.

Let $p^\star \in \arg\max_{p\geq 0} R(p)$ be an optimal posted price. Our first main result shows that if bidders collude in any (*ex-ante* incentive-)efficient manner, then posting a price of $p^\star$ is optimal. Formally, we bound the expected revenue of any mechanism under any (*ex-ante* incentive-)efficient collusion scheme by $R(p^\star)$, and show that a $p^\star$-posted-price mechanism yields *ex-post* revenue $r(p^\star;\theta)$ almost surely under any (*ex-ante* incentive-)efficient collusion scheme.[21]

**Theorem 1.** *(a) For any selling mechanism, expected revenue under any efficient (and therefore any ex-ante incentive-efficient) collusion scheme is at most $R(p^\star)$.*

*(b) For the $p^\star$-posted-price mechanism, ex-post revenue under each efficient (and therefore each ex-ante incentive-efficient) collusion scheme is almost surely $r(p^\star;\theta)$.*

The intuition behind Theorem 1 is that under (*ex-ante* incentive-)efficient collusion, it is as if the seller faces a single buyer whose value is the highest of the bidders' values. The ability of the grand coalition to send nonparticipation messages ensures *ex-post* individual rationality (*ex-post* IR) for the coalition. In turn, in a single-buyer setting with an *ex-post* IR constraint, a posted price is optimal, and (with a continuous value distribution) almost surely fully implements expected-revenue maximization.

3.3. **Proof of Theorem 1.** We first derive some simple implications of efficient collusion. Since bidders' utilities are quasilinear, efficient collusion schemes must almost surely send messages that maximize bidders' joint surplus, which values the aggregate allocation at the highest bidder's value and accounts for the aggregate payment. Hence, it is as if the good is sold to a single buyer whose private value is the highest of the bidders' values (subject to the standard IC and *ex-post* IR constraints).

**Lemma 2.** *Let $(\mu,\hat{q},\hat{t})$ be an efficient collusion scheme for a selling mechanism $(M,q,t,\sigma)$.*

*(a) The message function $\mu$ must satisfy*

$$\mu(\theta) \in \arg\max_{\eta\in\Delta(M)} \left\{ \Big(\max_{i\in I}\theta_i\Big)\mathbb{E}_{\tilde{m}\sim\eta}\Big[\sum_{i\in I} q_i(\tilde{m})\Big] - \mathbb{E}_{\tilde{m}\sim\eta}\Big[\sum_{i\in I} t_i(\tilde{m})\Big]\right\} \quad \text{almost surely.} \tag{3}$$

*(b) The collusive aggregate allocation $\sum_{i\in I}\hat{q}_i(\theta)$ and the seller's ex-post revenue $\sum_{i\in I} t_i(m)$, where $m\sim\mu(\theta)$, define a mechanism for selling to a single buyer with type $\theta$ and value $\max_{i\in I}\theta_i$ that is almost surely IC and almost surely IR.*

*Proof.* As utility is transferable, efficiency of the collusion scheme implies that

$$\big(\mu(\theta),\hat{q}(\theta)\big) \in \arg\max_{(\eta,q)\in\Delta(M)\times\mathcal{Q}\,|\,\sum_{i\in I} q_i = \mathbb{E}_{\tilde{m}\sim\eta}[\sum_{i\in I} q_i(\tilde{m})]} \left\{\sum_{i\in I}\theta_i q_i - \mathbb{E}_{\tilde{m}\sim\eta}\Big[\sum_{i\in I} t_i(\tilde{m})\Big]\right\} \quad \text{a.s.}$$

[21] Note that seller's *ex-post* revenue $\sum_{i\in I} t_i(m)$ under a collusion scheme $(\mu,\hat{q},\hat{t})$ depends on both the realized type profile $\theta$ and the realized collusive message $m\sim\mu(\theta)$.

Substituting an optimal choice of $q$ for any fixed $\eta$ yields (3), as the total value $\sum_{i\in I}\theta_i q_i$ is maximized by allocating $\mathbb{E}_{\tilde{m}\sim\eta}\left[\sum_{i\in I} q_i(\tilde{m})\right]$ to bidders with the highest value $\max_{i\in I}\theta_i$.

To prove Part (b), note that Part (a) implies almost-sure IC for the single buyer. The presence of the message profile $NP$ ensures that the maximum in (3) is nonnegative, and almost-sure IR for the single buyer follows. □

Both parts of Theorem 1 follow straightforwardly from Lemma 2.

*Proof of Theorem 1.* Focusing first on the case of efficient collusion, Part (a) is implied by Lemma 2(b) and the optimality of a posted price for selling to a single buyer (Myerson, 1981; Riley and Zeckhauser, 1983). As for Part (b), for the $p^\star$-posted-price mechanism, (3) requires almost surely sending message profiles $m \neq NP$ when $\max_{i\in I}\theta_i > p^\star$, and message profile $NP$ when $\max_{i\in I}\theta_i < p^\star$. Hence, due Lemma 2(a), *ex-post* revenue under any efficient collusion scheme must be $r(p^\star;\theta)$ almost surely.

By Proposition 1(b), both parts extend to *ex-ante* incentive-efficient collusion. □

3.4. **Other Optimal Mechanisms.** Theorem 1 shows that an optimal posted-price mechanism is optimal under efficient collusion, but not that it is uniquely optimal. We now characterize optimal mechanisms under efficient collusion by means of two properties that are sufficient and generically necessary. The first property is that the mechanisms always charge the bidders an aggregate payment of at least $p^\star$ per unit. The second is that they allow the bidders to jointly obtain the good with probability 1 for an aggregate payment equal to $p^\star$.

The necessity of these properties to obtain the non-collusive posted-price expected revenue $R(p^\star)$ under efficient collusion relies on the optimal price $p^\star$ being unique, which holds generically. Versions of these properties are also sufficient to ensure an *ex-post* revenue of $r(p;\theta)$ almost surely under efficient collusion for any price $p$.

**Proposition 2.** *(a) Let $p$ be a price. If a selling mechanism $(M,q,t,\sigma)$ satisfies*

$(\mathrm{AP}_{\geq p})$ $\quad \sum_{i\in I} t_i(m) \geq p\sum_{i\in I} q_i(m)$ *for all message profiles $m\in M$,*

$(\mathrm{AP}^{\hat{m}}_{=p})$ *there exists a message profile $\hat{m}\in M$ with $\sum_{i\in I} q_i(\hat{m}) = 1$ and $\sum_{i\in I} t_i(\hat{m}) = p$,*

*then, under any efficient collusion scheme, ex-post revenue is $r(p;\theta)$ almost surely.*

*(b) Conversely, if the optimal posted price $p^\star$ is unique, then every selling mechanism that yields expected revenue $R(p^\star)$ under some efficient collusion scheme satisfies $(\mathrm{AP}_{\geq p})$ and $(\mathrm{AP}^{\hat{m}}_{=p})$ for $p = p^\star$.*

*Proof of Proposition 2(a).* Let $(\mu, \hat{q}, \hat{t})$ be an efficient collusion scheme for $(M, q, t, \sigma)$. First, note that ($\text{AP}_{\geq p}$) implies that the coalitional surplus in the maximization problem in (3) is bounded above by

$$\sup_{m \in M} \left[ \left( \max_{i \in I} \theta_i - p \right) \sum_{i \in I} q_i(m) \right] \leq \max \left\{ \left( \max_{i \in I} \theta_i - p \right), 0 \right\}. \tag{4}$$

On the other hand, due to ($\text{AP}^{\hat{m}}_{=p}$), the coalitional surplus in (3) attains the right-hand side in (4) by sending the message profile $\hat{m}$ for type profiles with $\max_{i \in I} \theta_i > p$ and the message profile $NP$ for type profiles with $\max_{i \in I} \theta_i \leq p$. Hence, (3) requires almost surely sending message profiles $m$ with $\sum_{i \in I} q_i(m) = 1$ when $\max_{i \in I} \theta_i > p$ and $\sum_{i \in I} q_i(m) = 0$ when $\max_{i \in I} \theta_i < p$, and for which ($\text{AP}_{\geq p}$) holds with equality. Hence, by Lemma 2(a), *ex-post* revenue under the collusion scheme is almost surely

$$\sum_{i \in I} t_i(\tilde{m}) = p \mathbf{1} \left( \max_{i \in I} \theta_i \geq p \right) = r(p; \theta),$$

as claimed. □

To prove Proposition 2(b), we use the fact that the uniqueness of the optimal posted price means that there is essentially a unique optimal mechanism for selling to the grand coalition. The formal argument is in the online appendix.

For example, consider a standard auction with the optimal posted price $p^\star$ used as a reserve, as proposed by Graham and Marshall (1987). Such auctions allow the bidders to obtain the good for $p^\star$ (by one bidder bidding the reserve, and no one bidding above it), and always require the bidders to pay at least $p^\star$ to obtain the good. Therefore, by Proposition 2(a), they guarantee the non-collusive expected revenue of an optimal posted price under efficient collusion. Moreover, they also leave the possibility of obtaining higher expected revenue if bidders do not collude, so may be preferable if a seller is concerned only about efficient collusion but thinks there is a (small) probability that collusion does not occur.

If bidders are symmetric, then the optimal posted price is strictly increasing in the number of bidders. This comparative static was shown by McAfee and McMillan (1987) for the case of regular value distributions, while in Online Appendix D, we establish it without assuming regularity (instead assuming only a continuous, positive density). In particular, as argued by McAfee and McMillan (1987), it follows that with regular value distributions and at least two bidders, the optimal posted price is greater than Myerson's (1981) optimal reserve, which is the optimal posted price for one bidder. From this perspective, our results show that the best the seller can do to

counter the possibility of *efficient* collusion using any selling mechanism is simply to set a higher reserve—as highlighted by Cassady (1967, pp. 228–230).

## 4. Difficulties in Guaranteeing Revenue under Inefficient Collusion

In this section, we provide examples that illustrate issues with guaranteeing the optimal non-collusive posted-price revenue under inefficient collusion. In Section 4.1, we show that neither an optimal posted-price mechanism nor a standard auction that uses the optimal posted price as a reserve guarantees the optimal non-collusive posted-price expected revenue under *inefficient collusion*. A natural strategy to ensure robustness to collusion would be to combine the posted-price mechanism with an efficient collusion scheme. In Section 4.2, we show that this strategy does not always work, as it can leave the seller vulnerable to *collusion by smaller rings* than the grand coalition that harms non-colluding bidders. In Section 4.3, we show that the seller can also be left vulnerable to collusion by the grand coalition if *a bidding ring can commit to punishing bidders who refuse to join*. Finally, in Section 4.4, we provide a realistic setting in which these issues arise, which is when bidders collude implicitly by bidding via a demand-side platform that seeks to maximize its own expected revenue.

### 4.1. Vulnerability to Inefficient Collusion.

It turns out that posted-price mechanisms leave the seller vulnerable to inefficient collusion, even via IC and *ex-ante* Pareto-improving collusion schemes.

*Example* 1 (Vulnerability of Posted Prices to Inefficient Collusion). There are two bidders whose values are each distributed uniformly between 1 and 2. Consider the posted-price mechanism with optimal posted price $p^\star = 4/3$. Each bidder's *ex-ante* expected utility is $V_i(M, q, t, \sigma) = 4/27$, and expected revenue is $R(p^\star) = 32/27$.

There is a way for bidders to collude to buy the good only if the higher bidder's value is at least $\hat{p} = 3/2$, thereby decreasing the probability the good is sold and hence lowering expected revenue. The collusion scheme that we construct is IC, and allocates the good efficiently among the bidders if it is sold (as a bidder with a lower value never participates). While it is (*ex-ante* incentive-)inefficient, it improves allocative efficiency enough relative to the rationing entailed by a posted price to be *ex-ante* Pareto-improving.

Formally, consider the collusion scheme with deterministic message function

$$\mu_i(\theta) = \begin{cases} P_i & \text{if } \theta_i \geq \hat{p} \text{ and } \theta_i \geq \theta_{-i} \\ NP_i & \text{otherwise} \end{cases},$$

allocation rule $\hat{q}(\theta) = q(\mu(\theta))$, and payment rule

$$\hat{t}_i(\theta) = p^\star \hat{q}_i(\theta) + \left(\tfrac{1}{2}(\theta_i - \hat{p})^2 + (\theta_i - 1)(\hat{p} - p^\star)\right)\mathbf{1}(\theta_i \geq \hat{p}) - \left(\tfrac{1}{2}(\theta_{-i} - \hat{p})^2 + (\theta_{-i} - 1)(\hat{p} - p^\star)\right)\mathbf{1}(\theta_{-i} \geq \hat{p}).$$

(1) and (2) hold by construction. The collusion scheme is IC.[22] It delivers an *ex-ante* expected utility of $U_i(\hat{q}, \hat{t}) = 1/6 > 4/27 = V_i(M, q, t, \sigma)$ to each bidder, so is *ex-ante* Pareto-improving. But it leads to an expected revenue of $1 < 32/27 = R(p^\star)$. □

For standard auctions with reserve $p^\star$, efficient collusion already lowers expected revenue to the non-collusive posted-price expected revenue $R(p^\star)$ (by Proposition 2(a)). We show that inefficient collusion can lower expected revenue even below that.

*Example* 2 (Inefficient Collusion Can Lower the Expected Revenue of Standard Auctions Below the Non-Collusive Posted-Price Expected Revenue). As in Example 1, there are two bidders whose values are each distributed uniformly between 1 and 2, and the optimal posted price is $p^\star = 4/3$. Consider the (monotone, symmetric) Bayes–Nash equilibrium of any standard auction with reserve $p^\star$. Each bidder's *ex-ante* expected utility is $V_i(M, q, t, \sigma) = 10/81$.

There is a way for bidders to collude to buy the good only if the higher bidder's value is at least $\hat{p} = 3/2$, and only pay $p^\star = 4/3$ if they buy. This lowers expected revenue even below the non-collusive posted-price expected revenue. The collusion scheme that we construct is IC. While it is (*ex-ante* incentive-)inefficient, it lowers the (aggregate) payment to the seller conditional on sale to the reserve, so improves coalitional surplus for the bidders enough to be *ex-ante* Pareto-improving.

Formally, consider the collusion scheme with deterministic message function

$$\mu_i(\theta) = \begin{cases} p^\star & \text{if } \theta_i \geq \hat{p} \text{ and } \theta_i \geq \theta_{-i} \\ 0 & \text{otherwise} \end{cases},$$

and allocation and payment rules as in Example 1. (1) and (2) hold by construction, and the collusion scheme is IC. It delivers an *ex-ante* expected utility of $U_i(\hat{q}, \hat{t}) = 1/6 > 10/81 = V_i(M, q, t, \sigma)$ to each bidder, so is *ex-ante* Pareto-improving. But, as in Example 1, it leads to an expected revenue of $1 < 32/27 = R(p^\star)$. □

Collusion is not always efficient in practice, especially with asymmetric bidders (Asker, 2010). Hence, in the remainder of the paper, we seek to design mechanisms

[22] Indeed, the allocation rule is monotone and expected payments satisfy an envelope formula.

that guarantee the optimal non-collusive posted-price expected revenue even under inefficient collusion.

4.2. **Vulnerability to Collusion by Smaller Rings.** Our overall strategy to construct mechanisms that are robust to collusion is to use the logic of Laffont and Martimort's (1997) Collusion-Proofness Principle: to deter bidders from colluding, we combine a posted-price mechanism with a specific efficient collusion scheme implemented by the seller on behalf of the bidders. While this approach does ensure robustness to *ex-ante* Pareto-improving collusion, it may leave the seller vulnerable to collusion that harms some bidders, such as collusion by smaller rings that harms non-colluding bidders.[23]

For an illustration, consider combining the $p^\star$-posted-price mechanism with the efficient collusion scheme proposed by Mailath and Zemsky (1991). In the symmetric case, that collusion scheme is an all-pay knockout auction for the right to buy the good at price $p^\star$ in which each bidder's knockout bid is paid to the other bidders (dividing it equally among them).[24] With three bidders, the combined selling mechanism is vulnerable to collusion by a pair of bidders running a private knockout auction to determine who bids in the selling mechanism's knockout auction, thereby saving the side payment the loser of the private knockout auction would have made to the third bidder in the seller's knockout auction.[25] Along the way, the ring can end up lowering the probability the good is sold, thereby lowering revenue.

*Example* 3 (Collusion by Smaller Rings). There are three bidders, whose values are each distributed uniformly between 0 and 1. The optimal posted price is $p^\star = 1/\sqrt[3]{4} \approx 0.63$.

If the seller posted the price $p^\star$ or ran any standard auction with reserve $p^\star$, Mailath and Zemsky's (1991) efficient collusion scheme would allocate the good to the highest bidder, provided that his value were at least $p^\star$, and pay $p^\star$ to the seller in that case.

[23]Indeed, the Collusion-Proofness Principle does not generally hold in our setting because collusion may not be Pareto-improving for the bidders. Moreover, collusion schemes need not provide ring members with nonparticipation messages, while the seller is required to provide each bidder with a nonparticipation message, which can prevent the seller from implementing collusion on behalf of the bidders. Nevertheless, the logic behind the Collusion-Proofness Principle guides our construction.

[24]Alternative efficient collusion schemes have been proposed for the symmetric case by Graham and Marshall (1987) and McAfee and McMillan (1992). As Graham and Marshall's (1987) collusion scheme is not *ex-post* budget-balanced, combining a posted price with it would leave the seller vulnerable even to efficient collusion by the grand coalition.

[25]This vulnerability is closely related to the vulnerability to collusion of Cramton et al.'s (1987, Section 4) efficient partnership dissolution mechanism, which corresponds to the case of a posted price of 0. Cramton et al. (1987, page 624) and Mailath and Zemsky (1991, page 476) noted that their all-pay mechanisms are vulnerable to collusion.

In the all-pay knockout auction, the equilibrium bid of type $\theta_i$ would be

$$g(\theta_i) = (\theta_i - p^\star)^2 \frac{2\theta_i + p^\star}{3} \mathbf{1}(\theta_i \geq p^\star),$$

making bidder $i$'s side payment be given by

$$g(\theta_i) - \frac{1}{2} \sum_{j \in I \setminus \{i\}} g(\theta_j).$$

Combining this collusion scheme with the $p^\star$-posted-price mechanism (or any standard auction with reserve $p^\star$) gives a direct selling mechanism with payment rule[26]

$$t_i(\theta) = p^\star q_i(\theta) + g(\theta_i) - \frac{1}{2} \sum_{j \in I \setminus \{i\}} g(\theta_j).$$

For this mechanism, under truthful revelation, expected revenue is $R(p^\star) \approx 0.47$. Moreover, since the mechanism is efficient for the bidders given the price, any *ex-ante* Pareto-improving collusion scheme must almost surely leave the allocation unchanged, and hence leave expected revenue unchanged at $R(p^\star)$ as well.

However, the mechanism leaves the seller vulnerable to collusion by rings consisting of pairs of bidders that harms the third bidder. Specifically, we construct collusion schemes in which bidders 1 and 2 collude so that the lower-value bidder between them reports zero; and the higher-value bidder reports truthfully provided that his value is at least $\hat{p} = 2/3 > p^\star$, and reports zero otherwise. This results in an efficient allocation of the good between bidders 1 and 2, and they therefore do not reallocate it.[27]

We first suppose that bidder 3 is unaware of the collusion between bidders 1 and 2, and therefore plays the mechanism non-collusively—i.e., reports truthfully. We construct side payments that ensure that the collusion scheme is IC for bidders 1 and 2.[28] Under this collusion scheme, bidders 1 and 2 each obtain an *ex-ante* expected

[26]The selling mechanism must also specify what happens when some bidders do not participate, which is not material for this example but is important for the next one.

[27]Thus, breaking ties in favor of bidder 1 for simplicity, bidders 1 and 2 collude to send message profile

$$\mu^{\{1,2\}}(\theta_1, \theta_2) = \begin{cases} (\theta_1, 0) & \text{if } \theta_1 \geq \theta_2 \text{ and } \theta_1 \geq \hat{p} \\ (0, \theta_2) & \text{if } \theta_2 > \theta_1 \text{ and } \theta_2 \geq \hat{p} \\ (0, 0) & \text{otherwise} \end{cases}.$$

The collusive allocation rule is $\hat{q}(\theta) = q(\mu^{\{1,2\}}(\theta_{1,2}), \theta_3)$.

[28]In the case in which bidder 3 is unaware of collusion, the side payment from bidder 1 to bidder 2 is

$$\Big(\frac{2\theta_1^3 + \hat{p}^3}{3} - \theta_1^2 p^\star - \theta_1 g(\theta_1) - \frac{1}{2} \int_{\hat{p}}^{\theta_1} g(x)\, dx\Big) \mathbf{1}(\theta_1 \geq \hat{p}) - \Big(\frac{2\theta_2^3 + \hat{p}^3}{3} - \theta_2^2 p^\star - \theta_2 g(\theta_2) - \frac{1}{2} \int_{\hat{p}}^{\theta_2} g(x)\, dx\Big) \mathbf{1}(\theta_2 \geq \hat{p}).$$

utility of $\approx 0.0535$, which exceeds their *ex-ante* expected utilities $V_i(M,q,t,\sigma) \approx 0.0531$ from reporting truthfully. The collusion scheme lowers expected revenue to $\approx 0.45 < R(p^\star)$. This decline in revenue is possible because the collusion scheme harms bidder 3: his *ex-ante* expected utility under the collusion scheme is $\approx 0.0519$, compared to $V_3(M,q,t,\sigma) \approx 0.0531$ if all bidders reported truthfully.

We next suppose that bidder 3 is aware of the collusion, and therefore plays a best response to it. In this case, bidder 3 shades his reports to take advantage of the missing mass of competing reports between $p^\star$ and $\hat{p}$,[29] but is still harmed by collusion: his *ex-ante* expected utility is $\approx 0.0520$, compared to $V_3(M,q,t,\sigma) \approx 0.0531$ if all bidders reported truthfully. We can still construct collusive side payments that ensure that the collusion scheme is IC for bidders 1 and 2.[30] Under this collusion scheme, bidders 1 and 2 each obtain an *ex-ante* expected utility of $\approx 0.0534$, which exceeds their *ex-ante* expected utilities $V_i(M,q,t,\sigma) \approx 0.0531$ from reporting truthfully. The collusion scheme still lowers expected revenue to $\approx 0.45 < R(p^\star)$. □

Our constructions ensure robustness to collusion by all rings by combining a posted price with an efficient collusion scheme that, unlike Mailath and Zemsky's (1991) scheme, is itself robust to collusion that may harm some bidders.

4.3. **Vulnerability to Collusion by Rings that Commit to Punish.** Even when we restrict attention to collusion that involves all the bidders, inefficient collusion may be enforced by the ring's commitment to punish bidders who refuse to join it, as in Quesada (2005) and Dequiedt (2007).

*Example* 4 (Collusion by Rings that Commit to Punish). Consider again the two-bidder setting of Example 1. Consider the direct mechanism that combines the optimal posted price $p^\star = 4/3$ (or any standard auction with reserve $p^\star$) with Mailath and

[29] Specifically, bidder 3's best response to the collusion by bidders 1 and 2 under consideration is to shade his report to $p^\star$ if his value is slightly above $p^\star$—i.e., to use the strategy

$$\hat{\sigma}_3(\theta_3) = \begin{cases} p^\star & \text{if } p^\star \le \theta_3 \le \frac{-p^\star + \sqrt{3(4\hat{p}^2 - (p^\star)^2)}}{2} \\ \theta_3 & \text{otherwise} \end{cases}.$$

[30] In the case in which bidder 3 plays a best response to collusion, writing $\hat{\theta}_3 = \frac{-p^\star + \sqrt{3(4\hat{p}^2 - (p^\star)^2)}}{2}$ for the cutoff value below which bidder 3 shades, the side payment from bidder 1 to bidder 2 is

$$\begin{aligned} &\Big(\frac{4\theta_1^3 + 3\hat{\theta}_3\hat{p}^2 - \hat{\theta}_3^3}{6} - \theta_1^2 p^\star\Big)\mathbf{1}(\theta_1 > \hat{\theta}_3) + \Big(\frac{\theta_1^2\hat{\theta}_3 + \hat{\theta}_3\hat{p}^2}{2} - \theta_1\hat{\theta}_3 p^\star\Big)\mathbf{1}(\hat{\theta}_3 \ge \theta_1 \ge \hat{p}) \\ &\quad - \Big(\theta_1 g(\theta_1) + \frac{1}{2}\int_{\hat{p}}^{\theta_1} g(x)\,dx\Big)\mathbf{1}(\theta_1 \ge \hat{p}) + \Big(\theta_2 g(\theta_2) + \frac{1}{2}\int_{\hat{p}}^{\theta_2} g(x)\,dx\Big)\mathbf{1}(\theta_2 \ge \hat{p}) \\ &\quad - \Big(\frac{4\theta_2^3 + 3\hat{\theta}_3\hat{p}^2 - \hat{\theta}_3^3}{6} - \theta_2^2 p^\star\Big)\mathbf{1}(\theta_2 > \hat{\theta}_3) - \Big(\frac{\theta_2^2\hat{\theta}_3 + \hat{\theta}_3\hat{p}^2}{2} - \theta_2\hat{\theta}_3 p^\star\Big)\mathbf{1}(\hat{\theta}_3 \ge \theta_2 \ge \hat{p}). \end{aligned}$$

Zemsky's (1991) efficient collusion scheme. This mechanism allocates the good to the highest bidder, provided that his value exceeds $p^\star$. Its payment rule is given by

$$t_i(\theta) = p^\star q_i(\theta) + \tfrac{1}{2}(\theta_i - p^\star)^2 \mathbf{1}(\theta_i \geq p^\star) - \tfrac{1}{2}(\theta_{-i} - p^\star)^2 \mathbf{1}(\theta_{-i} \geq p^\star).$$

Each bidder's *ex-ante* expected utility under truthtelling is $\approx 0.173$.

Now suppose that if either bidder sends a nonparticipation message, the seller implements the null outcome. Then, inefficient collusion can be sustained by one bidder threatening the other bidder with nonparticipation. For example, bidders can sustain the IC collusion scheme with no reallocation or side payments in which bidder 2 always reports 1, and bidder 1 reports 5/3 when $\theta_1 > 25/18$ and reports 1 otherwise. This collusion scheme gives bidder 1 an *ex-ante* expected utility of $\approx 0.187 > 0.173$. It gives bidder 2 an expected rebate of $\approx 0.034$, and he is induced to participate in it by bidder 1's threat of not participating in the selling mechanism—thereby canceling the selling mechanism—if collusion does not occur. Note that this collusion scheme reduces the expected revenue to $22/27 < R(p^\star) = 32/27$. □

This example demonstrates, in particular, that to avoid the possibility that bidders force other bidders to collude using a threat of nonparticipation, outcomes following nonparticipation by some bidders must be defined more carefully.

4.4. **Implicit Collusion via a Demand-Side Platform.** We next tie together the issues highlighted in the previous subsections in a more realistic example. Specifically, we show that either posting a price of $p^\star$ or running a standard auction with reserve $p^\star$ does not yield the non-collusive posted-price revenue when collusion occurs implicitly by bidding via a demand-side platform (DSP) that maximizes its own revenue. The DSP faces interim individual rationality (interim IR) constraints in enticing bidders to join, and can commit to punishing bidders who refuse to join via the bids of other bidders. To maximize its own expected revenue, the DSP must run its own auction among the bidders to determine who gets to buy at price $p^\star$. Due to double marginalization, the reserve price in the DSP's optimal auction is above $p^\star$—which makes the probability that the good is sold lower, and hence reduces the seller's expected revenue below what it would be under efficient collusion.

We focus here on collusion via a DSP in a standard auction with reserve $p^\star$; Online Appendix C shows that posted-price mechanisms have a similar vulnerability.[31]

[31]In that case, the DSP-revenue-maximizing collusion scheme is more involved due to the bidders' interim reservation utilities being positive and type-dependent, which leads to the DSP procuring the good with a positive probability even if all bidders' values are below the DSP's optimal reserve.

*Example* 5 (Implicit Collusion via a DSP in a Standard Auction). There are two bidders whose values are each distributed uniformly between 0 and 1. Consider a standard auction with reserve equal to the optimal posted price $p^\star = 1/\sqrt{3} \approx 0.58$. Under efficient collusion, expected revenue is $R(p^\star) \approx 0.38$.

Note that the DSP can punish a bidder who refuses to participate in the ring by subsidizing a high bid for the other bidder, driving the defector's utility to 0. Thus, the DSP faces interim IR constraints with the usual outside option of 0.

The DSP designs a message function $\mu : \Theta \to \Delta(M)$ and an IC direct mechanism $(\hat{q}, \hat{t}) : \Theta \to \mathcal{Q} \times \mathbb{R}^2$ satisfying the allocation feasibility constraint (1) and interim IR to maximize its own expected revenue

$$\mathbb{E}_{\tilde{\theta}}\Big[\sum_{i\in I} \hat{t}_i(\tilde{\theta}) - \mathbb{E}_{\tilde{m}\sim\mu(\tilde{\theta})}\Big[\sum_{i\in I} t_i(\tilde{m})\Big]\Big].$$

Since the DSP can procure the good for price $p^\star$ by making one bidder bid $p^\star$ and the other bidder not bid, its revenue-maximization problem reduces to maximizing

$$\mathbb{E}_{\tilde{\theta}}\Big[\sum_{i\in I} \hat{t}_i(\tilde{\theta}) - p^\star \sum_{i\in I} \hat{q}_i(\tilde{\theta})\Big]$$

over all IC and interim IR mechanisms $(\hat{q}, \hat{t})$.

This is a standard optimal auction problem with cost $p^\star$, whose solutions (almost surely) allocate the good to the bidder with the highest value provided that his value exceeds the DSP's optimal reserve $\hat{p} = (1 + p^\star)/2 \approx 0.79$ and (almost surely) do not allocate the good otherwise. Hence, in contrast to efficient collusion, the good is (almost surely) not sold if the higher bidder's value is between $p^\star \approx 0.58$ and $\hat{p} \approx 0.79$—lowering the probability that the good is allocated from $2/3 \approx 0.67$ to $(4-\sqrt{3})/6 \approx 0.38$. Thus, implicit collusion via a DSP that maximizes its own expected revenue lowers the seller's expected revenue to $\approx 0.22 < 0.38 \approx R(p^\star)$. □

## 5. Strong Collusion-Robustness

In this section, we formulate a model of collusion that need not be *ex-ante* Pareto-improving for bidders, generalizing the examples in Section 4. We then show that guaranteeing the optimal non-collusive posted-price expected revenue under all the forms of collusion that we consider requires protecting bidders' *ex-ante* utilities from the same forms of collusion—a property that we term "strong collusion-robustness." Last, we develop the implications of strong collusion-robustness and compare it to other collusion-proofness concepts from the literature.

While our focus is single-item auctions where rings can reallocate the good, most of the results of this section apply more broadly to other mechanism design settings.[32]

5.1. **Modeling Collusion That May Harm Collusion-Aware Bidders.** Collusion by some bidders may harm non-colluding bidders who are aware of it (see Example 3). Moreover, if a ring can commit to punish bidders who refuse to join the collusion, then bidders may find it in their interest to join collusion even if it harms them (see Example 4). We first consider the latter possibility, and then argue that the former possibility is a special case of the latter.

Suppose that the grand coalition, acting as a bidding ring in mechanism $(M, q, t, \sigma)$, can commit to punishing a bidder $i$ who refused to join *ex ante* by an arbitrary (measurable) plan $\nu^i : \Theta_{-i} \to \Delta(M_{-i})$ specifying the other bidders' (randomized) messages. If bidder $i$ responds with a (measurable) strategy $\sigma_i' : \Theta_i \to M_i$, his *ex-ante* expected utility is given by

$$V_i(M, q, t, \sigma_i', \nu^i) = \mathbb{E}_{\tilde{\theta}}\Big[\mathbb{E}_{\tilde{m}_{-i} \sim \nu^i(\tilde{\theta}_{-i})}\big[\tilde{\theta}_i q_i\big(\sigma_i'(\tilde{\theta}_i), \tilde{m}_{-i}\big) - t_i\big(\sigma_i'(\tilde{\theta}_i), \tilde{m}_{-i}\big)\big]\Big].$$

**Definition 5.** A collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism $(M, q, t, \sigma)$ is *individually rational (IR) for bidder $i$* if there exists a (measurable) punishment scheme $\nu^i : \Theta_{-i} \to \Delta(M_{-i})$ such that[33]

$$U_i(\hat{q}, \hat{t}) \geq V_i(M, q, t, \sigma_i', \nu^i) \quad \text{for all (measurable) strategies } \sigma_i' : \Theta_i \to M_i. \tag{5}$$

A collusion scheme is *individually rational (IR)* if it is IR for all bidders.

In words, a collusion scheme is IR if, for each bidder $i$, the remaining bidders $j \in I \setminus \{i\}$ can ensure that $i$ will find it in his interest to participate in the collusion scheme at the *ex-ante* stage by threatening him with a punishment in case he does not join, even if $i$ plays a best response to such punishment.[34]

IR is also satisfied when the collusive ring does not include all the bidders, provided that the noncolluding bidders are aware of collusion and play best responses to it

[32]Specifically, all of the results of this section except for Proposition 3(b) hold for general allocation problems (under continuity conditions on the type distributions for statements about revenue) and apply regardless of whether bidding rings can reallocate goods.

[33]Note that (i) the restriction that bidder $i$ respond to punishment with a pure strategy $\sigma_i'$ is without loss by the randomization lemma for probability kernels (Kallenberg, 2002, Lemma 2.22); (ii) without loss, $\nu^i$ can be taken to be constant in $\theta_{-i}$, which would trivially ensure incentive compatibility for the punishing bidders; and (iii) we allow message profiles $m_{-i} \sim \nu^i(\theta_{-i})$ to have component messages that are correlated across punishing bidders $j \in I \setminus \{i\}$.

[34]If bidders decide whether to participate in collusion after having learned some information (e.g., at the interim stage), then, by the law of iterated expectations, our (*ex-ante*) IR property also holds.

and to each others' strategies. Indeed, we can take a punishment scheme $\nu^i$ for a noncolluding bidder $i$ that consists of the ring's collusive messages and the responses of the other noncolluding bidders. See Online Appendix F for a formal statement.

In particular, this applies when the "ring" is empty and so the "collusion scheme" simply involves playing a Bayes–Nash equilibrium with no reallocation or side payments.[35] For ease of discussion, given a strategy profile $\hat{\sigma}$, we refer to the corresponding collusion scheme with no reallocation or side payments by simply $\hat{\sigma}$.

**Lemma 3.** *Each Bayes–Nash equilibrium of a selling mechanism defines an IC and IR collusion scheme for that selling mechanism.*

*Proof.* Given a Bayes–Nash equilibrium $\hat{\sigma}$, letting the punishment scheme for bidder $i$ be $\nu^i = \hat{\sigma}_{-i}$, (5) follows from the (*ex-ante*) best-response property for $\hat{\sigma}_i$. The IC property is immediate from the definition of Bayes–Nash equilibrium. □

Finally, note that if $\sigma$ itself is IR (in particular, as per Lemma 3, if $\sigma$ is a Bayes–Nash equilibrium), then each *ex-ante* Pareto-improving collusion scheme is also IR.

5.2. **Modeling Collusion that Some Bidders Are Unaware Of.** We next consider collusion where bidders $i \in S$ are unaware of collusion, so follow the strategies $\sigma_i$ prescribed by the selling mechanism. This may occur, for example, for a (private, short-term) joint deviation from non-collusive play by the bidders in $I \setminus S$.

Formally, we consider collusion schemes in which bidders $i \in S$ follow $\sigma_i$, the bidders in $I \setminus S$ (who are aware of collusion) send joint messages according to a message function $\mu_{-S}(\theta_{-S})$ depending only on their types, and reallocation and side payments do not involve the unaware bidders $i \in S$.

**Definition 6.** Let $S \subseteq I$. A collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism $(M, q, t, \sigma)$ is *$S$-unaware* if there exists a function $\mu_{-S} : \Theta_{-S} \to \Delta(M_{-S})$ such that we have $\mu(\theta) = \mu_{-S}(\theta_{-S}) \times \sigma_S(\theta_S)$ for all type profiles $\theta \in \Theta$, and

$$\left(\hat{q}_i(\theta), \hat{t}_i(\theta)\right) = \mathbb{E}_{\tilde{m} \sim \mu(\theta)}\left[\left(q_i(\tilde{m}), t_i(\tilde{m})\right)\right] \quad \text{for all } i \in S. \tag{6}$$

Since bidders $i \in S$ just follow their prescribed strategies, they do not generally play best responses to collusion, so their IR constraint need not be satisfied. To avoid having to distinguish collusion-aware and collusion-unaware bidders in our arguments, it is convenient for us to work with a weaker notion of IR (and therefore even stronger notions of collusion and robustness to collusion), which constrains any defector to respond to punishments with his prescribed strategy under the selling mechanism.

[35] We consider *ex-ante* Bayes–Nash equilibrium throughout for technical convenience.

**Definition 7.** A collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism $(M, q, t, \sigma)$ is *weakly individually rational (weakly IR)* if for each bidder $i$, there exists a (measurable) punishment scheme $\nu^i : \Theta_{-i} \to \Delta(M_{-i})$ such that[36]

$$U_i(\hat{q}, \hat{t}) \geq V_i(M, q, t, \sigma_i, \nu^i). \tag{7}$$

Constraining defectors to using their prescribed strategies means that weakly IR collusion subsumes any form of collusion that is IR for all bidders who are aware of it.[37]

**Lemma 4.** *(a) For all $S \subseteq I$, any $S$-unaware collusion scheme for a selling mechanism that is IR for each bidder $i \in I \setminus S$ is weakly IR.*[38]

*(b) In particular, every IR collusion scheme for a selling mechanism is weakly IR.*

*Proof.* To prove Part (a), consider a selling mechanism $(M, q, t, \sigma)$ and an $S$-unaware collusion scheme $(\mu, \hat{q}, \hat{t})$ that is IR for all $i \in I \setminus S$. For each bidder $i \in I \setminus S$, consider a punishment scheme $\nu^i$ that satisfies (5). Taking $\sigma_i' = \sigma_i$, we can see that (7) holds for the same punishment scheme $\nu^i$. Letting $\mu_{-S} : \Theta_{-S} \to \Delta(M_{-S})$ be such that $\mu(\theta) = \mu_{-S}(\theta_{-S}) \times \sigma_S(\theta_S)$, for bidders $i \in S$, (6) implies that (7) holds for the punishment scheme $\nu^i(\theta_{-i}) = \mu_{-S}(\theta_{-S}) \times \sigma_{S\setminus\{i\}}(\theta_{S\setminus\{i\}})$.

Part (b) follows from Part (a) by taking $S = \varnothing$. □

*Remark* 1. Lemma 4(a) can be strengthened to apply to collusion of which bidders in $S \subseteq I$ are unaware, where subsequent to the mechanism, members of $I \setminus S$ can resell the item to bidders in $S$. If the resale makes the members of $S$ no worse off *ex post*, it makes them no worse off *ex ante*, hence the collusion scheme incorporating such resale is weakly IR.

Due to their incorrect beliefs, collusion is generally not IC for collusion-unaware bidders $i \in S$. We therefore do not impose IC when considering weakly IR collusion. This also allows a ring to observe its participants' types, so as not to face incentive constraints among members—the only restriction (implicit in our definitions of IR and weak IR) is that colluding bidders cannot observe non-colluding bidders' types, so cannot use that information to punish bidders who do not join. Allowing for a broader range of collusion schemes strengthens our conclusions about the robustness of the mechanisms constructed below. However, the results below can also be derived under a restriction to weakly IR collusion schemes that are also IC.

[36]Just like in Definition 5, $\nu^i$ can be restricted to be constant in $\theta_{-i}$ without loss, but here it can be further restricted without loss to send a constant message profile $m_{-i} \in M_{-i}$ with probability 1.

[37]By construction, every *ex-ante* Pareto-improving collusion scheme is weakly IR.

[38]Note that we allow punishments to involve bidders in $S$, which makes collusion easier and so strengthens our notion of robustness to weakly IR collusion.

5.3. **Strong Collusion-Robustness.** Our objective is to obtain the best expected-revenue guarantee against all weakly IR collusion. However, as foreshadowed by Examples 3 and 4, it turns out that achieving this also requires making bidders' *ex-ante* expected utilities robust to all such collusion, in the following sense.

**Definition 8.** A selling mechanism $(M, q, t, \sigma)$ is *strongly collusion-robust* if all weakly IR collusion schemes for it give the same profile of bidders' *ex-ante* expected utilities as if all bidders used $\sigma$.[39]

We now show that strong collusion-robustness is both sufficient for revenue to be robust to all weakly IR collusion, and necessary to guarantee at least the optimal non-collusive posted-price expected revenue $R(p^\star)$.

**Proposition 3.** *(a) If a selling mechanism $(M, q, t, \sigma)$ is strongly collusion-robust, then it almost surely yields the same ex-post revenue under each weakly IR collusion scheme for it as if all bidders used $\sigma$.*

*(b) Each selling mechanism that yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes for it is strongly collusion-robust.*[40]

We prove Proposition 3 in the online appendix. For an intuition for Part (b), a violation of strong collusion-robustness allows bidders to sustain an inefficient weakly IR collusion scheme with punishment threats (Jagadeesan et al., 2026, Proposition 2(b)). The following lemma implies that there is an *ex-ante* Pareto-dominating (and hence also weakly IR) collusion scheme under which expected revenue is less than $R(p^\star)$.

**Lemma 5.** *Given a selling mechanism, for each inefficient collusion scheme, there exists an ex-ante Pareto-dominating IC collusion scheme under which expected revenue is strictly less than $R(p^\star)$.*

[39] Our May 6, 2025 working paper also imposed two additional conditions: that weakly IR collusion schemes affect neither the allocation of the good (almost surely) nor the seller's expected revenue. Subsequently, Csóka et al. (2025) argued that these two conditions are not natural beyond the auction context. We therefore dropped those two conditions from the definition, and instead show that they are implied by the robustness of the bidders' *ex-ante* expected-utility profile. (Proposition 3(a) shows this for revenue, and its proof shows this for the allocation of the good.)
Strong collusion-robustness also implies that the same three robustness properties hold for collusion that is only budget-balanced *ex ante*. Indeed, any collusion scheme for which the weak budget balance condition (2) only holds *ex ante* can be made into one for which (2) holds without affecting bidders' *ex-ante* utilities, the allocation, or expected revenue, by adjusting one of the bidders' transfers to balance the collusive budget. (This argument exploits the possibility that collusion is not IC.)

[40] Alternatively, the same conclusion obtains if all weakly IR collusion schemes are required to yield the *same* positive expected revenue, no matter what it is.

The proof of Lemma 5 (in the online appendix) constructs a collusion scheme by mixing nonparticipation with probability $\varepsilon > 0$ and efficient collusion with probability $1 - \varepsilon$. Since the expected revenue under efficient collusion is at most $R(p^\star)$ (by Theorem 1(a)), the expected revenue under this mixed collusion scheme is at most $(1-\varepsilon)R(p^\star) < R(p^\star)$.

For an intuition for Proposition 3(a), note that under strong collusion-robustness, all weakly IR collusion schemes must be efficient, hence almost surely yield the same outcome as if there were a single buyer with the highest of the bidders' values. Since the value distributions are continuous, the *ex-post* revenue must therefore also be the same almost surely across all weakly IR collusion schemes, including $\sigma$.

5.4. **Role of the Prescribed Strategies and Connection to Full Implementation.** Strong collusion-robustness depends on the choice of prescribed strategies $\sigma$. In particular, it requires that the strategies form a Bayes–Nash equilibrium.

**Proposition 4.** *If a selling mechanism $(M, q, t, \sigma)$ is strongly collusion-robust, then $\sigma$ is a Bayes–Nash equilibrium.*

*Proof.* Note that for each bidder $i$, any (*ex-ante*) better response $\sigma_i'$ to $\sigma_{-i}$ defines a weakly IR collusion scheme with deterministic message function $\mu(\theta) = (\sigma_i'(\theta_i), \sigma_{-i}(\theta_{-i}))$ and no reallocation or side payments, sustained by the punishment schemes $\nu^i = \sigma_{-i}$ and $\nu^j(\theta_{-j}) = \sigma_i'(\theta_i) \times \sigma_{-i,j}(\theta_{-i,j})$ for $j \in I \setminus \{i\}$. Strong collusion-robustness implies that $\sigma_i'$ cannot be a (*ex-ante*) strictly better response. □

A strongly collusion-robust mechanism $(M, q, t, \sigma)$ may have another Bayes–Nash equilibrium $\hat{\sigma}$ such that $(M, q, t, \hat{\sigma})$ is not strongly collusion-robust. Nevertheless, strong collusion-robustness implies the weaker property of robustness to IR collusion, which does not depend on $\sigma$.

**Definition 9.** A selling mechanism is *collusion-robust* if all IR collusion schemes for it give the same profile of bidders' *ex-ante* expected utilities.

Csóka et al. (2026) introduced a similar concept under the name "collusion-proofness" in a setting where collusive rings cannot reallocate the good.[41]

Since every IR collusion scheme is weakly IR (Lemma 4(b)), every strongly collusion-robust mechanism is collusion-robust, but the converse is not generally true. (The relationship between these concepts is developed in Jagadeesan et al. (2026).)

[41] Csóka et al. (2026) considered collusion under which collusive *ex-ante* utilities exceed (correlated) *ex-ante* maxmin utilities, while our concept of IR collusion requires collusive *ex-ante* utilities to exceed (correlated) *ex-ante* minmax utilities. These concepts coincide for finite mechanisms but differ in general. See Jagadeesan et al. (2026) for further discussion.

(Strong) collusion-robustness also implies a form of full implementation.

**Proposition 5.** *If a selling mechanism is collusion-robust, then all Bayes–Nash equilibria give the same profile of bidders' ex-ante expected utilities, and yield the same ex-post revenue almost surely.*

The invariance of bidders' *ex-ante* expected utilities across Bayes–Nash equilibria holds due to the definition of collusion-robustness because every Bayes–Nash equilibrium is an IR collusion scheme (Lemma 3). The invariance of *ex-post* revenue for *strongly* collusion-robust mechanisms follows from Proposition 3(a). The proof in the online appendix extends this logic to mechanisms that are only collusion-robust.

5.5. **Relationship to Other Collusion-Proofness Concepts.** Strong collusion-robustness also implies other collusion-proofness concepts from the literature.

For each set $S$ of unaware bidders, any $S$-unaware collusion scheme that is Pareto-improving for bidders in $I \setminus S$ is weakly IR, with a punishment scheme for each bidder being the others playing their prescribed strategies. Therefore, strong collusion-robustness implies that bidders' *ex-ante* expected utilities are robust to all such collusion schemes, which is a form of *ex-ante* strong Nash equilibrium that Safronov (2018) introduced under the name "coalition-proofness."[42] For the special case in which $S = \varnothing$, this property implies robustness to collusion that is *ex-ante* Pareto-improving for all bidders, which we considered in Section 3.

Strong collusion-robustness and collusion-robustness also imply robustness to the broadest forms of interim collusion considered by Laffont and Martimort (1997, Section 8) in their "strong collusion-proofness" concept.[43] However, strong collusion-robustness and collusion-robustness are logically independent of "robust collusion-proofness" in the sense of Che and Kim (2006), who did not restrict attention to collusion schemes that are weakly IR in the sense of our Definition 7. Indeed, for our setting, Che and Kim's (2006) concept would allow all collusion schemes that deliver each type of each bidder a nonnegative interim expected utility, regardless of what utility the bidder could guarantee by using his prescribed strategy in the mechanism. On the other hand, Che and Kim's (2006) concept focuses on collusion after bidders learn their values, so does not ensure robustness to *ex-ante* collusion.

[42] Safronov (2018) allowed for side payments among deviators but not the reallocation of the good.

[43] Laffont and Martimort's (1997, Section 8) concept of "strong collusion-proofness" requires that there be no equilibrium with non-null collusion, while (strong) collusion-robustness instead only requires that all bidders be *ex-ante* indifferent to any collusion of the forms that we consider. Laffont and Martimort (1997) also assumed that the proposed collusion scheme maximizes bidders' total *ex-ante* utility subject to the constraints, which we do not assume here.

## 6. Constructing Strongly Collusion-Robust Selling Mechanisms

In this section, we construct two strongly collusion-robust mechanisms that guarantee the revenue of an optimal posted price. The first mechanism works in the case of symmetric bidders; the second mechanism is more complicated, but works in general.

Both constructions rely on a set of conditions that are sufficient for a selling mechanism to be strongly collusion-robust and guarantee the *ex-post* revenue $r(p;\theta)$ of a posted price of $p$. The same conditions are necessary for a mechanism to guarantee at least the optimal non-collusive posted-price expected revenue $R(p^\star)$ under weakly IR collusion in the generic case in which the optimal posted price is unique.

**Proposition 6.** *(a) Let $p$ be a price. If a selling mechanism $(M,q,t,\sigma)$ satisfies*

$$(\mathrm{EA}_p^\sigma)\quad q_i(\sigma(\theta)) = \begin{cases} 1 & \text{if } \theta_i > \max_{j\in I\setminus\{i\}}\theta_j \text{ and } \theta_i > p \\ 0 & \text{if } \theta_i < \max_{j\in I\setminus\{i\}}\theta_j \text{ or } \theta_i < p \end{cases} \quad \text{for all } i\in I \text{ almost surely,}$$

$$(\mathrm{AP}_{=p}^\sigma)\qquad \sum_{i\in I} t_i(\sigma(\theta)) - p\sum_{i\in I} q_i(\sigma(\theta)) = 0 \quad \text{almost surely,}$$

$$(\mathrm{AP}_{\geq p})\qquad \sum_{i\in I} t_i(m) - p\sum_{i\in I} q_i(m) \geq 0 \quad \text{for all message profiles } m\in M,$$

$$(\mathrm{GU}^\sigma)\qquad V_i(M,q,t,\sigma_i,m_{-i}) \geq V_i(M,q,t,\sigma) \quad \text{for all } i\in I \text{ and } m_{-i}\in M_{-i},$$

*then it is strongly collusion-robust and yields ex-post revenue $r(p;\theta)$ almost surely under each weakly IR collusion scheme.*

*(b) Conversely, if the optimal posted price $p^\star$ is unique, then every selling mechanism that yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes satisfies $(\mathrm{EA}_p^\sigma)$, $(\mathrm{AP}_{=p}^\sigma)$, $(\mathrm{AP}_{\geq p})$, and $(\mathrm{GU}^\sigma)$ for $p=p^\star$.*

$(\mathrm{EA}_p^\sigma)$ and $(\mathrm{AP}_{=p}^\sigma)$ require that if bidders follow the prescribed strategies $\sigma$, the mechanism must give an Efficient Allocation for the bidders given the price $p$, and the Aggregate Payment must equal the corresponding $p$-posted price payment. $(\mathrm{AP}_{\geq p})$, which appeared in Proposition 2, requires that the Aggregate Payment always be at least $p$ times the aggregate allocation. $(\mathrm{GU}^\sigma)$ is a Guaranteed-Utility property: if any bidder $i$ follows his prescribed strategy $\sigma_i$, then his *ex-ante* expected utility cannot be reduced by others deviating from their prescribed strategies $\sigma_{-i}$ (to any fixed message profile, and hence to any contingent message profile independent of $i$'s type).

*Proof of Proposition 6(a).* Consider any weakly IR collusion scheme $(\mu,\hat{q},\hat{t})$ for $(M,q,t,\sigma)$. Since bidders' values are independent and private, weak IR and $(\mathrm{GU}^\sigma)$ together imply

$$U_i(\hat{q},\hat{t}) \geq V_i(M,q,t,\sigma). \tag{8}$$

Moreover, ($\mathrm{EA}^{\sigma}_{p}$), ($\mathrm{AP}^{\sigma}_{=p}$), and ($\mathrm{AP}_{\geq p}$) imply that $\sigma$ is efficient. Hence, we have

$$\sum_{i\in I} U_i(\hat{q},\hat{t}) \leq \sum_{i\in I} V_i(M,q,t,\sigma). \tag{9}$$

Comparing the sum of (8) over all bidders to (9), we see that (8) and (9) must both hold with equality. Equality in (8) implies strong collusion-robustness.

Equality in (9) implies $(\mu,\hat{q},\hat{t})$ is efficient. To show that *ex-post* revenue under $(\mu,\hat{q},\hat{t})$ is almost surely $r(p;\theta)$, we consider two cases. If $\max_{i\in I}\tilde{\theta}_i > p$ with positive probability, then this conclusion follows from Proposition 2(a). Last, in the trivial case in which $\max_{i\in I}\tilde{\theta}_i \leq p$ holds almost surely, since the value distributions are continuous, we have $\max_{i\in I}\tilde{\theta}_i < p$ almost surely, and it then follows from Lemma 2(a) and ($\mathrm{AP}_{\geq p}$) that *ex-post* revenue under efficient collusion is $0 = r(p;\theta)$ almost surely. □

The proof of Proposition 6(a) uses ($\mathrm{GU}^{\sigma}$) and the efficiency of $\sigma$ to establish strong collusion-robustness. Subsequent to our original paper introducing strong collusion-robustness, Csóka et al. (2025, Proposition C.13) showed that in general transferable-utility settings, a selling mechanism $(M,q,t,\sigma)$ is strongly collusion-robust if and only if ($\mathrm{GU}^{\sigma}$) holds and $\sigma$ is efficient. In the online appendix, we derive Proposition 6(b) using the "only if" direction of that equivalence: an expected-revenue guarantee of at least $R(p^\star)$ under all weakly IR collusion implies strong collusion-robustness (by Proposition 3(b)), and obtaining expected revenue $R(p^\star)$ under $\sigma$ when $\sigma$ is efficient significantly restricts the selling mechanism (due to Proposition 2(b)).

In our constructions below, we ensure ($\mathrm{EA}^{\sigma}_{p}$) and ($\mathrm{AP}^{\sigma}_{=p}$) by combining a posted price with an efficient, strictly *ex-post* budget-balanced collusion scheme. Obtaining ($\mathrm{AP}_{\geq p}$) and ($\mathrm{GU}^{\sigma}$) is more subtle. In particular, it requires careful treatment of non-participation. For example, if nonparticipation by a bidder resulted in the cancellation of the mechanism, ($\mathrm{GU}^{\sigma}$) would fail—thereby enabling collusion sustained by threats of nonparticipation off-path (as in Example 4), as well as creating zero-revenue non-participation equilibria. For ($\mathrm{GU}^{\sigma}$) to hold, nonparticipation by some bidders must allow the mechanism to proceed in a way that does not harm participants. On the other hand, for ($\mathrm{AP}_{\geq p}$) to hold, nonparticipation by some bidders must not allow participants to obtain the good for an aggregate payment below $p$—indeed, otherwise the seller would be vulnerable to efficient collusion involving nonparticipation on-path. To thread the needle between these two considerations, our constructions stipulate that following nonparticipation by some bidders, the mechanisms proceed for the remaining bidders as if the nonparticipants had submitted the prescribed messages of the lowest types.

6.1. **The Symmetric Case.** Suppose that bidders are *symmetric* in that their value distributions are the same distribution $F_i = F$. We combine the $p$-posted-price mechanism with a first-price knockout auction for the ability to purchase the good for $p$. The seller conducts the knockout auction on behalf of the bidders and remits the knockout revenue back to the bidders, giving the winner a share $\alpha$ of the knockout revenue and dividing the remainder equally among the losers. Here, $\alpha$ is a parameter that can lie in $[0, 1/N]$, so each loser receives at least as much of the knockout revenue as the winner. The knockout auction for the case of $\alpha = 0$ was considered by McAfee and McMillan (1992). We call the combined mechanism the *First-Price $\alpha$-Knockout $p$-Posted Price ($FPK_\alpha$-$p$) mechanism.*

Formally, the message spaces of $\text{FPK}_\alpha$-$p$ are $M_i = [0, \bar{b}_i] \cup \{NP_i\}$, where $\bar{b}_i$ is sufficiently large. If all bidders participate—i.e., each bidder $i$ submits a message $b_i \in [0, \bar{b}_i]$—then the good is allocated to the highest bidder (breaking ties uniformly) as long as the highest bid is at least $p$. Thus, the allocation is

$$q_i^p(b) = \begin{cases} 0 & \text{if } b_i < p \text{ or } b_i < \max_{j \in I \setminus \{i\}} b_j \\ 1/|\{j \mid b_j = b_i\}| & \text{otherwise} \end{cases}. \tag{10}$$

The winner then pays his bid, and the knockout revenue (the winner's bid net of the posted price $p$) is remitted to the bidders, resulting in the following net payments:

$$t_i^{\text{FPK}_\alpha\text{-}p}(b) = \underbrace{b_i q_i^p(b)}_{\text{first price}} - \underbrace{\alpha\,(b_i - p) q_i^p(b)}_{\substack{\text{knockout revenue} \\ \text{in case } i \text{ wins}}} - \underbrace{\frac{1-\alpha}{N-1} \sum_{j \in I \setminus \{i\}} (b_j - p) q_j^p(b)}_{\substack{\text{knockout revenue} \\ \text{in case } i \text{ loses}}}.$$

When some bidders do not participate, the mechanism proceeds for participants as if the nonparticipants had bid 0.[44] Thus, for any nonparticipating coalition $S \subseteq I$, the participants' outcomes are

$$\left(q_i^p(NP_S, b_{-S}), t_i^{\text{FPK}_\alpha\text{-}p}(NP_S, b_{-S})\right) = \left(q_i^p(0_S, b_{-S}), t_i^{\text{FPK}_\alpha\text{-}p}(0_S, b_{-S})\right) \text{ for } i \in I \setminus S. \tag{11}$$

The following theorem is our key robustness result for the symmetric case.

**Theorem 2.** *With symmetric bidders, for each $\alpha \in [0, 1/N]$, there is a symmetric profile $\sigma^{FPK_\alpha\text{-}p}$ of strategies that are strictly increasing on a $F$-full-measure set of types, with $\sigma_i^{FPK_\alpha\text{-}p}(\theta_i) \le \theta_i$, such that the mechanism $FPK_\alpha$-$p$ is strongly collusion-robust and almost surely yields ex-post revenue $r(p;\theta)$ under each weakly IR collusion scheme.*

[44] Alternatively, instead of letting the seller keep nonparticipants' shares of the knockout revenue, we could divide them among the participants—thereby making ($\text{AP}_{\ge p}$) hold with equality and strengthening ($\text{GU}^\sigma$), and hence preserving strong collusion-robustness.

By Proposition 4, $\sigma^{\mathrm{FPK}_\alpha\text{-}p}$ must be a Bayes–Nash equilibrium of the FPK$_\alpha$-$p$ mechanism. The proof (whose full details are in the appendix) constructs strategies $\sigma^{\mathrm{FPK}_\alpha\text{-}p}$ explicitly as a candidate Bayes–Nash equilibrium and verifies the conditions of Proposition 6(a). ($\mathrm{EA}_p^\sigma$) holds as strategies $\sigma^{\mathrm{FPK}_\alpha\text{-}p}$ are symmetric and strictly increasing on a $F$-full-measure set of types by construction, ($\mathrm{AP}_{=p}^\sigma$) holds as the knockout auction revenue is distributed among the bidders if all of them participate, and ($\mathrm{AP}_{\geq p}$) holds as the seller keeps the shares of the knockout revenue of any nonparticipating bidders.

Finally, ($\mathrm{GU}^\sigma$) follows from the first-price structure of the knockout component and the sharing of the knockout revenue among bidders. Intuitively, note that as each bidder is only affected by the others' highest bid (due to the first-price structure of the knockout component), it suffices to show that no bidder $i$ can be harmed by another bidder $j$. As the knockout revenue is shared and the equilibrium allocation is efficient given the posted price, a small change in $j$'s bid has a second-order effect on the bidders' total utility. As the change cannot have a first-order impact on $j$ in Bayes–Nash equilibrium, it cannot have a first-order impact on the aggregate utility of bidders other than $j$. By symmetry, it then cannot have a first-order impact on $i$.

To better understand this intuition, we use first-order conditions for a symmetric Bayes–Nash equilibrium to informally derive ($\mathrm{GU}^\sigma$). We focus on the case of FPK$_\alpha$-$p$ for $\alpha = 1/N$—i.e., when the knockout revenue is shared equally among bidders regardless of who won—which, for brevity, we refer to simply as FPK-$p$.[45] This focus is without loss because the games induced by FPK$_\alpha$-$p$ are isomorphic across different $\alpha$,[46] but it simplifies the analysis by separating the remittance of the knockout revenue from the allocation problem. Note that a small decrease in the bid of type $\theta_i$ from bid $b_i = \sigma_i(\theta_i)$ only affects outcomes if bidder $i$ were winning. Conditioning on this event and denoting the reverse hazard rate of one bidder's equilibrium bids by $h(b_i)$,[47] type $\theta_i$'s first-order condition for (symmetric) Bayes–Nash equilibrium in FPK-$p$ becomes

$$-\underbrace{\big(\theta_i - b_i\big)(N-1)h\left(b_i\right)}_{\text{win less often}} + \underbrace{1-\tfrac{1}{N}}_{\text{lower payment}} = 0, \tag{12}$$

as the reverse hazard rate of the highest competing bid is $(N-1)h(b_i)$ and FPK-$p$ remits a fraction $1/N$ of the knockout revenue to each bidder. On the other hand, a

[45] In the case of $p = 0$, FPK-$p$ coincides with the first-price partnership dissolution mechanism of Cramton et al. (1987, Section 6), but with different nonparticipation outcomes. Thus, our analysis implies that their mechanism is (strongly) robust to collusion that does not involve nonparticipation.
[46] Indeed, a bid of $b_i > p$ in FPK$_\alpha$-$p$ corresponds to a bid of $p + N(1-\alpha)(b_i - p)/(N-1)$ in FPK$_{1/N}$-$p$.
[47] The discussion here assumes that $h$ exists, though our Theorem 2 does not rely on this property.

small decrease in the highest bid $y > p$ of bidders other than $i$ only affects $i$ when he loses. If all bidders use the same (almost-surely) increasing strategy $\sigma_i$, the marginal effect of such a decrease on $i$'s (*ex-ante*) expected utility conditioned on him losing is

$$\underbrace{(\theta_i - y)h(y)}_{\text{lose less often}} - \underbrace{\tfrac{1}{N}}_{\text{lower knockout revenue}} \qquad \text{where } \sigma_i(\theta_i) = y, \tag{13}$$

as the reverse hazard rate of $i$'s bid is $h(y)$ and he receives a fraction $1/N$ of the knockout revenue. As (13) is $-1/(N-1)$ times the left-hand side of (12), the first-order condition for Bayes–Nash equilibrium implies an *ex-ante* utility guarantee in FPK-$p$.

We also note that the symmetric Bayes–Nash equilibrium of FPK$_{1/N}$-$p$ coincides with the symmetric Bayes–Nash equilibrium of the standard first-price auction with $N+1$ bidders and reserve $p$, which is given by the strategies[48]

$$\sigma_i^{\text{FPK-}p}(\theta_i) = \theta_i - \frac{\mathbf{1}(\theta_i \geq p)}{F(\theta_i)^N} \int_p^{\theta_i} F(x)^N dx. \tag{14}$$

Intuitively, relative to standard first-price auctions with $N+1$ bidders, the sharing of the knockout revenue in FPK-$p$ lowers the marginal benefit of bid reduction by a fraction $1/N$. On the other hand, the removal of one of the $N$ competing bidders lowers the marginal cost of bid reduction by a fraction $1/N$—making the first-order conditions for Bayes–Nash equilibrium in the two auctions proportional.[49]

6.2. **The General Case.** We next turn to the general case with asymmetric bidders and construct a more complicated mechanism that is strongly collusion-robust. We combine the posted price with a collusion scheme whose side payments correspond to expected externalities when bidders are treated as submitting reports sequentially. (The seller executes this collusion scheme on behalf of the bidders.)

We call the mechanism the *Sequential Expected Externality Knockout $p$-Posted Price (SEEK-$p$) mechanism*. It is direct, so bidder $i$'s message space is $M_i = \Theta_i \cup \{NP_i\}$ and his strategy is $\sigma_i(\theta_i) = \theta_i$. It uses the same allocation rule as the FPK$_\alpha$-$p$ mechanisms. (In particular, this is efficient given the price, so makes (EA$_p^\sigma$) hold.)

[48]Here, we use the convention that $0/0 = 0$.

[49]More precisely, type $\theta_i$'s first-order condition for equilibrium in a $(N+1)$-bidder first-price auction with reserve $p$ conditioning on bidder $i$ winning takes the form

$$-\underbrace{(\theta_i - b_i)Nh(b_i)}_{\text{win less often}} + \underbrace{1}_{\text{lower payment}} = 0,$$

which is proportional to (12).

We begin by describing payments when all bidders participate—i.e., $m = \theta \in \Theta$. Denote the bidders' maximal aggregate surplus net of price $p$ by

$$S^p(\theta) = \max_{q \in \mathcal{Q}} \left\{ \sum_{i \in I} (\theta_i - p) q_i \right\} = \max \left\{ \max_{i \in I} \{\theta_i - p\}, 0 \right\}.$$

Taking an arbitrary order of the bidders, we define the payments $t_i^{\text{SEEK-}p}(\theta)$ so that

$$\underbrace{\theta_i q_i^p(\theta) - t_i^{\text{SEEK-}p}(\theta)}_{\textit{ex-post}\text{ utility}} = \underbrace{\mathbb{E}_{\tilde{\theta}_{\geq i}}\left[ S^p\left(\theta_{<i}, \theta_i, \tilde{\theta}_{>i}\right) - S^p\left(\theta_{<i}, \tilde{\theta}_i, \tilde{\theta}_{>i}\right) \right]}_{\text{internalizing impact from } i \text{ reporting } \theta_i} + \gamma_i^p, \tag{15}$$

where the constants $\gamma_i^p$ are chosen to make $t_i^{\text{SEEK-}p}(0_I) = 0$ for each bidder $i$—i.e.,

$$\gamma_i^p = -\mathbb{E}_{\tilde{\theta}_{\geq i}}\left[ S^p\left(0_{<i}, 0_i, \tilde{\theta}_{>i}\right) - S^p\left(0_{<i}, \tilde{\theta}_i, \tilde{\theta}_{>i}\right) \right]. \tag{16}$$

The intuition for (15) is that bidders are treated as submitting reports sequentially (in an order that can be fixed arbitrarily, or randomized). Each bidder is charged an amount that makes him internalize the impact of his report on aggregate surplus based on the preceding bidders' reported values, in expectation over the succeeding bidders' values. Each bidder $i$ also receives a lump-sum payment $\gamma_i^p$ so that no one pays under the lowest type profile $0_I$.

We specify outcomes following nonparticipation by some bidders analogously to (11): the mechanism proceeds for participants as if the nonparticipants had reported 0. Thus, for any nonparticipating coalition $S \subseteq I$, the participants' outcomes are

$$\left(q_i^p(NP_S, \theta_{-S}), t_i^{\text{SEEK-}p}(NP_S, \theta_{-S})\right) = \left(q_i^p(0_S, \theta_{-S}), t_i^{\text{SEEK-}p}(0_S, \theta_{-S})\right) \text{ for } i \in I \setminus S. \tag{17}$$

The construction of the payments is based on the ideas of Safronov (2018).[50] ($\text{AP}^{\sigma}_{=p}$) can be verified when all bidders participate by adding (15) across bidders using a telescoping sum. (15) also ensures that when bidder $i$ reports truthfully, his *ex-ante* expected utility equals $\gamma_i^p$ for all possible type reports by the other bidders, which verifies ($\text{GU}^{\sigma}$) when all bidders participate. In this logic, the specific choice of the profile of constants $\gamma_i^p$ is not critical—only the sum $\sum_{i \in I} \gamma_i^p$ is relevant.

A key difference from Safronov (2018) lies in ensuring ($\text{AP}_{\geq p}$) and ($\text{GU}^{\sigma}$) when some bidders do not participate. (Safronov's (2018) constructions ensure *ex-ante* or interim IR, but not robustness to collusion on participation.) Here, the choice of constants

[50] To map (15) to the payments of Safronov (2018), consider the problem of efficiently allocating the right to buy the good for $p$, which is worth $\max\{\theta_i - p, 0\}$ to bidder $i$, and suppose that the mechanism does not have to have a nonparticipation message. While Safronov (2018) defined the mechanism using bilateral externality payments between agents, for this problem, the total payments can be expressed according to (15) with a different choice of constants $\gamma_i^p$, namely $\mathbb{E}_{\tilde{\theta}}[(\tilde{\theta}_i - p) q_i^p(\tilde{\theta})]$. See also Csóka (2015) for related ideas.

$\gamma_i^p$ is crucial. Namely, we show that with the lump-sum payments $\gamma_i^p$ constructed in (16), the lowest types $\theta_i = 0$ *never* make positive payments. Hence, when some bidders do not participate, (17) and ($\mathrm{AP}^{\sigma}_{=p}$) ensure that the participants pay at least $p$ per purchased unit in aggregate, making ($\mathrm{AP}_{\geq p}$) hold. (17) also ensures that ($\mathrm{GU}^{\sigma}$) holds if some bidders do not participate.[51]

Thus, with all the conditions of Proposition 6(a) verified, we obtain the following robustness theorem (the full proof is in the appendix).

**Theorem 3.** *The SEEK-p mechanism is strongly collusion-robust and almost surely yields ex-post revenue $r(p;\theta)$ under each weakly IR collusion scheme.*

6.3. **Role of Beliefs.** How do beliefs matter for our strongly collusion-robust mechanisms? From the bidders' perspective, neither the $\mathrm{FPK}_\alpha$-$p$ mechanism nor the SEEK-$p$ mechanism has dominant strategies, and therefore the mechanisms rely on bidders using their beliefs to find a Bayes–Nash equilibrium. This reliance is unavoidable: as shown by Goldberg and Hartline (2005), the only dominant-strategy selling mechanisms that are immune to collusion (even when bidding rings cannot reallocate the good) are those that sell separate allocations of the good to different buyers—which leads to lower expected revenue than an optimal posted price.

However, both the $\mathrm{FPK}_\alpha$-$p$ and SEEK-$p$ mechanisms make it secure for a bidder to participate regardless of his beliefs. That is, both mechanisms are *ex-post* IR—a property not shared by any other collusion-proof mechanisms constructed in the literature (such as those of Laffont and Martimort (1997), Che and Kim (2006, 2009), Pavlov (2008), Safronov (2018), and Csóka et al. (2026)).[52]

**Proposition 7.** *$\mathrm{FPK}_\alpha$-$p$ for $\alpha \in [0, 1/N]$ and SEEK-p are ex-post IR mechanisms.*[53]

The proof is in the online appendix. In the case of $\mathrm{FPK}_\alpha$-$p$, this follows from the property that each type bids at most his value. In the case of SEEK-$p$, this follows from the construction of $\gamma_i^p$: the lowest type receives the lowest *ex-post* utility, and, hence, ensuring the lowest type never pays implies *ex-post* IR.

[51] An alternative approach to ensure robustness to collusion on participation would be to use a construction of Csóka et al. (2026, Section 7.1), whereby the ordering of agents is adjusted after agents submit reports to place all nonparticipating agents first. One advantage of our construction is that it also ensures *ex-post* IR, as we show in the sequel.

[52] If we required that selling mechanisms be *ex-post* IR instead of requiring that they include nonparticipation messages for each bidder, our theorems would persist unchanged.

[53] Formally, we say a selling mechanism $(M, q, t, \sigma)$ is *ex-post individually rational (ex-post IR)* if for all bidders $i$ and type profiles $\theta \in \Theta$, we have $\theta_i q_i(\sigma(\theta)) - t_i(\sigma(\theta)) \geq 0$.

From the seller's perspective, the FPK$_\alpha$-$p$ mechanisms are much less reliant on beliefs, as they are indirect mechanisms that only use knowledge of the value distribution to set the price $p$. This contrasts with the SEEK-$p$ mechanism and other collusion-proof mechanisms in the literature, which rely more strongly on knowing bidders' beliefs to construct the payment rules (e.g., Laffont and Martimort (1997), Che and Kim (2006, 2009), Pavlov (2008), Safronov (2018), and Csóka et al. (2026)).

## 7. Undomination

In this section, we formalize senses in which the mechanisms that we construct are undominated under a range of models of *ex-ante* collusion that may harm bidders.

In light of Proposition 3, to guarantee at least the optimal non-collusive posted-price expected revenue $R(p^\star)$ over all weakly IR collusion, the seller must give up the possibility of obtaining a higher revenue under any form of weakly IR collusion, including if no collusion occurs.[54] In addition, if the optimal posted price $p^\star$ is unique, it follows from Proposition 6 that the *ex-post* revenue is also (almost surely) pinned down to that of the optimal posted price. (The formal proof is in the online appendix.)

**Corollary 1.** *(a) Each selling mechanism that yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes yields expected revenue $R(p^\star)$ under all weakly IR collusion schemes.*

*(b) If the optimal posted price $p^\star$ is unique, then each such selling mechanism almost surely yields ex-post revenue $r(p^\star;\theta)$ under each weakly IR collusion scheme.*

Since weakly IR collusion is extremely permissive, an expected-revenue guarantee over all such collusion is a very strong hypothesis. We now show that similar conclusions obtain if we only require an expected-revenue guarantee over collusion that is IC and IR—i.e., when all bidders are aware of collusion and respond optimally both to it and to punishments.[55] (The proofs are in the online appendix.)

**Proposition 3′.** *(a) If a selling mechanism $(M, q, t, \sigma)$ is collusion-robust, then it almost surely yields the same ex-post revenue under any two IR collusion schemes for it.*

[54]Note that if the seller is only concerned with (*ex-ante* incentive-)efficient collusion by the grand coalition, using the optimal posted price as a reserve in a standard auction (as in Section 3.4) improves the non-collusive revenue without harming the revenue guarantee under collusion. However, as shown in Example 2, such auctions leave the seller vulnerable to inefficient collusion.

[55]We also use Proposition 3′(a) in the proof of Proposition 5.

*(b) Each selling mechanism that yields expected revenue at least $R(p^\star)$ under all IC and IR collusion schemes for it is collusion-robust.*[56]

**Corollary 1′.** *(a) Each selling mechanism that yields expected revenue at least $R(p^\star)$ under all IC and IR collusion schemes yields expected revenue $R(p^\star)$ under all IR collusion schemes.*

*(b) If the optimal posted price $p^\star$ is unique, then each such selling mechanism yields ex-post revenue $r(p^\star;\theta)$ almost surely under each IR collusion scheme.*

As playing any Bayes–Nash equilibrium is a form of IR collusion (Lemma 3), Corollary 1′ implies that an optimal expected-revenue guarantee over IC and IR collusion precludes the seller from obtaining higher revenue in any Bayes–Nash equilibrium. Moreover, it turns out that the same conclusion obtains even if we only require the expected-revenue guarantee to hold for collusion that can be sustained without commitment to punishment—i.e., where bidders simply play a Bayes–Nash equilibrium if any bidder refuses to collude. (The proof is in the online appendix.)

**Definition 10.** A collusion scheme $(\mu, \hat{q}, \hat{t})$ for a selling mechanism $(M, q, t, \sigma)$ is *Nash-reversion individually rational (Nash-IR)* if the mechanism has a Bayes–Nash equilibrium $\hat{\sigma}$ such that $U_i(\hat{q}, \hat{t}) \geq V_i(M, q, t, \hat{\sigma})$ for all bidders $i$.[57,58]

**Proposition 8.** *(a) Each selling mechanism that yields expected revenue at least $R(p^\star)$ under all IC and Nash-IR collusion schemes yields expected revenue $R(p^\star)$ in all Bayes–Nash equilibria.*

*(b) If the optimal posted price $p^\star$ is unique, then each such selling mechanism yields ex-post revenue $r(p^\star;\theta)$ almost surely in each Bayes–Nash equilibrium.*

One intuition for the results of this section comes from the theory of repeated games. The possibility of achieving higher revenue requires tougher competition, but this means that bidders have ways to punish each other, which can in turn sustain revenue-reducing inefficient collusion (due to Lemma 5).[59]

[56]Alternatively, the same conclusion obtains if the mechanism is required to yield the *same* positive expected revenue under all IC and IR collusion schemes.

[57]By abuse of notation, we write $V_i(M, q, t, \hat{\sigma}) = \mathbb{E}_{\tilde{\theta}}\big[\mathbb{E}_{\tilde{m}\sim\hat{\sigma}(\tilde{\theta})}\big[\tilde{\theta}_i q_i(\tilde{m}) - t_i(\tilde{m})\big]\big]$ for bidder $i$'s *ex-ante* expected utility in a potentially mixed-strategy Bayes–Nash equilibrium $\hat{\sigma}$.

[58]It follows from Lemma 3 that every Nash-IR collusion scheme is IR.

[59]By similar logic, Chassang and Ortner (2019) argued that repeated procurement auctions may benefit from a price floor, which deters collusion by limiting bidders' ability to punish each other.

## 8. The Role of Payments to the Bidders

One unusual property of the FPK$_\alpha$-$p$ and SEEK-$p$ mechanisms is that they generally involve payments to losing bidders. In this section, we show that this property is necessary for guaranteeing the optimal non-collusive posted-price expected revenue $R(p^\star)$ under collusion, except in degenerate cases in which this expected revenue can be obtained from selling to one single bidder. For a stronger result, we require the expected-revenue guarantee only over collusion schemes that are *ex-ante* Pareto-improving over the prescribed strategies, which we assume here to comprise a Bayes–Nash equilibrium of the mechanism.

**Proposition 9.** *Suppose that each bidder's value distribution is supported on an interval and admits a positive density, and no IC and IR mechanism that sells to a single bidder yields expected revenue $R(p^\star)$. If a mechanism $(M, q, t, \sigma)$ for which $\sigma$ is a Bayes–Nash equilibrium yields expected revenue at least $R(p^\star)$ under all IC and ex-ante Pareto-improving collusion schemes for it, then there exists a bidder $i$ such that*

$$\Pr_{\tilde{\theta}}\left[q_i(\sigma(\tilde{\theta})) = 0 \text{ and } t_i(\sigma(\tilde{\theta})) < 0\right] > 0. \tag{18}$$

We prove Proposition 9 in the online appendix. For intuition, note that $\sigma$ must be efficient by (the contrapositive of) Lemma 5, so the mechanism must incorporate efficient collusion on the behalf of the bidders. The efficiency gain relative to the random rationing entailed by a posted price generates additional revenue: due to Revenue Equivalence (which applies when value distributions are supported on intervals and admit positive densities), the seller effectively obtains the expected revenue of an auction rather than a posted price. The seller must then return this revenue to the bidders, which requires making payments to some losing bidders in equilibrium.

In the same setting, when there are two bidders with uniformly-distributed values, the following example illustrates a mechanism that is optimal among mechanisms that do not make payments to the bidders in a restricted class.

*Example* 6. There are two bidders whose values are each distributed uniformly between 0 and 1. We consider the worst-case expected revenue over all IC and *ex-ante* Pareto-improving collusion schemes. We optimize this quantity over all mechanisms $(M, q, t, \sigma)$ with $t \geq 0$ that satisfy ($\mathrm{AP}_{\geq p}$) and ($\mathrm{AP}^{\hat{m}}_{=p}$) for some price $p$ (i.e., offer a posted price to the grand coalition if bidders collude efficiently), and for which $\sigma$ is a Bayes–Nash equilibrium. We first solve this problem for a fixed effective price $p$, and then optimize over $p$. Detailed arguments are in Online Appendix G.

We find that maximizing the worst-case expected revenue over a class of mechanisms that satisfy ($\mathrm{AP}_{\geq p}$) and ($\mathrm{AP}^{\hat{m}}_{=p}$) for a given $p \in (0,1)$ is equivalent to maximizing the bidders' total *ex-ante* expected utility over the same class of mechanisms. (Intuitively, this shrinks the set of *ex-ante* Pareto-improving collusion schemes.) In turn, the latter problem is solved by the direct mechanism given by[60]

$$\left(q_i^{p\text{-DNP}}(\theta), t_i^{p\text{-DNP}}(\theta)\right) = \begin{cases} \big(1-\rho(\theta_{-i}), p\big) & \text{if } \theta_i > \theta_{-i} \text{ and } \theta_i \geq \check{p}, \\ \big(\rho(\theta_i), 0\big) & \text{if } \theta_{-i} > \theta_i \geq \check{p}, \\ (1/2, p/2) & \text{if } \theta_i = \theta_{-i} \geq \check{p} \\ (0,0) & \text{if } \theta_i < \check{p}, \end{cases}$$

where $\check{p}$ is the unique solution in $(0,p)$ to $\frac{1-\check{p}^2}{2} = -p\log\check{p}$, and $\rho : [0,1] \to [0,1/2]$ is an increasing rationing function given by

$$\rho(\theta_i) = \begin{cases} \frac{1}{2} - p\frac{\theta_i - 1 - \log\theta_i}{(1-\theta_i)^2} & \text{if } \theta_i \geq \check{p}, \\ 0 & \text{if } \theta_i < \check{p}. \end{cases}$$

This mechanism allocates the good when the higher bidder's value is at least $\check{p}$, in which case the aggregate payment is $p$. If both bidders' values are at least $\check{p}$, then, rather than paying the lower-value bidder, the mechanism gives him a chance of obtaining the good for free—charging the full price $p$ to the higher-value bidder. The lower-value bidder's probability of obtaining the good is given by rationing function $\rho$ evaluated at his type.

The choice of $p$ that maximizes the worst-case expected revenue over all IC and *ex-ante* Pareto-improving collusion schemes is $p \approx 0.63$. This price is higher than the optimal posted price $p^\star = 1/\sqrt{3} \approx 0.58$—reflecting that the inability to pay bidders reduces the seller's ability to deter collusion, making her use a less efficient way of raising revenue. This mechanism yields a worst-case expected revenue of $\approx 0.33$, which is lower than the noncollusive posted-price expected revenue $R(p^\star) \approx 0.38$. □

As *ex-ante* Pareto-improving collusion is the weakest model of collusion that we consider, the necessity of payments to losing bidders in equilibrium persists for mechanisms that guarantee expected revenue $R(p^\star)$ in our models of collusion that may harm some bidders. Since the selection of any Bayes–Nash equilibrium is an IC and IR form of collusion (Lemma 3), payments to losing bidders are necessary in *all* Bayes–Nash equilibria to guarantee expected revenue under all IC and IR collusion.

[60] If either bidder sends a nonparticipation message, then the null outcome is implemented.

In particular, this applies to the prescribed strategies $\sigma$ if expected revenue $R(p^\star)$ is to be guaranteed under all weakly IR collusion (due to Propositions 3(b) and 4). (The formal argument is in the online appendix.)

**Corollary 2.** *Suppose that each bidder's value distribution is supported on an interval and admits a positive density, and no IC and IR mechanism that sells to a single bidder yields expected revenue $R(p^\star)$.*

*(a) If a mechanism $(M, q, t, \sigma)$ yields expected revenue at least $R(p^\star)$ under all IC and IR collusion schemes for it,*[61] *then for each Bayes–Nash equilibrium $\hat{\sigma}$, there exists a bidder $i$ such that*

$$\Pr_{\tilde{\theta}, \tilde{m} \sim \hat{\sigma}(\tilde{\theta})}\left[q_i(\tilde{m}) = 0 \text{ and } t_i(\tilde{m}) < 0\right] > 0.$$

*(b) In particular, if a mechanism $(M, q, t, \sigma)$ yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes for it, then there exists a bidder $i$ such that (18) holds.*

## 9. Discussion

This paper shows that if bidders can collude *ex ante* (i.e., before learning their values or deciding to participate), then the seller cannot guarantee a higher expected revenue than the non-collusive expected revenue of an optimal posted price. That is, *ex-ante* collusion leads to a complete unraveling of competition.

Bidders can overcome their *ex-post* private information and *ex-ante* Pareto improvement constraints to collude efficiently. If only efficient collusion is possible, the seller can guarantee the optimal posted-price expected revenue by posting an optimal price, or running a standard auction with the posted price as a reserve. However, those mechanisms leave the seller vulnerable to inefficient collusion.

Rather than taking a stance on what forms of collusion are reasonable, we construct mechanisms that guarantee the optimal posted-price expected revenue against a broad class of collusion schemes. Our concept of weakly IR collusion provides a unified framework that captures collusion in which bidders who refuse to collude may be punished, and some bidders may not participate in or be aware of collusion. We show that guaranteeing the optimal posted-price revenue against all weakly IR collusion requires *strong collusion-robustness*—i.e., that the bidders' *ex-ante* expected-utility

[61] As the proof shows, the same also holds for IC and Nash-IR collusion schemes.

| class of distributions | cost with $N$ bidders | cost as $N \to \infty$ | references |
|---|---|---|---|
| symmetric, regular | $\left(1 - \frac{1}{N}\right)^N$ | $\frac{1}{e} \approx 36.8\%$ | Dütting et al. (2016)<br>Chawla et al. (2010) |
| symmetric, general | $\frac{1}{2} - \frac{1}{4N-2}$ | $\frac{1}{2} = 50\%$ | Dütting et al. (2016)<br>Hartline (2016) |
| asymmetric, regular | $\approx 61.8\%$ (not tight) | $\approx 61.8\%$ | Jin et al. (2019) |
| asymmetric, general | $1 - \frac{1}{N}$ | 1 | Alaei et al. (2015) |

TABLE 1. Cost of Collusion to the Seller as Worst-Case Fraction of Myerson (1981) Revenue. All bounds are tight except as noted.

profile be robust to the same forms of collusion. We characterize strongly collusion-robust selling mechanisms and construct two that guarantee the optimal posted-price revenue. These mechanisms are max-min optimal for a range of models of collusion.

Conceptually, our main contribution is to delineate the limits of auctions when bidders can collude perfectly, even when they are subject to asymmetric information. In practice, it may be possible to achieve more expected revenue through auctions, but only due to legal or institutional features that limit bidders' ability to collude.

9.1. **Quantifying the Cost of Collusion.** Our results imply that the cost of *ex-ante* collusion is the revenue loss from using a posted price instead of Myerson's (1981) optimal auction. A literature in computer science (surveyed in Hartline (2016, Section 4.4)) has studied this revenue loss, and has shown that in many natural cases, it is at most a constant fraction of the Myerson (1981) revenue in the worst case regardless of the number of bidders. Table 1 summarizes the implications of this literature for the worst-case cost of collusion. All the bounds in the table are tight, except for the case of asymmetric regular distributions with a fixed number of bidders.

Based on our results, the table indicates that when bidders have either symmetric or regular value distributions, strong collusion-robustness can be ensured by sacrificing at most a fraction of the Myerson (1981) revenue that is bounded by $\approx 61.8\%$ regardless of the number $N$ of bidders. Also, in the symmetric (resp. symmetric regular) case, this fraction is increasing with $N$ and bounded by 50% (resp. $\approx 36.8\%$).[62] However, in the case of general asymmetric distributions, the worst-case cost of collusion as a fraction of the Myerson (1981) revenue goes to 1 as $N$ goes to infinity.

[62]By contrast, the revenue loss from running an efficient auction instead of the Myerson (1981) optimal auction is at most $1/N$ fraction of the Myerson (1981) revenue in the symmetric regular case (Bulow and Klemperer, 1996; Hartline, 2016), which goes to 0 as $N$ goes to infinity.

9.2. **Beyond Single-Item Auctions.** Some of our results extend to general allocation problems. For example, it is generally true that IC and *ex-ante* Pareto-improving over non-collusive play do not preclude efficient collusion (Proposition 1). Moreover, if that occurs and the seller optimizes, the seller would obtain the optimal expected revenue from selling to the grand coalition acting as a single buyer (Theorem 1). However, beyond our single-item setting, it is not generally true that efficient collusion is a worst case for the seller: the optimal expected revenue under efficient collusion cannot generally be guaranteed against all IC and IR collusion.

*Example* 7. Suppose that the seller has two identical units for sale. Suppose furthermore there are two bidders, each of whom wants to buy just one unit, with i.i.d. values drawn from a distribution with a continuous density. If the bidders play a Bayes–Nash equilibrium, then the seller's problem separates them, so it is optimal to post the optimal monopoly price for each unit (which gives a strongly collusion-robust mechanism). Furthermore, each Bayes–Nash equilibrium gives a collusion scheme that is IC and IR (by Lemma 3) and whose revenue cannot exceed the optimal expected revenue from separately offering the units to the bidders. Thus, provided that a Bayes–Nash equilibrium exists, "no collusion" constitutes a worst case for the seller. If instead the bidders colluded efficiently and acted as a single buyer, the seller could obtain a higher expected revenue by slightly lowering the price for two units and/or slightly raising the price for one unit (McAfee et al., 1989). □

Building on the current paper and a follow-up paper (Jagadeesan et al., 2026), Jagadeesan and Yang (2026) characterize when efficient collusion is the worst case for the seller, and provide insights into the structure of the selling mechanism with the optimal expected-revenue guarantee both when this is the case and when it is not.

9.3. **Beyond Revenue Maximization.** From a methodological perspective, our concept of strong collusion-robustness may be useful as a desideratum for market design more broadly. For example, Jagadeesan et al. (2026) develops our construction of SEEK-$p$ to provide optimal welfare guarantees under collusion when agents have property rights and can threaten to hold out.

## Appendix A. Omitted Proofs Supporting Theorems 1, 2, and 3

A.1. **Proof of Lemma 1.** Consider the aggregate outcome function $o : M \to [0,1]\times\mathbb{R}$ defined by $o(m) = \big(\sum_{i\in I} q_i(m), \sum_{i\in I} t_i(m)\big)$. Let $A = o(M)$ be the range of $o$, which is assumed to be compact. By Berge's Maximum Theorem, the correspondence

$$\mathcal{W}(\theta) := \arg\max_{(q_\Sigma, t_\Sigma)\in A} \left\{ \Big( \max_{i\in I} \theta_i \Big) q_\Sigma - t_\Sigma \right\}$$

is upper hemicontinuous and has nonempty compact values, hence has a closed graph.

Define a correspondence $\mathcal{E} : \Theta \rightrightarrows M$ by $\mathcal{E}(\theta) = \{m \,|\, o(m) \in \mathcal{W}(\theta)\}$, which is nonempty-valued since $\mathcal{W} : \Theta \rightrightarrows A$ is nonempty-valued. The graph of $\mathcal{E}$ is Borel because $\mathcal{W}$ has a closed (hence Borel) graph and $o$ is measurable. Hence, by the von Neumann–Aumann Measurable Selection Theorem (Aliprantis and Border, 2006, Corollary 18.27), there exists a measurable function $\mu : \Theta \to M$ such that $\mu(\theta) \in \mathcal{E}(\theta)$ almost surely. In particular, we have $o(\mu(\theta)) \in \mathcal{W}(\theta)$ almost surely.

Regarding $\mu$ as a deterministic message function, we can then define collusive allocation and payment rules by allocating the good and charging the payment to the highest bidder (breaking ties uniformly)—i.e.,

$$\big(\hat{q}_i(\theta), \hat{t}_i(\theta)\big) = \begin{cases} 0 & \text{if } \theta_i < \max_{j\in I\setminus\{i\}} \theta_j \\ \left( \frac{\sum_{j\in I} q_j(\mu(\theta))}{|\{j \,|\, \theta_j = \theta_i\}|}, \frac{\sum_{j\in I} t_j(\mu(\theta))}{|\{j \,|\, \theta_j = \theta_i\}|} \right) & \text{otherwise} \end{cases}.$$

The collusion feasibility conditions (1) and (2) hold by construction, with equality in (2). Equality in (2), and the facts that $\hat{q}$ never allocates the good to a bidder without a highest value and that $o(\mu(\theta)) \in \mathcal{W}(\theta)$ almost surely, imply that $(\mu, \hat{q}, \hat{t})$ is efficient.

A.2. **Proof of Proposition 1.** Let $(M, q, t, \sigma)$ be a selling mechanism.

To prove Part (a), let $(\mu, \hat{q}, \hat{t})$ be an efficient collusion scheme. Define auxiliary Vickrey–Clarke–Groves-style side payments based on the externalities exerted through the (collusive) allocation rule $\hat{q}(\theta)$ and the payments $\mathbb{E}_{\tilde{m}\sim\mu(\theta)}[t(\tilde{m})]$ due to the seller under the collusive message function by

$$\hat{t}_i^{\text{side}}(\theta) = -\sum_{j\in I\setminus\{i\}} \Big( \theta_j \hat{q}_j(\theta) - \mathbb{E}_{\tilde{m}\sim\mu(\theta)}\big[t_j(\tilde{m})\big] \Big).$$

Following d'Aspremont and Gérard-Varet (1979), we can then define a payment rule by

$$\hat{t}'_i(\theta) = \mathbb{E}_{\tilde{m}\sim\mu(\theta)}\big[t_i(\tilde{m})\big] + \mathbb{E}_{\tilde{\theta}_{-i}}\big[\hat{t}_i^{\text{side}}(\theta_i, \tilde{\theta}_{-i})\big] - \frac{1}{N-1} \sum_{j\in I\setminus\{i\}} \mathbb{E}_{\tilde{\theta}_{-j}}\big[\hat{t}_j^{\text{side}}(\theta_j, \tilde{\theta}_{-j})\big].$$

By construction, $(\mu, \hat{q}, \hat{t}')$ satisfies the collusion budget-balancedness condition (2) with equality. Hence, we have defined a collusion scheme, which is efficient due to the

efficiency of $(\mu, \hat{q}, \hat{t})$ and equality in (2). It is IC by the arguments of d'Aspremont and Gérard-Varet (1979).[63] By adding budget-balanced lump-sum side payments, we can make it deliver the same *ex-ante* expected utilities as $(\mu, \hat{q}, \hat{t})$.

To prove Part (b), we prove the contrapositive. Let $(\mu, \hat{q}, \hat{t})$ be an inefficient collusion scheme. As efficient collusion schemes exist (Lemma 1) and utility is transferable, $(\mu, \hat{q}, \hat{t})$ is *ex-ante* Pareto-dominated by an efficient collusion scheme, which can be taken to be IC by Part (a). Hence, $(\mu, \hat{q}, \hat{t})$ is not *ex-ante* incentive-efficient.

A.3. **Proof of Theorem 2.** We first show that we can reduce to the case of $\alpha = 1/N$ by exhibiting an isomorphism between the games induced by FPK$_\alpha$-$p$ and FPK$_{1/N}$-$p$. Define a strictly increasing bijection $\Phi_\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by

$$\Phi_\alpha(b_i) = \begin{cases} p + N(1-\alpha)(b_i - p)/(N-1) & \text{if } b_i \geq p \\ b_i & \text{if } b_i < p \end{cases}.$$

Writing $\Phi_\alpha(NP_i) = NP_i$, the outcome at message profile $m$ in FPK$_\alpha$-$p$ is the same as the outcome at message profile $\Phi_\alpha(m)$ in FPK$_{1/N}$-$p$. Hence, given a symmetric strategy profile $\sigma^{\text{FPK}_{1/N}\text{-}p}$ of strategies that are strictly increasing on a $F$-full-measure set of types that makes FPK$_{1/N}$-$p$ (under maximum bid $\Phi_\alpha(\overline{b}_i)$) strongly collusion-robust, $\Phi_\alpha^{-1} \circ \sigma^{\text{FPK}_{1/N}\text{-}p}$ is a symmetric strategy profile of strategies that are strictly increasing on a $F$-full-measure set of types that makes FPK$_\alpha$-$p$ (under maximum bid $\overline{b}_i$) strongly collusion-robust. Moreover, as $\alpha \leq 1/N$, we have $\Phi_\alpha(b_i) \geq b_i$, and hence $\Phi_\alpha^{-1}(b_i) \leq b_i$. In particular, if $\sigma^{\text{FPK}_{1/N}\text{-}p}(\theta_i) \leq \theta_i$, then we have $\Phi_\alpha^{-1}(\sigma^{\text{FPK}_{1/N}\text{-}p}(\theta_i)) \leq \theta_i$. Therefore, it suffices to prove the theorem for $\alpha = 1/N$.

We abbreviate FPK$_{1/N}$-$p$ to FPK-$p$, and consider the strategies $\sigma_i^{\text{FPK-}p}(\theta_i)$ given by (14), which are defined for $\theta_i \in \mathbb{R}_{\geq 0}$. These strategies are continuous and nondecreasing on $\mathbb{R}_{\geq 0}$, and strictly increasing on $\{\theta_i \mid \Pr_{\tilde{\theta}_i}[\tilde{\theta}_i \in (\theta_i - \varepsilon, \theta_i)] > 0 \text{ for all } \varepsilon > 0\}$, which

[63] For completeness, we give the argument. By construction, for all bidders $i$ and types $\theta_i$, we have

$$\mathbb{E}_{\tilde{\theta}_{-i}}[\hat{t}_i'(\theta_i, \tilde{\theta}_{-i})] = \mathbb{E}_{\tilde{\theta}_{-i}}\Big[\sum_{k \in I} \mathbb{E}_{\tilde{m} \sim \mu(\theta_i, \tilde{\theta}_{-i})}[t_k(\tilde{m})] - \sum_{j \in I \setminus \{i\}} \tilde{\theta}_j \hat{q}_j(\theta_i, \tilde{\theta}_{-i})\Big] - \frac{1}{N-1} \sum_{j \in I \setminus \{i\}} \mathbb{E}_{\tilde{\theta}}\big[\hat{t}_j^{\text{side}}(\tilde{\theta})\big].$$

It follows that for all bidders $i$ and $F_i$-almost all types $\theta_i$, for all reports $\theta_i'$, we have

$$\begin{aligned}
&\mathbb{E}_{\tilde{\theta}_{-i}}\big[\theta_i \hat{q}_i(\theta_i, \tilde{\theta}_{-i}) - \hat{t}_i'(\theta_i, \tilde{\theta}_{-i})\big] \\
=&\mathbb{E}_{\tilde{\theta}_{-i}}\Big[\theta_i \hat{q}_i(\theta_i', \tilde{\theta}_{-i}) + \sum_{j \in I \setminus \{i\}} \tilde{\theta}_j \hat{q}_j(\theta_i, \tilde{\theta}_{-i}) - \sum_{k \in I} \mathbb{E}_{\tilde{m} \sim \mu(\theta_i, \tilde{\theta}_{-i})}[t_k(\tilde{m})]\Big] + \frac{1}{N-1} \sum_{j \in I \setminus \{i\}} \mathbb{E}_{\tilde{\theta}}\big[\hat{t}_j^{\text{side}}(\tilde{\theta})\big] \\
\geq&\mathbb{E}_{\tilde{\theta}_{-i}}\Big[\theta_i \hat{q}_i(\theta_i', \tilde{\theta}_{-i}) + \sum_{j \in I \setminus \{i\}} \tilde{\theta}_j \hat{q}_j(\theta_i', \tilde{\theta}_{-i}) - \sum_{k \in I} \mathbb{E}_{\tilde{m} \sim \mu(\theta_i', \tilde{\theta}_{-i})}[t_k(\tilde{m})]\Big] + \frac{1}{N-1} \sum_{j \in I \setminus \{i\}} \mathbb{E}_{\tilde{\theta}}\big[\hat{t}_j^{\text{side}}(\tilde{\theta})\big] \\
=&\mathbb{E}_{\tilde{\theta}_{-i}}\big[\theta_i \hat{q}_i(\theta_i', \tilde{\theta}_{-i}) - \hat{t}_i'(\theta_i', \tilde{\theta}_{-i})\big],
\end{aligned}$$

where the inequality follows from the hypothesis that $(\mu, \hat{q}, \hat{t})$ is an efficient collusion scheme.

has $F$-full measure because $F$ is continuous. They also satisfy $\sigma_i^{\text{FPK-}p}(\theta_i) \le \theta_i$. To complete the proof, we verify the conditions of Proposition 6(a) in turn.

($\text{EA}_p^\sigma$) holds as the strategies $\sigma^{\text{FPK-}p}$ are symmetric, strictly increasing on a $F$-full-measure set of types, and satisfy $\sigma_i^{\text{FPK-}p}(p) = p$. To verify ($\text{AP}_{=p}^\sigma$), note

$$\sum_{i\in I} t_i^{\text{FPK-}p}(b) = p \sum_{i\in I} q_i^p(b) \tag{A.1}$$

holds for all bid profiles $b \in \times_{i\in I}[0, \bar{b}_i]$ by construction. To verify ($\text{AP}_{\ge p}$), note that given any message $m = (NP_S, b_{-S})$, by construction, we have

$$t_i^{\text{FPK-}p}(m) - pq_i^p(m) = 0 \ge t_i^{\text{FPK-}p}(0_S, b_{-S}) = t_i^{\text{FPK-}p}(0_S, b_{-S}) - pq_i^p(0_S, b_{-S})$$

for all $i \in S$. By the construction of participants' outcomes if some bidders do not participate in (11), and using (A.1) for the bid profile $(0_S, b_{-S})$, we have

$$\sum_{i\in I} t_i^{\text{FPK-}p}(m) - p\sum_{i\in I} q_i^p(m) \ge \sum_{i\in I} t_i^{\text{FPK-}p}(0_S, b_{-S}) - p\sum_{i\in I} q_i^p(0_S, b_{-S}) = 0.$$

To verify ($\text{GU}^\sigma$), consider a message profile $m_{-i} = (NP_S, b_{-(S\cup\{i\})}) \in M_{-i}$. Let $z = \max\big(\{p\} \cup \{b_j \,|\, j \in I \setminus \{i\} \setminus S\}\big)$ be the larger of $p$ and the other bidders' highest bid. Let $\xi$ be the largest value in $[0, \max\{p, \bar{\theta}_i\}]$ such that $\sigma_i^{\text{FPK-}p}(\xi) \le z$, which exists since $\sigma_i^{\text{FPK-}p}$ is continuous, and satisfies $\xi \ge p$ since $\sigma_i^{\text{FPK-}p}(p) = p \le z$. Since $\sigma_i^{\text{FPK-}p}$ is nondecreasing, and strictly increasing on a $F$-full-measure set of types, we have $\sigma_i^{\text{FPK-}p}(\theta_i) > z$ if and only if $\theta_i > \xi$ almost surely. If bidder $i$ uses $\sigma_i$ and the others send message profile $m_{-i}$, then $i$'s expected bid shading conditional on winning is

$$\begin{aligned}
&\int_\xi^\infty \big(\theta_i - \sigma_i^{\text{FPK-}p}(\theta_i)\big)\, dF(\theta_i) \\
=&\int_\xi^\infty \frac{1}{F(\theta_i)^N} \int_p^{\theta_i} F(x)^N\, dx\, dF(\theta_i) \\
=&\int_p^\infty \int_{\max\{\xi,x\}}^\infty \frac{F(x)^N}{F(\theta_i)^N}\, dF(\theta_i)\, dx \\
=&\frac{1}{N-1}\int_p^\infty \Big(\frac{F(x)^N}{F(\max\{\xi,x\})^{N-1}} - F(x)^N\Big)\, dx \\
=&\frac{1}{N-1}\Bigg(\int_p^\infty \Big(\frac{F(x)^N}{F(\max\{\xi,x\})^{N-1}} - 1\Big)\, dx + \int_p^\infty \big(1 - F(x)^N\big)\, dx\Bigg) \\
=&\frac{1}{N-1}\Bigg(\int_p^\xi \Big(\frac{F(x)^N}{F(\xi)^{N-1}} - 1\Big)\, dx + \int_\xi^\infty \big(F(x) - 1\big)\, dx + \int_p^\infty \big(1 - F(x)^N\big)\, dx\Bigg),
\end{aligned}$$

where the first equality is by the definition of $\sigma^{\text{FPK-}p}$ since $\xi \ge p$, the second by Fubini's Theorem, the third by evaluating the inner integral, the fourth by decomposing the

integrand, and the fifth by dividing the domain of integration. Evaluating the second and third terms using the identity $\int_y^\infty \Pr_{\tilde{X}}[\tilde{X} \geq x]\,dx = \mathbb{E}_{\tilde{X}}[\max\{\tilde{X} - y, 0\}]$ and the fact that $F(x)^N$ is the distribution of the highest bidder's value, we then have

$$\begin{aligned}
&(N-1)\mathbb{E}_{\tilde{\theta}_i}\Big[\big(\tilde{\theta}_i - \sigma_i^{\text{FPK-}p}(\tilde{\theta}_i)\big)\mathbf{1}\big(\tilde{\theta}_i \geq \xi\big)\Big] \\
=&\big(\xi - \sigma_i^{\text{FPK-}p}(\xi)\big)F(\xi) - (\xi - p) - \mathbb{E}_{\tilde{\theta}_i}\Big[\max\big\{\tilde{\theta}_i - \xi, 0\big\}\Big] + \mathbb{E}_{\tilde{\theta}}\Big[\max\Big\{\max_{j\in I}\big\{\tilde{\theta}_j - p\big\}, 0\Big\}\Big] \\
=&\big(p - \sigma_i^{\text{FPK-}p}(\xi)\big)F(\xi) - (\xi - p)\big(1 - F(\xi)\big) - \mathbb{E}_{\tilde{\theta}_i}\Big[\big(\tilde{\theta}_i - \xi\big)\mathbf{1}\big(\tilde{\theta}_i \geq \xi\big)\Big] \\
&\qquad + \mathbb{E}_{\tilde{\theta}}\Big[\max\Big\{\max_{j\in I}\big\{\tilde{\theta}_j - p\big\}, 0\Big\}\Big] \\
=&\big(p - \sigma_i^{\text{FPK-}p}(\xi)\big)F(\xi) - \mathbb{E}_{\tilde{\theta}_i}\Big[\big(\tilde{\theta}_i - p\big)\mathbf{1}\big(\tilde{\theta}_i \geq \xi\big)\Big] + \mathbb{E}_{\tilde{\theta}}\Big[\max\Big\{\max_{j\in I}\big\{\tilde{\theta}_j - p\big\}, 0\Big\}\Big],
\end{aligned}$$

where the second equality is by subtracting $(\xi - p)F(\xi)$ from each of the first two terms, and the third equality is by combining the second and third terms.

It follows that bidder $i$'s *ex-ante* utility from using $\sigma_i^{\text{FPK-}p}$ in response to $m_{-i}$ is

$$\begin{aligned}
&V_i\big(M, q^p, t^{\text{FPK-}p}, \sigma_i^{\text{FPK-}p}, m_{-i}\big) \\
\geq&\underbrace{\frac{\sigma_i^{\text{FPK-}p}(\xi) - p}{N}F(\xi)}_{\text{losing}} + \underbrace{\mathbb{E}_{\tilde{\theta}_i}\Big[\Big(\tilde{\theta}_i - \sigma_i^{\text{FPK-}p}(\tilde{\theta}_i) + \frac{\sigma_i^{\text{FPK-}p}(\tilde{\theta}_i) - p}{N}\Big)\mathbf{1}\big(\tilde{\theta}_i \geq \xi\big)\Big]}_{\text{winning}} \\
=&\frac{\sigma_i^{\text{FPK-}p}(\xi) - p}{N}F(\xi) + \mathbb{E}_{\tilde{\theta}_i}\Big[\Big(\frac{\tilde{\theta}_i - p}{N} + \frac{N-1}{N}\big(\tilde{\theta}_i - \sigma_i^{\text{FPK-}p}(\tilde{\theta}_i)\big)\Big)\mathbf{1}\big(\tilde{\theta}_i \geq \xi\big)\Big] \\
=&\frac{1}{N}\mathbb{E}_{\tilde{\theta}}\Big[\max\Big\{\max_{j\in I}\big\{\tilde{\theta}_j - p\big\}, 0\Big\}\Big] = V_i\big(M, q^p, t^{\text{FPK-}p}, \sigma^{\text{FPK-}p}\big),
\end{aligned}$$

where the inequality is by the definition of $\xi$; and the last equality by ($\text{EA}_p^\sigma$), ($\text{AP}_{=p}^\sigma$), and symmetry. The theorem then follows by Proposition 6(a).

A.4. **Proof of Theorem 3.** We verify the conditions of Proposition 6(a) in turn.

($\text{EA}_p^\sigma$) holds by construction.

For notation, we let $v_i^{\text{SEEK-}p}(\theta_i, m_{-i}) = \theta_i q_i^p(\theta_i, m_{-i}) - t_i^{\text{SEEK-}p}(\theta_i, m_{-i})$. To verify ($\text{AP}_{=p}^\sigma$), note that the construction of $t^{\text{SEEK-}p}$ in (15) implies

$$\begin{aligned}
\sum_{i\in I} v_i^{\text{SEEK-}p}(\theta) &= \sum_{i\in I}\mathbb{E}_{\tilde{\theta}_{\geq i}}\big[S^p\big(\theta_{<i}, \theta_i, \tilde{\theta}_{>i}\big) - S^p\big(\theta_{<i}, \tilde{\theta}_i, \tilde{\theta}_{>i}\big)\big] + \sum_{i\in I}\gamma_i^p \\
&= S^p(\theta) - \mathbb{E}_{\tilde{\theta}}\big[S^p\big(\tilde{\theta}\big)\big] + \sum_{i\in I}\gamma_i^p = \sum_{i\in I}(\theta_i - p)q_i^p(\theta) - \mathbb{E}_{\tilde{\theta}}\big[S^p\big(\tilde{\theta}\big)\big] + \sum_{i\in I}\gamma_i^p,
\end{aligned}$$

where the second equality uses a telescoping property of the sum involving expectations of differences of $S^p$, and the third is by the efficiency of $q_i^p$ given $p$. It follows that

$$p\sum_{i\in I} q_i^p(\theta) - \sum_{i\in I} t_i^{\text{SEEK-}p}(\theta) = \sum_{i\in I}\left(v_i^{\text{SEEK-}p}(\theta) - (\theta_i - p)q_i^p(\theta)\right) = -\mathbb{E}_{\tilde{\theta}}\left[S^p\left(\tilde{\theta}\right)\right] + \sum_{i\in I}\gamma_i^p.$$

Hence, the left-hand side is constant in $\theta$. By the construction of the allocation rule in (10) and the lump-sum payments $\gamma_i^p$ in (16), it equals 0 for the lowest type profile $\theta = 0_I$. Therefore, it equals 0 for all type profiles $\theta$.

To verify ($\text{AP}_{\geq p}$), note that $S^p(\theta)$ is the maximum of monotone functions of individual components $\theta_i$, and is hence a submodular function on the lattice $\Theta = \bigtimes_{i\in I}[0,\overline{\theta}_i]$ (Topkis, 1998, Example 2.6.2(f)). The construction of $t^{\text{SEEK-}p}$ in (15) then implies

$$-t_i^{\text{SEEK-}p}(0_i,\theta_{-i}) = v_i^{\text{SEEK-}p}(0_i,\theta_{-i}) \geq v_i^{\text{SEEK-}p}(0_I) = 0, \tag{A.2}$$

where the equality is by the construction of the lump-sum payments $\gamma_i^p$ in (16). It follows that given any message $m = (NP_S, \theta_{-S})$, we have

$$t_i^{\text{SEEK-}p}(m) - pq_i^p(m) = 0 \geq t_i^{\text{SEEK-}p}(0_S,\theta_{-S}) = t_i^{\text{SEEK-}p}(0_S,\theta_{-S}) - pq_i^p(0_S,\theta_{-S})$$

for all $i \in S$. Using ($\text{AP}^{\sigma}_{=p}$) for the message $(0_S,\theta_{-S})$ and the construction of participants' outcomes if some bidders do not participate in (17), we have

$$\sum_{i\in I} t_i^{\text{SEEK-}p}(m) - p\sum_{i\in I} q_i^p(m) \geq \sum_{i\in I} t_i^{\text{SEEK-}p}(0_S,\theta_{-S}) - p\sum_{i\in I} q_i^p(0_S,\theta_{-S}) = 0.$$

To verify ($\text{GU}^{\sigma}$), note that for all type profiles $\theta_{-i} \in \Theta_{-i}$, (15) implies

$$\mathbb{E}_{\tilde{\theta}_i}\left[v_i^{\text{SEEK-}p}\left(\tilde{\theta}_i,\theta_{-i}\right)\right] = \mathbb{E}_{\tilde{\theta}_i}\left[\mathbb{E}_{\tilde{\theta}_{>i}}\left[S^p\left(\theta_{<i},\tilde{\theta}_i,\tilde{\theta}_{>i}\right)\right]\right] - \mathbb{E}_{\tilde{\theta}_{\geq i}}\left[S^p\left(\theta_{<i},\tilde{\theta}_i,\tilde{\theta}_{>i}\right)\right] + \gamma_i^p = \gamma_i^p,$$

where the second equality is by the law of iterated expectations. Since, in particular, taking expectations over $\theta_{-i}$ yields $V_i(M,q^p,t^{\text{SEEK-}p},\sigma) = \gamma_i^p$, ($\text{GU}^{\sigma}$) holds with equality for messages $m_{-i} \in \Theta_{-i}$. If bidders $j \in S \subseteq I \setminus \{i\}$ do not participate, then by (17), we have

$$\mathbb{E}_{\tilde{\theta}_i}\left[v_i^{\text{SEEK-}p}\left(\tilde{\theta}_i, NP_S, \theta_{-(S\cup\{i\})}\right)\right] = \mathbb{E}_{\tilde{\theta}_i}\left[v_i^{\text{SEEK-}p}\left(\tilde{\theta}_i, 0_S, \theta_{-(S\cup\{i\})}\right)\right] = V_i\left(M,q^p,t^{\text{SEEK-}p},\sigma\right),$$

so ($\text{GU}^{\sigma}$) still holds. The theorem then follows by Proposition 6(a).

## References

Alaei, S., J. Hartline, R. Niazadeh, E. Pountourakis, and Y. Yuan (2015). Optimal auctions vs. anonymous pricing. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science*, FOCS 2015, pp. 1446–1463. Institute of Electrical and Electronics Engineers.

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Arrow, K. J. (1979). The property rights doctrine and demand revelation under incomplete information. In M. J. Boskin (Ed.), *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, pp. 23–39. Elsevier.

Asker, J. (2010). A study of the internal organization of a bidding cartel. *American Economic Review 100*(3), 724–762.

Bulow, J. and P. Klemperer (1996). Auctions versus negotiations. *American Economic Review 86*(1), 180–194.

Cassady, R. (1967). *Auctions and Auctioneering*. University of California Press.

Chassang, S. and J. Ortner (2019). Collusion in auctions with constrained bids: Theory and evidence from public procurement. *Journal of Political Economy 127*(5), 2269–2300.

Chawla, S., J. D. Hartline, D. L. Malec, and B. Sivan (2010). Multi-parameter mechanism design and sequential posted pricing. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC 2010, pp. 311–320. Association for Computing Machinery.

Che, Y.-K. and J. Kim (2006). Robustly collusion-proof implementation. *Econometrica 74*(4), 1063–1107.

Che, Y.-K. and J. Kim (2009). Optimal collusion-proof auctions. *Journal of Economic Theory 144*(2), 565–603.

Cramton, P., R. Gibbons, and P. Klemperer (1987). Dissolving a partnership efficiently. *Econometrica 55*(3), 615–632.

Csóka, E. (2015). Efficient teamwork. Working paper.

Csóka, E., H. Liu, A. Rodivilov, and A. Teytelboym (2026). Efficiency and collusion-proofness in dynamic mechanism design. Working paper.

Csóka, E., A. Pongrácz, and A. Rodivilov (2025). Guaranteed utility equilibrium. Working paper.

d'Aspremont, C. and L.-A. Gérard-Varet (1979). Incentives and incomplete information. *Journal of Public Economics 11*(1), 25–45.

Dequiedt, V. (2007). Efficient collusion in optimal auctions. *Journal of Economic Theory 136*(1), 302–323.

Dütting, P., F. Fischer, and M. Klimm (2016). Revenue gaps for static and dynamic posted pricing of homogeneous goods. Working paper.

Goldberg, A. V. and J. D. Hartline (2005). Collusion-resistant mechanisms for single-parameter agents. In *Proceedings of the Sixteenth Annual ACM–SIAM Symposium on Discrete Algorithms*, SODA '05, pp. 620–629. Society for Industrial and Applied Mathematics.

Graham, D. A. and R. C. Marshall (1987). Collusive bidder behavior at single-object second-price and English auctions. *Journal of Political Economy 95*(6), 1217–1239.

Hartline, J. D. (2016). *Mechanism Design and Approximation*.

Jagadeesan, R., I. Segal, and A. Tordjman (2026). Robustly efficient implementation under collusion: Holdout threats and the anti-core. Working paper.

Jagadeesan, R. and F. Yang (2026). Multidimensionally screening cartels. In preparation.

Jin, Y., P. Lu, Q. Qi, Z. G. Tang, and T. Xiao (2019). Tight approximation ratio of anonymous pricing. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pp. 674–685. Association for Computing Machinery.

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer.

Laffont, J.-J. and D. Martimort (1997). Collusion under asymmetric information. *Econometrica 65*(4), 875–911.

Laffont, J.-J. and D. Martimort (2000). Mechanism design with collusion and correlation. *Econometrica 68*(2), 309–342.

Mailath, G. J. and P. Zemsky (1991). Collusion in second price auctions with heterogeneous bidders. *Games and Economic Behavior 3*(4), 467–486.

McAfee, R. P. and J. McMillan (1987). Auctions and bidding. *Journal of Economic Literature 25*(2), 699–738.

McAfee, R. P. and J. McMillan (1992). Bidding rings. *American Economic Review 82*(3), 579–599.

McAfee, R. P., J. McMillan, and M. D. Whinston (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *Quarterly Journal of Economics 104*(2), 371–383.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*(2), 265–281.

Pavlov, G. (2008). Auction design in the presence of collusion. *Theoretical Economics 3*(3), 383–429.

Pesendorfer, M. (2000). A study of collusion in first-price auctions. *Review of Economic Studies 67*(3), 381–411.

Quesada, L. (2005). Collusion as an informed principal problem. Working paper.

Riley, J. and R. Zeckhauser (1983). Optimal selling strategies: When to haggle, when to hold firm. *Quarterly Journal of Economics 98*(2), 267–289.

Robinson, M. S. (1985). Collusion and the choice of auction. *RAND Journal of Economics 16*(1), 141–145.

Safronov, M. (2018). Coalition-proof full efficient implementation. *Journal of Economic Theory 177*, 659–677.

Segal, I. and M. Whinston (2012). Property rights. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 100–158. Princeton University Press.

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.

# ONLINE APPENDIX

## Appendix B. Further Omitted Proofs

B.1. **Proof of Proposition 2(b).** Let $(\mu, \hat{q}, \hat{t})$ be an efficient collusion scheme for $(M, q, t, \sigma)$ that yields expected revenue $R(p^\star)$. As the uniqueness of the optimal posted price implies that posting a price of $p^\star$ is the almost-surely unique IC and IR mechanism for selling to a single buyer with type $\theta$ and value $\max_{i\in I}\theta_i$ that yields the optimal expected revenue $R(p^\star)$,[64] it follows from Lemma 2(b) that

$$\sum_{i\in I}\hat{q}_i(\theta) = \mathbf{1}\Big(\max_{i\in I}\theta_i \geq p^\star\Big) \text{ and } \mathbb{E}_{\tilde{m}\sim\mu(\theta)}\Big[\sum_{i\in I}t_i(\tilde{m})\Big] = p^\star\sum_{i\in I}\hat{q}_i(\theta) \text{ almost surely.} \tag{B.1}$$

To prove that ($\mathrm{AP}_{\geq p}$) holds for $p = p^\star$, suppose for the sake of deriving a contradiction that there exists a message profile $m$ such that

$$\varepsilon := p^\star\sum_{i\in I}q_i(m) - \sum_{i\in I}t_i(m) > 0.$$

As $p^\star$ is an optimal posted price, we must have $\Pr_{\tilde{\theta}}\big[(\max_{i\in I}\tilde{\theta}_i) \in [p^\star, p^\star+\varepsilon)\big] > 0$ (or else a posted price of $p^\star+\varepsilon$ would yield higher expected revenue). By Lemma 2(a) and the collusion feasibility conditions (1) and (2), almost surely conditional on $(\max_{i\in I}\theta_i) \in [p^\star, p^\star+\varepsilon)$, we have

$$\begin{aligned}\Big(\max_{i\in I}\theta_i\Big)\sum_{i\in I}\hat{q}_i(\theta) - \mathbb{E}_{\tilde{m}\sim\mu(\theta)}\Big[\sum_{i\in I}t_i(\tilde{m})\Big] &\geq \Big(\max_{i\in I}\theta_i\Big)\sum_{i\in I}q_i(m) - \sum_{i\in I}t_i(m)\\ &\geq p^\star\sum_{i\in I}q_i(m) - \sum_{i\in I}t_i(m) = \varepsilon.\end{aligned}$$

But due to (B.1), almost surely conditional on $(\max_{i\in I}\theta_i) \in [p^\star, p^\star+\varepsilon)$, we also have

$$\Big(\max_{i\in I}\theta_i\Big)\sum_{i\in I}\hat{q}_i(\theta) - \mathbb{E}_{\tilde{m}\sim\mu(\theta)}\Big[\sum_{i\in I}t_i(\tilde{m})\Big] = \Big(\max_{i\in I}\theta_i\Big) - p^\star < \varepsilon$$

—a contradiction. Hence, ($\mathrm{AP}_{\geq p}$) must hold for $p = p^\star$.

Letting $\theta$ be any type profile with $\max_{i\in I}\theta_i \geq p^\star$ for which the equalities in (B.1) hold, the collusion feasibility condition (1) implies that we must have

$$\sum_{i\in I}q_i(\tilde{m}) = 1 \quad \text{almost surely conditional on } \theta \text{ if } \tilde{m}\sim\mu(\theta).$$

Due to ($\mathrm{AP}_{\geq p}$) for $p = p^\star$, we must also have

$$\sum_{i\in I}t_i(\tilde{m}) = p^\star\sum_{i\in I}q_i(\tilde{m}) \quad \text{almost surely conditional on } \theta \text{ if } \tilde{m}\sim\mu(\theta).$$

[64] This assertion holds because the extreme points among monotone allocation rules in the single-buyer problem correspond to the allocation rules of posted prices.

Taking $\hat{m}$ such that both of these conditions hold yields ($\mathrm{AP}^{\hat{m}}_{=p}$) for $p = p^\star$.

B.2. **Proof of Proposition 3.** We prove the two parts via separate arguments.

*Proof of Part (a).* By Jagadeesan et al. (2026, Proposition 2(b)), strong collusion-robustness implies that all weakly IR collusion schemes are efficient. Since $\sigma$ is a weakly IR collusion scheme, it suffices to prove that all efficient collusion schemes for $(M, q, t, \sigma)$ almost surely yield the same *ex-post* revenue, as following lemma shows.[65]

**Lemma 6.** *Each pair of efficient collusion schemes for a given selling mechanism almost surely yield the same ex-post revenue, and almost surely give the same allocation of the good.*

*Proof.* Let $(M, q, t, \sigma)$ be a selling mechanism. Define a function $W : \mathbb{R} \to \mathbb{R}$ by

$$W(\psi) = \max_{\eta \in \Delta(M)} \left\{ \mathbb{E}_{\tilde{m}\sim\eta}\Big[\psi \sum_{i\in I} q_i(\tilde{m}) - \sum_{i\in I} t_i(\tilde{m})\Big] \right\},$$

which is defined due to our compactness assumption on the space of aggregate allocations and transfers. Note that $W$ is a convex function, hence it is differentiable outside a countable set $K \subseteq \mathbb{R}$ (Rockafellar, 1970, Theorem 25.3), and its derivative at $\psi \notin K$ is given by the Envelope Theorem (Milgrom and Segal, 2002, Theorem 1):[66]

$$\mathbb{E}_{\tilde{m}\sim\eta}\Big[\sum_{i\in I} q_i(\tilde{m})\Big] = W'(\psi) \quad \text{for all } \eta \in \arg\max_{\eta\in\Delta(M)} \left\{ \mathbb{E}_{\tilde{m}\sim\eta}\Big[\psi \sum_{i\in I} q_i(\tilde{m}) - \sum_{i\in I} t_i(\tilde{m})\Big] \right\}.$$

Now consider any efficient collusion scheme $(\mu, \hat{q}, \hat{t})$. As the distributions of values are continuous, we have $\max_{i\in I} \theta_i \notin K$ almost surely. It then follows from Lemma 2(a) that if $m \sim \mu(\theta)$, we have, almost surely, that

$$\sum_{i\in I} q_i(m) = W'\Big(\max_{i\in I} \theta_i\Big) \quad \text{and} \quad \sum_{i\in I} t_i(m) = \Big(\max_{i\in I} \theta_i\Big) \sum_{i\in I} q_i(m) - W\Big(\max_{i\in I} \theta_i\Big).$$

In particular, *ex-post* revenue under any efficient collusion scheme is almost surely

$$\Big(\max_{i\in I} \theta_i\Big) W'\Big(\max_{i\in I} \theta_i\Big) - W\Big(\max_{i\in I} \theta_i\Big),$$

and is hence almost surely determined by the function $W$.

Also, efficient collusion almost surely allocates the good a highest-value bidder, who is unique almost surely due to the continuity of the value distributions. Due to

[65] The lemma also shows that all efficient collusion schemes almost surely give the same allocation of the good, which implies that each weakly IR collusion scheme for a strongly collusion-robust mechanism almost surely gives the same allocation of the good as $\sigma$ does (see Footnote 39).

[66] The argument here follows the proof of Börgers (2015, Lemma 2.2).

the collusion allocation feasibility condition (1), it follows that

$$\hat{q}_i(\theta) = \begin{cases} W'(\theta_i) & \text{if } \theta_i = \max_{j \in I} \theta_j \\ 0 & \text{if } \theta_i \neq \max_{j \in I} \theta_j \end{cases} \quad \text{almost surely.}$$

Hence, the allocation $\hat{q}(\theta)$ of any efficient collusion scheme $(\mu, \hat{q}, \hat{t})$ is also determined almost surely by the function $W$. □

*Proof of Part (b).* Suppose that $(M, q, t, \sigma)$ yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes. As every collusion scheme that *ex-ante* Pareto-dominates any weakly IR collusion scheme is itself weakly IR (by taking the same punishment schemes), the contrapositive of Lemma 5 implies every weakly IR collusion scheme must be efficient. By Jagadeesan et al. (2026, Proposition 2(b)), $(M, q, t, \sigma)$ is therefore strongly collusion-robust.

B.3. **Proof of Lemma 5.** Consider any inefficient collusion scheme $(\mu, \hat{q}, \hat{t})$. As efficient collusion schemes exist (Lemma 1) and utility is transferable, there exists an efficient collusion scheme $(\mu', \hat{q}', \hat{t}')$ such that $U_i(\hat{q}', \hat{t}') > U_i(\hat{q}, \hat{t})$ for all bidders $i$. By Proposition 1(a), we can assume that $(\mu', \hat{q}', \hat{t}')$ is IC without loss of generality.

Let $\varepsilon > 0$ be small enough that $(1-\varepsilon)U_i(\hat{q}', \hat{t}') > U_i(\hat{q}, \hat{t})$ for all bidders $i$. Consider the collusion scheme $(\mu'', \hat{q}'', \hat{t}'')$ defined by $\mu''(\theta) = (1-\varepsilon)\mu'(\theta) + \varepsilon\delta_{NP}$ and $(\hat{q}''(\theta), \hat{t}''(\theta)) = \big((1-\varepsilon)\hat{q}'(\theta), (1-\varepsilon)\hat{t}'(\theta)\big)$. By construction, $(\mu'', \hat{q}'', \hat{t}'')$ *ex-ante* Pareto dominates $(\mu, \hat{q}, \hat{t})$. Moreover, as $(\mu', \hat{q}', \hat{t}')$ is IC, so is $(\mu'', \hat{q}'', \hat{t}'')$.

By construction, expected revenue under $(\mu'', \hat{q}'', \hat{t}'')$ is $(1-\varepsilon)$ fraction of the expected revenue under $(\mu', \hat{q}', \hat{t}')$, which, in turn, is at most $R(p^\star)$ by Theorem 1(a). As we have $R(p^\star) > 0$, expected revenue under $(\mu'', \hat{q}'', \hat{t}'')$ is less than $R(p^\star)$.

B.4. **Proof of Proposition 5.** Follows from Lemma 3 and Proposition 3′(a).

B.5. **Proof of Proposition 6(b).** Let $(M, q, t, \sigma)$ be a selling mechanism that yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes. By Proposition 3(b), $(M, q, t, \sigma)$ must be strongly collusion-robust. By Csóka et al. (2025, Proposition C.13), strong collusion-robustness implies ($\mathrm{GU}^\sigma$) holds and $\sigma$ is efficient.

As $\sigma$ is weakly IR, expected revenue under it must be at least $R(p^\star)$. Hence, by Theorem 1(a), expected revenue under $\sigma$ must equal $R(p^\star)$. Proposition 2(b) then implies that ($\mathrm{AP}_{\geq p}$) and ($\mathrm{AP}^{\hat{m}}_{=p}$) must hold for $p = p^\star$.

($\mathrm{AP}_{\geq p}$) for $p = p^\star$ implies that *ex-post* coalitional surplus $\sum_{i \in I} \big(\theta_i \hat{q}_i(\theta) - \hat{t}_i(\theta)\big)$ under the collusion feasibility conditions (1) and (2) is at most $\max\big\{\max_{i \in I}\{\theta_i - p^\star\}, 0\big\}$,

with equality only if $\sum_{i\in I}\hat{t}_i(\theta) = p^\star \sum_{i\in I}\hat{q}_i(\theta)$ and

$$\begin{cases}\hat{q}_i(\theta) = 1 & \text{if } \theta_i > \max_{j\in I\setminus\{i\}}\theta_j \text{ and } \theta_i > p^\star\\ \hat{q}_i(\theta) = 0 & \text{if } \theta_i < \max_{j\in I\setminus\{i\}}\theta_j \text{ or } \theta_i < p^\star\end{cases}.$$

Due to ($\mathrm{AP}^{\hat{m}}_{=p}$) for $p = p^\star$, this bound can be attained by sending the message profile $\hat{m}$ (resp. $NP$) if $\max_{i\in I}\theta_i > p^\star$ (resp. $\max_{i\in I}\theta_i \le p^\star$), allocating any purchased quantity of the good to any highest bidder, and any strictly budget-balanced collusive payments. Hence, the efficiency of $\sigma$ implies ($\mathrm{EA}^\sigma_p$) and ($\mathrm{AP}^\sigma_{=p}$) must hold for $p = p^\star$.

B.6. **Proof of Proposition 7.** For the $\mathrm{FPK}_\alpha$-$p$ mechanisms, as $\sigma_i^{\mathrm{FPK}_\alpha\text{-}p}(\theta_i) \le \theta_i$ holds for all types $\theta_i$ by construction, we have

$$\theta_i q_i^p\big(\sigma^{\mathrm{FPK}_\alpha\text{-}p}(\theta)\big) - t_i^{\mathrm{FPK}_\alpha\text{-}p}\big(\sigma^{\mathrm{FPK}_\alpha\text{-}p}(\theta)\big) \ge \big(\theta_i - \sigma_i^{\mathrm{FPK}_\alpha\text{-}p}(\theta_i)\big)q_i^p\big(\sigma^{\mathrm{FPK}_\alpha\text{-}p}(\theta)\big) \ge 0.$$

For the SEEK-$p$ mechanism, note that as $S^p(\theta)$ is nondecreasing in $\theta_i$, (15) implies

$$v_i^{\text{SEEK-}p}(\theta) \ge v_i^{\text{SEEK-}p}(0_i, \theta_{-i}) \ge 0,$$

where $v_i^{\text{SEEK-}p}$ is defined as in Appendix A.4 and the second inequality is (A.2).

B.7. **Proof of Corollary 1.** Consider a selling mechanism $(M, q, t, \sigma)$ that yields expected revenue at least $R(p^\star)$ under all weakly IR collusion schemes.

*Proof of Part (a).* As efficient collusion schemes exist (Lemma 1) and utility is transferable, there exists an efficient collusion scheme $(\mu, \hat{q}, \hat{t})$ that weakly *ex-ante* Pareto-dominates $\sigma$. By Theorem 1(a), expected revenue under $(\mu, \hat{q}, \hat{t})$ is at most $R(p^\star)$, hence must equal $R(p^\star)$. Proposition 3 then implies that expected revenue under all weakly IR collusion schemes is $R(p^\star)$.

*Proof of Part (b).* Proposition 6(b) implies that ($\mathrm{EA}^\sigma_p$), ($\mathrm{AP}^\sigma_{=p}$), ($\mathrm{AP}_{\ge p}$), and ($\mathrm{GU}^\sigma$) must hold for $p = p^\star$. Part (b) then follows by Proposition 6(a).

B.8. **Proof of Proposition 3′.** The proof is similar to the proof of Proposition 3.

*Proof of Part (a).* By Jagadeesan et al. (2026, Proposition 2(a)), collusion-robustness implies that all IR collusion schemes are efficient. The conclusion follows by Lemma 6.

*Proof of Part (b).* Suppose that $(M, q, t, \sigma)$ yields expected revenue at least $R(p^\star)$ under all IC and IR collusion schemes. As every collusion scheme that *ex-ante* Pareto-dominates any IR collusion scheme is itself IR (by taking the same punishment schemes), the contrapositive of Lemma 5 implies that every IR collusion scheme must

be efficient. By Jagadeesan et al. (2026, Proposition 2(a)), $(M, q, t, \sigma)$ is therefore collusion-robust.

B.9. **Proof of Corollary 1′.** Let $(M, q, t, \sigma)$ be a selling mechanism that yields expected revenue at least $R(p^\star)$ under all IC and IR collusion schemes. Consider any IR collusion scheme $(\mu, \hat{q}, \hat{t})$ (if none exist, then both parts of the corollary are tautologically true). As efficient collusion schemes exist (Lemma 1) and utility is transferable, there exists an efficient collusion scheme $(\mu', \hat{q}', \hat{t}')$ that weakly *ex-ante* Pareto-dominates $(\mu, \hat{q}, \hat{t})$, and hence is IR (by taking the same punishment schemes).

*Proof of Part (a).* By Proposition 1(a), we can assume that $(\mu', \hat{q}', \hat{t}')$ is IC without loss of generality. By Theorem 1(a), expected revenue under $(\mu', \hat{q}', \hat{t}')$ is at most $R(p^\star)$, hence must equal $R(p^\star)$. Proposition 3′ then implies that expected revenue under all IR collusion schemes is $R(p^\star)$.

*Proof of Part (b).* Proposition 2(b) implies that $(\mathrm{AP}^{\hat{m}}_{=p})$ and $(\mathrm{AP}_{\geq p})$ must hold for $p = p^\star$. By Proposition 2(a), *ex-post* revenue under $(\mu', \hat{q}', \hat{t}')$ is then almost surely $r(p^\star; \theta)$. Proposition 3′(a) then implies that *ex-post* revenue under each IR collusion scheme must be $r(p^\star; \theta)$ almost surely.

B.10. **Proof of Proposition 8.** Let $(M, q, t, \sigma)$ be a selling mechanism that yields expected revenue at least $R(p^\star)$ under all IC and Nash-IR collusion schemes. Consider any Bayes–Nash equilibrium $\hat{\sigma}$. As every collusion scheme that weakly *ex-ante* Pareto-dominates $\hat{\sigma}$ is Nash-IR (by reverting to $\hat{\sigma}$ if collusion is rejected), the contrapositive of Lemma 5 implies $\hat{\sigma}$ must be efficient.

*Proof of Part (a).* By Theorem 1(a), expected revenue under $\hat{\sigma}$ is at most $R(p^\star)$. By Lemma 3, $\hat{\sigma}$ defines an IC collusion scheme, which is also Nash-IR (as $\hat{\sigma}$ weakly *ex-ante* Pareto-dominates itself). Hence, expected revenue under $\hat{\sigma}$ must equal $R(p^\star)$.

*Proof of Part (b).* Proposition 2(b) implies that $(\mathrm{AP}^{\hat{m}}_{=p})$ and $(\mathrm{AP}_{\geq p})$ must hold for $p = p^\star$. By Proposition 2(a), *ex-post* revenue under $\hat{\sigma}$ is then almost surely $r(p^\star; \theta)$.

B.11. **Proof of Proposition 9.** For each bidder $i \in I$, write $[\underline{\theta}_i, \overline{\theta}_i]$ for the support of $F_i$. Letting $\underline{\psi} = \max_{i \in I} \underline{\theta}_i$ and $\overline{\psi} = \max_{i \in I} \overline{\theta}_i$, the distribution of $\max_{i \in I} \theta_i$ admits a positive density on $(\underline{\psi}, \overline{\psi}]$.

By the contrapositive of Lemma 5, $\sigma$ must define an efficient collusion scheme. It follows from Lemma 2(b) there is a nondecreasing function $G : [\underline{\psi}, \overline{\psi}] \to [0, 1]$

such that $G(\max_{i\in I}\theta_i) = \sum_{i\in I} q_i(\sigma(\theta))$ almost surely.[67] Since the bidders' value distributions are continuous, we can assume that $G$ is right-continuous without loss of generality. Writing $G(x) = 0$ for $x < \underline{\psi}$ and $G(x) = G(\overline{\psi})$ for $x > \overline{\psi}$, and treating $G$ as a distribution on $\mathbb{R}_{\geq 0} \cup \{\infty\}$, note that the aggregate allocation $\sum_{i\in I} q_i(\sigma(\theta))$ almost surely coincides with the aggregate allocation rule under a random posted price $\tilde{p} \sim G$.

Considering the single-buyer problem with value $\max_{i\in I}\theta_i$ described in Lemma 2(b), it follows from Revenue Equivalence that expected revenue under $\sigma$ is at most $\mathbb{E}_{\tilde{p}}[R(\tilde{p})]$. Since $\sigma$ itself defines an *ex-ante* Pareto-improving collusion scheme, expected revenue under it must be at least $R(p^\star)$. Hence, posted prices $\tilde{p} \sim G$ must be almost-surely optimal and expected revenue under $\sigma$ must be $\mathbb{E}_{\tilde{p}}[R(\tilde{p})] = R(p^\star)$.

We claim that there exists $\varepsilon > 0$ such that $G(\underline{\psi} + \varepsilon) = 0$. Suppose for the sake of deriving a contradiction that $G(\underline{\psi} + \varepsilon) > 0$ for all $\varepsilon > 0$. Then, for each $\varepsilon > 0$, there exists optimal posted price $p \leq \underline{\psi} + \varepsilon$. As the set of optimal prices is closed (due to the continuity of $R$), taking a limit of a subsequence as $\varepsilon \to 0^+$ gives an optimal price $p \leq \underline{\psi}$. We then have $R(p^\star) = R(p) = p$. Letting $i$ be a bidder such that $\underline{\psi} = \underline{\theta}_i$, the (direct version) of the $p$-posted price mechanism is then an (IC and IR) mechanism for selling to $i$ with expected revenue $p = R(p^\star)$—a contradiction.

As efficient collusion requires almost surely not allocating the good to bidders whose values are not the highest, the allocation $q(\sigma(\theta))$ must almost surely coincide with the allocation rule $\hat{q}'(\theta)$ of a second-price auction with random reserve $\tilde{p} \sim G$. Since each bidder $i$'s type distribution has positive density on $[\underline{\theta}_i, \overline{\theta}_i]$, this equality must hold almost everywhere on $\times_{i\in I}[\underline{\theta}_i, \overline{\theta}_i]$.

Now suppose for the sake of deriving a contradiction that for each bidder $i$, we almost surely have that $q_i(\sigma(\theta)) > 0$ or $t_i(\sigma(\theta)) \geq 0$. Since each bidder $i$'s type

[67] Formally, let $B$ be the set of type profiles $\theta$ such that (3) holds for the message function $\mu(\theta) = \sigma(\theta)$, which is ascribed probability 1 by Lemma 2(a). Define a function $G : [\underline{\psi}, \overline{\psi}] \to [0,1]$ by

$$G(x) = \sup\left\{ \sum_{i\in I} q_i(\sigma(\theta)) \,\middle|\, \theta \in B \text{ and } \max_{i\in I}\theta_i \leq x \right\} \quad \text{for } x \in (\underline{\psi}, \overline{\psi}],$$

and let $G(\underline{\psi}) = \inf\{G(x) \,|\, x \in (\underline{\psi}, \overline{\psi}]\}$. By construction, the function $G$ is nondecreasing, and we have $\sum_{i\in I} q_i(\sigma(\theta)) \leq G(\max_{i\in I}\theta_i)$ for all $\theta \in B$. Let $K \subseteq \mathbb{R}$ be the set of points $x \in [\underline{\psi}, \overline{\psi}]$ at which $G$ is discontinuous, which is countable since $G$ is nondecreasing. If $\theta \in B$ is such that $\max_{i\in I}\theta_i \notin K \cup \{\underline{\psi}\}$, then since the distribution of $\max_{i\in I}\theta_i$ has full support on $[\underline{\psi}, \overline{\psi}]$, for each $\varepsilon > 0$, there exists $x < \max_{i\in I}\theta_i$ such that $G(x) > G(\max_{i\in I}\theta_i) - \varepsilon/2$. There then exists $\theta' \in B$ with $\max_{i\in I}\theta'_i < \max_{i\in I}\theta_i$ such that $\sum_{i\in I} q_i(\sigma(\theta')) > G(\max_{i\in I}\theta_i) - \varepsilon$, and it follows from (3) that

$$\sum_{i\in I} q_i(\sigma(\theta)) \geq \sum_{i\in I} q_i(\sigma(\theta')) > G(\max_{i\in I}\theta_i) - \varepsilon.$$

Taking $\varepsilon \to 0^+$ yields $\sum_{i\in I} q_i(\sigma(\theta)) \geq G(\max_{i\in I}\theta_i)$, and hence we have $\sum_{i\in I} q_i(\sigma(\theta)) = G(\max_{i\in I}\theta_i)$ almost surely.

distribution has positive density on $[\underline{\theta}_i, \overline{\theta}_i]$, this property must hold almost everywhere on $\times_{i\in I}[\underline{\theta}_i, \overline{\theta}_i]$. Since $G(\underline{\psi}+\varepsilon) = 0$ and $\underline{\psi} \geq \underline{\theta}_i$, for almost all types $\theta_i \in [\underline{\theta}_i, \min\{\underline{\theta}_i+\varepsilon, \overline{\theta}_i\}]$, we almost surely have $q_i(\sigma(\theta)) = \hat{q}_i'(\theta) = 0$ and hence $t_i(\sigma(\theta)) \geq 0$. In particular, for each bidder $i$ and $\delta > 0$, there exists a positive measure of types $\theta_i \in [\underline{\theta}_i, \underline{\theta}_i+\delta)$ such that

$$\mathbb{E}_{\tilde{\theta}_{-i}}\left[\theta_i q_i(\sigma(\theta_i, \tilde{\theta}_{-i})) - t_i(\sigma(\theta_i, \tilde{\theta}_{-i}))\right] \leq 0.$$

Since $\sigma$ is a Bayes–Nash equilibrium, Revenue Equivalence then implies that expected revenue under $\sigma$ must then be at least that of a second-price auction with random reserve $\tilde{p}$. As expected revenue under $\sigma$ is $\mathbb{E}_{\tilde{p}}[R(\tilde{p})]$, there must be 0 probability that more than one bidder's value is greater than $\tilde{p}$. Hence, letting $\underline{p} = \inf\{\psi \,|\, G(\psi) > 0\}$, there must be 0 probability that more than one bidder's value is greater than $\underline{p}$. Since bidders' values are independent and drawn from continuous distributions, this requires that at most one bidder $i$ have $F_i(\underline{p}) < 1$. Supposing that $F_j(\underline{p}) = 1$ for all $j \in I \setminus \{i\}$, since $q(\sigma(\theta)) = \hat{q}'(\theta)$ almost surely, we have $q_j(\sigma(\theta)) = 0$ almost surely for all $j \in I \setminus \{i\}$. Hence, we have $t_j(\sigma(\theta)) \geq 0$ almost surely for all $j \in I \setminus \{i\}$; since $\sigma$ is a Bayes–Nash equilibrium and each bidder has a nonparticipation message, we also have $\mathbb{E}_{\tilde{\theta}}[t_j(\sigma(\tilde{\theta}))] \leq 0$ for all $j \in I \setminus \{i\}$. Hence, we have $\mathbb{E}_{\tilde{\theta}}[t_j(\sigma(\tilde{\theta}))] = 0$ for all $j \in I \setminus \{i\}$, and so $\mathbb{E}_{\tilde{\theta}}[t_i(\sigma(\tilde{\theta}))] = R(p^\star)$. It follows that the (IC and IR) mechanism for selling to $i$ with allocation rule $\mathbb{E}_{\tilde{\theta}_{-i}}[q_i(\sigma(\theta_i, \tilde{\theta}_{-i}))]$ and payment rule $\mathbb{E}_{\tilde{\theta}_{-i}}[t_i(\sigma(\theta_i, \tilde{\theta}_{-i}))]$ yields expected revenue $R(p^\star)$—a contradiction.

B.12. **Proof of Corollary 2.** We prove the two parts separately.

*Proof of Part (a).* Without loss of generality, by expanding the message spaces if necessary, we can assume that $\hat{\sigma}$ has pure strategies. Every *ex-ante* Pareto-improving collusion scheme for $(M, q, t, \hat{\sigma})$ is a Nash-IR collusion scheme for $(M, q, t, \sigma)$ (by reverting to $\hat{\sigma}$ if collusion is rejected), and hence an IR collusion scheme for $(M, q, t, \sigma)$ due to Lemma 3. The result follows by Proposition 9 applied to $(M, q, t, \hat{\sigma})$.

*Proof of Part (b).* By Proposition 3(b), $(M, q, t, \sigma)$ must be strongly collusion-robust, so $\sigma$ must be a Bayes–Nash equilibrium (by Proposition 4). Since every IR collusion scheme is weakly IR (Lemma 4(b)), the result follows from Part (a).[68]

[68]Alternatively, to deduce Part (b) from Proposition 9, note that since every *ex-ante* Pareto-improving collusion scheme is weakly IR (Footnote 37), Proposition 9 implies that (18) must hold.

## Appendix C. Implicit Collusion via a DSP under a Posted Price

To complement Example 5 from Section 4.4, this appendix provides an example showing that posted-price mechanisms also leave the seller vulnerable to implicit collusion via a DSP that maximizes its own revenue. Even without the standard-auction structure, a form of double marginalization means that the DSP will sometimes not buy at the posted price even if some bidder's value exceeds the posted price. As described in Footnote 31, a subtlety relative to Example 5 is that the posted price structure limits the DSP's ability to punish bidders who refuse to join, making the DSP's revenue-maximization problem have type-dependent outside options.

*Example* 8 (Implicit Collusion via a DSP in a Posted-Price Mechanism). As in Example 5, there are two bidders whose values are each distributed uniformly between 0 and 1. Consider the posted-price mechanism with optimal posted price $p^\star = 1/\sqrt{3}$, whose non-collusive expected revenue is $R(p^\star) \approx 0.38$.

If bidder 2 always sends a participation message, then type $\theta_1$ of bidder 1 can only get the good with probability 1/2. Thus, by committing that bidder 2 will always send a participation message if bidder 1 refuses to bid via the DSP, the DSP can restrict type $\theta_1$ to interim expected utility $\frac{1}{2}\max\{\theta_1 - p^\star, 0\}$.

Given this type-dependent outside-option utility level, the DSP can set a report function $\mu : \Theta \to \Delta(M)$ and an IC direct revelation mechanism $(\hat{q}, \hat{t}) : \Theta \to \mathcal{Q} \times \mathbb{R}^2$ satisfying the allocation feasibility constraint (1), and the interim IR constraint

$$\text{(C.1)}\quad \mathbb{E}_{\tilde{\theta}_{-i}}\big[\theta_i \hat{q}_i(\theta_i, \tilde{\theta}_{-i}) - \hat{t}_i(\theta_i, \tilde{\theta}_{-i})\big] \geq \tfrac{1}{2}\max\{\theta_i - p^\star, 0\} \text{ for all bidders } i \text{ and types } \theta_i.$$

The DSP's expected revenue is $\mathbb{E}_{\tilde{\theta}}\big[\sum_{i\in I} \hat{t}_i(\tilde{\theta}) - \mathbb{E}_{\tilde{m}\sim\mu(\tilde{\theta})}\big[\sum_{i\in I} t_i(\tilde{m})\big]\big]$. Since the seller posts a price of $p^\star$, revenue maximization for the DSP is equivalent to maximizing

$$\mathbb{E}_{\tilde{\theta}}\Big[\sum_{i\in I} \hat{t}_i(\tilde{\theta}) - p^\star \sum_{i\in I} \hat{q}_i(\tilde{\theta})\Big],$$

over all IC mechanisms $(\hat{q}, \hat{t}) : \Theta \to \mathcal{Q} \times \mathbb{R}^2$ that satisfy (C.1).

We will show that revenue maximization by the DSP implies that the good must be sold with probability $\approx 0.59$, instead of $2/3 \approx 0.67$ without collusion. Hence, posting a price leaves the seller vulnerable to implicit collusion via a revenue-maximizing DSP.

First, without loss of generality, we restrict to symmetric mechanisms. Next, by Border's (1991) Theorem, it is equivalent for the DSP to choose an IC reduced-form mechanism $(\hat{\bar{q}}, \hat{\bar{t}}) : \Theta_i \to [0,1] \times \mathbb{R}$ to maximize

$$\text{DSPRev} = 2\int_0^1 \Big(\hat{\bar{t}}(\theta_i) - p^\star \hat{\bar{q}}(\theta_i)\Big)\, d\theta_i$$

subject to the Border feasibility constraints

$$\int_{\theta_i}^{1} \hat{\bar{q}}(\theta_i')d\theta_i' \le \frac{1-\theta_i^2}{2} \tag{C.2}$$

and the IR constraints (C.1) rewritten as

$$\hat{\bar{u}}(\theta_i) \ge \tfrac{1}{2}\max\{\theta_i - p^\star, 0\}, \tag{C.3}$$

where we write $\hat{\bar{u}}(\theta_i) = \theta_i\hat{\bar{q}}(\theta_i) - \hat{\bar{t}}(\theta_i)$ for the indirect (interim) utility of type $\theta_i$.[69]

As is standard, we can eliminate transfers and write

$$\frac{\text{DSPRev}}{2} = \int_0^1 \left((\theta_i - p^\star)\hat{\bar{q}}(\theta_i) - \hat{\bar{u}}(\theta_i)\right) d\theta_i.$$

Breaking the integral into components at the DSP's optimal reserve $\hat{p} = \frac{1+p^\star}{2}$ yields

$$\frac{\text{DSPRev}}{2} = \int_{\hat{p}}^1 (\theta_i - p^\star)\hat{\bar{q}}(\theta_i)\, d\theta_i - \int_{\hat{p}}^1 \hat{\bar{u}}(\theta_i)\, d\theta_i + \int_0^{\hat{p}} (\theta_i - p^\star)\hat{\bar{q}}(\theta_i)\, d\theta_i - \int_0^{\hat{p}} \hat{\bar{u}}(\theta_i)\, d\theta_i.$$

Integrating the second and third summands by parts using the envelope formula, as in Myerson (1981) and Manelli and Vincent (2006), respectively, we have

$$\begin{aligned}\frac{\text{DSPRev}}{2} = \int_{\hat{p}}^1 (\theta_i - p^\star)\hat{\bar{q}}(\theta_i)\, d\theta_i + \left(\int_{\hat{p}}^1 (\theta_i - 1)\hat{\bar{q}}(\theta_i)\, d\theta_i - (1-\hat{p})\hat{\bar{u}}(\hat{p})\right) \\ + \left(-\int_0^{\hat{p}} \hat{\bar{u}}(\theta_i)\, d\theta_i + (\hat{p} - p^\star)\hat{\bar{u}}(\hat{p}) + p^\star\hat{\bar{u}}(0)\right) - \int_0^{\hat{p}} \hat{\bar{u}}(\theta_i)\, d\theta_i.\end{aligned}$$

As $1-\hat{p} = \hat{p} - p^\star$ by the definition of $\hat{p}$, the $\hat{\bar{u}}(\hat{p})$ terms cancel, so grouping terms gives

$$\frac{\text{DSPRev}}{2} = \int_{\hat{p}}^1 (2\theta_i - 1 - p^\star)\hat{\bar{q}}(\theta_i)\, d\theta_i - 2\int_0^{\hat{p}} \hat{\bar{u}}(\theta_i)\, d\theta_i + p^\star\hat{\bar{u}}(0).$$

Integrating the first term by parts and using $p^\star\hat{\bar{u}}(0) = \int_0^{p^\star} \hat{\bar{u}}(0)d\theta_i$ then yields

$$\frac{\text{DSPRev}}{2} = 2\int_{\hat{p}}^1 \int_{\theta_i}^1 \hat{\bar{q}}(\theta_i')\, d\theta_i'\, d\theta_i - 2\int_{p^\star}^{\hat{p}} \hat{\bar{u}}(\theta_i)\, d\theta_i - 2\int_0^{p^\star} \left(\hat{\bar{u}}(\theta_i) - \hat{\bar{u}}(0)\right) d\theta_i - p^\star\hat{\bar{u}}(0).$$

Using (C.2) to bound the first term, (C.3) to bound the second and the last term, and the monotonicity of $\hat{\bar{u}}$ to bound the third, it follows that

$$\frac{\text{DSPRev}}{2} \le \int_{\hat{p}}^1 (1-\theta_i^2)\, d\theta_i - \int_{p^\star}^{\hat{p}} (\theta_i - p^\star)\, d\theta_i - 0 - 0,$$

[69] This mechanism design problem has a similar structure to problems studied by Jullien (2000) and Dworczak and Muir (2025). However, due to the addition of a Border (1991) feasibility constraint, we cannot apply their results. Liu (2025) provided an analysis that allows such constraints, which is applicable here. In our example, we can instead solve the mechanism directly by exploiting the uniform distribution, which means ironing is not needed even with our IR constraint.

with equality if and only if (C.2) binds for almost all $\theta_i > \hat{p}$, (C.3) binds for almost all $\theta_i < \hat{p}$, and $\hat{\bar{u}}(0) = 0$. It follows that an IC reduced-form mechanism maximizes the DSP's expected revenue if and only if $\hat{\bar{t}}(0) = 0$ and

$$\hat{\bar{q}}(\theta_i) = \begin{cases} 0 & \text{if } \theta_i < p^\star \\ \frac{1}{2} & \text{if } p^\star < \theta_i < \hat{p} \\ \theta_i & \text{if } \theta_i > \hat{p} \end{cases}.$$

Intuitively, the DSP allocates the good efficiently to types above its optimal reserve $\hat{p}$, and rations types below it to satisfy the IR constraint with equality.[70] Hence, the good is allocated with probability $2\int_0^1 \hat{\bar{q}}(\theta_i)\,d\theta_i = (7-2\sqrt{3})/6 \approx 0.59$, rather than $2/3 \approx 0.67$ without collusion. Thus, implicit collusion via a DSP that maximizes its own revenue lowers the seller's expected revenue to $(7\sqrt{3}-6)/18 \approx 0.34 < 0.38 \approx R(p^\star)$.

## Appendix D. A Comparative Static for the Optimal Posted Price

This appendix shows that the optimal posted price is strictly increasing in the number of bidders in the symmetric case—extending a result of McAfee and McMillan (1987) to allow for an irregular distribution of values.

**Proposition 10.** *If each bidder's value is drawn from a distribution $F$ supported on an interval $[\underline{\phi}, \overline{\phi}] \subseteq \mathbb{R}_{\geq 0}$ with a continuous density $f$ that is positive on $(\underline{\phi}, \overline{\phi})$, then any selection of optimal posted prices $p_N^\star$ for $N \geq 1$ bidders is strictly increasing in $N$.*

*Proof.* Writing $R(p;N) = p(1-F(p)^N)$, we have $p_N^\star \in \arg\max_p R(p;N)$ by construction. First, note that $p_N^\star \in (\underline{\phi}, \overline{\phi})$ for all $N > 1$. Indeed, this follows from the facts that $R(\overline{\phi};N) = 0$, that $R(\underline{\phi};N) = \underline{\phi} \geq 0$, and that $\frac{\partial R(p;N)}{\partial p}\big|_{p=\underline{\phi}} = 1$ (for $N > 1$). Second, note that for $\underline{\phi} < p < \overline{\phi}$ and $N > 1$, we have

$$\frac{\partial^2 \log R(p;N)}{\partial N \partial p} = \frac{f(p)}{F(p)} \frac{x(x-1-\log x)}{(1-x)^2}\bigg|_{x=F(p)^N} > 0.$$

Thus, $\log R(p;N)$ has increasing marginal returns for $\underline{\phi} < p < \overline{\phi}$ in the sense of Edlin and Shannon (1998). The result follows by Edlin and Shannon (1998, Theorem 1). □

[70] An example of a revenue-maximizing allocation rule for the DSP is to allocate the good to the highest bidder if his value exceeds $\hat{p}$; and if no bidder's value exceeds $\hat{p}$, randomly allocate the good among bidders whose values exceed $p^\star$ with probability $1/(p^\star + \hat{p})$, and not allocate the good at all with complementary probability.

The intuition for this result is that (efficient) collusion effectively destroys competition between the bidders, forcing the seller to raise revenue in a less efficient way. It follows that the posted price exceeds the Myerson (1981) optimal reserve.

**Corollary 3** (McAfee and McMillan, 1987)**.** *Under the hypotheses of Proposition 10, if $F$ is regular and $N > 1$, then any optimal posted price is strictly higher than the Myerson (1981) optimal reserve.*

*Proof.* As shown by Myerson (1981), the optimal reserve is the optimal posted price for $N = 1$. Hence, the result follows from Proposition 10. □

## Appendix E. *Ex-Post* Feasibility Conditions on Collusion Schemes

In this appendix, we expand on Footnote 20 and show that any reallocation satisfying (1) and (2) can be implemented in a way that is feasible for any realization of messages $\tilde{m}$. To formalize this assertion, we explicitly allow outcomes in the collusion scheme to condition on realized messages $m$; we denote message-contingent collusive allocation and payment rules by $\hat{\mathfrak{q}}(\theta, m)$ and $\hat{\mathfrak{t}}(\theta, m)$, respectively.

**Lemma 7.** *If $(\mu, \hat{q}, \hat{t})$ is a collusion scheme for a selling mechanism $(M, q, t, \sigma)$, then there exist (bounded, measurable) functions $(\hat{\mathfrak{q}}, \hat{\mathfrak{t}}) : \Theta \times M \to \mathcal{Q} \times \mathbb{R}^I$ such that for all type profiles $\theta \in \Theta$ and message profiles $m \in M$, we have*

$$\text{(E.1)} \quad \mathbb{E}_{\tilde{m} \sim \mu(\theta)}[\hat{\mathfrak{q}}(\theta, \tilde{m})] = \hat{q}(\theta) \quad \textit{(equivalent allocation in expectation over } \tilde{m} \sim \mu(\theta)\textit{)}$$

$$\text{(E.2)} \quad \mathbb{E}_{\tilde{m} \sim \mu(\theta)}[\hat{\mathfrak{t}}(\theta, \tilde{m})] = \hat{t}(\theta) \quad \textit{(equivalent payments in expectation over } \tilde{m} \sim \mu(\theta)\textit{)}$$

$$\text{(E.3)} \quad \sum_{i \in I} \hat{\mathfrak{q}}_i(\theta, m) = \sum_{i \in I} q_i(m) \quad \textit{(feasibility of reallocation of goods for each } m\textit{)}$$

$$\text{(E.4)} \quad \sum_{i \in I} \hat{\mathfrak{t}}_i(\theta, m) \geq \sum_{i \in I} t_i(m) \quad \textit{(weak budget-balancedness of side payments for each } m\textit{)}.$$

For interpretation, given a collusion scheme $(\mu, \hat{q}, \hat{t})$, (E.1) and (E.2) require that the message-contingent allocation and payment rules $\hat{\mathfrak{q}}(\theta, m)$ and $\hat{\mathfrak{t}}(\theta, m)$ in fact induce expected allocation $\hat{q}(\theta)$ and payments $\hat{t}(\theta)$ (conditional on $\theta$) when messages are drawn according to the message function $\mu$. (E.3) and (E.4) require that the message-contingent allocation and payment rules be feasible for each message $m$.

*Proof of Lemma 7.* To construct $\hat{\mathfrak{q}}$, define a (measurable) function $\beta^q : \Theta \to \mathbb{R}^I$ by

$$\beta_i^q(\theta) = \begin{cases} \hat{q}_i(\theta) \big/ \left( \sum_{j \in I} \hat{q}_j(\theta) \right) & \text{if } \sum_{j \in I} \hat{q}_j(\theta) \neq 0 \\ 1/N & \text{if } \sum_{j \in I} \hat{q}_j(\theta) = 0 \end{cases}.$$

Note that $\sum_{i\in I}\beta_i^q(\theta)=1$ by construction. Hence, we can let $\hat{\mathfrak{q}}(\theta,m)=\beta^q(\theta)\sum_{i\in I}q_i(m)$, so (E.1) follows from (1), and (E.3) holds by construction.

To construct $\hat{\mathfrak{t}}$, let $\beta_i^t(\theta)=\hat{t}_i(\theta)-\mathbb{E}_{\tilde{m}\sim\mu(\theta)}[t_i(\tilde{m})]$. Defining $\hat{\mathfrak{t}}_i(\theta,m)=t_i(m)+\beta_i^t(\theta)$, (E.2) holds by construction. (2) implies that $\sum_{i\in I}\beta_i^t(\theta)\geq 0$, and (E.4) follows. □

## Appendix F. IR for Aware Non-Colluders

This appendix formally shows that if bidders who do not participate in collusion play best responses to collusion, they achieve IR.

We first define collusion schemes for a ring $R\subseteq I$: such schemes correspond to collusion schemes in which non-colluding bidders $i\in I\setminus R$ play strategies $\hat{\sigma}_i$, and the reallocation and side payments do not involve the non-colluding bidders. Thus, the definition extends the definition of $(I\setminus R)$-unaware collusion by allowing bidders $i\in I\setminus R$ to play any strategies $\hat{\sigma}_i$ rather than the prescribed strategy $\sigma_i$.

**Definition 11.** A collusion scheme $(\mu,\hat{q},\hat{t})$ for a selling mechanism $(M,q,t,\sigma)$ is a *collusion scheme for ring* $R\subseteq I$ if there exist $\mu^R:\Theta_R\to\Delta(M_R)$ and strategies $\hat{\sigma}_i:\Theta_i\to\Delta(M_i)$ for non-ring bidders $i\in I\setminus R$ such that for all type profiles $\theta\in\Theta$, we have $\mu(\theta)=\mu^R(\theta_R)\times\hat{\sigma}_{-R}(\theta_{-R})$ and

$$\big(\hat{q}_i(\theta),\hat{t}_i(\theta)\big)=\mathbb{E}_{\tilde{m}\sim\mu(\theta)}\big[\big(q_i(\tilde{m}),t_i(\tilde{m})\big)\big]\quad\text{for all } i\in I\setminus R. \tag{F.1}$$

We can then show that if the strategy $\hat{\sigma}_i$ is a best response for a non-colluding bidder $i\in I\setminus R$, then the collusion scheme is IR for $i$. Indeed, the best-response property ensures that $i$ obtains his highest expected utility against the punishment scheme given by other bidders following the collusion scheme. For notation, we denote bidder $i$'s *ex-ante* utility from playing a (measurable) mixed strategy $\hat{\sigma}_i:\Theta_i\to\Delta(M_i)$ in response to a punishment scheme $\nu^i:\Theta_{-i}\to\Delta(M_{-i})$ by

$$V_i(M,q,t,\hat{\sigma}_i,\nu^i)=\mathbb{E}_{\tilde{\theta}}\big[\mathbb{E}_{\tilde{m}_i\sim\hat{\sigma}_i(\tilde{\theta}_i)}\big[\mathbb{E}_{\tilde{m}_{-i}\sim\nu^i(\tilde{\theta}_{-i})}\big[\tilde{\theta}_i q_i(\tilde{m})-t_i\big(\tilde{m}\big)\big]\big]\big].$$

**Lemma 8.** *Let $(\mu,\hat{q},\hat{t})$ be a collusion scheme for a ring $R\subseteq I$ for a selling mechanism $(M,q,t,\sigma)$, and let $\mu^R$ and $\hat{\sigma}_{-R}$ be as in Definition 11.*

*(a) Let $i\in I\setminus R$, and write $\nu^i(\theta_{-i})=\mu^R(\theta_R)\times\hat{\sigma}_{-(R\cup\{i\})}(\theta_{-(R\cup\{i\})})$. If*

$$V_i(M,q,t,\hat{\sigma}_i,\nu^i)\geq V_i(M,q,t,\sigma_i',\nu^i)\ \text{for all (measurable) strategies } \sigma_i':\Theta_i\to M_i, \tag{F.2}$$

*then $(\mu,\hat{q},\hat{t})$ is IR for $i$.*

*(b) In particular, if $(\mu,\hat{q},\hat{t})$ is IR for each $i\in R$ and (F.2) holds for each $i\in I\setminus R$, then $(\mu,\hat{q},\hat{t})$ is an IR collusion scheme.*

Here, (F.2) requires that $\hat{\sigma}_i$ be a best response for the non-colluding bidder $i$ to the collusion by members of $R$ and the strategies $\hat{\sigma}_j$ of the other non-colluding bidders $j$. (In particular, the case of Lemma 8(b) for $R = \varnothing$ recovers Lemma 3.)

*Proof of Lemma 8.* To prove Part (a), note (F.1) implies $U_i(\hat{q}, \hat{t}) = V_i(M, q, t, \hat{\sigma}_i, \nu^i)$. Hence, taking $\nu^i$ to be the punishment scheme for bidder $i$, (F.2) implies that (5) holds for all strategies $\sigma_i'$. Part (b) is immediate from Part (a). □

Note that in Lemma 8(b), we implicitly allowed punishment schemes for a defecting ring member to involve non-ring members. While this property may be too permissive for rings smaller than the grand coalition, it means that ensuring robustness to IR collusion also ensures robustness to collusion by arbitrary rings that can commit to punish defectors—provided that all bidders are aware of collusion.

## Appendix G. Supporting Details for Example 6

*Reduction to Welfare Maximization for Each $p$.* First we argue that for any two mechanisms $(M, q, t, \sigma)$ and $(M', q', t', \sigma')$ satisfying ($\text{AP}_{\geq p}$) and ($\text{AP}^{\hat{m}}_{=p}$) with the same $p$ such that $\sum_{i \in I} V_i(M, q, t, \sigma) \geq \sum_{i \in I} V_i(M', q', t', \sigma')$, for each IC and *ex-ante* Pareto improving collusion scheme $(\mu, \hat{q}, \hat{t})$ for $(M, q, t, \sigma)$, there exists an IC and *ex-ante* Pareto improving collusion scheme $(\mu', \hat{q}', \hat{t}')$ for $(M', q', t', \sigma')$ that delivers a weakly lower expected revenue. Indeed, we can let $\hat{q}' = \hat{q}$ and let $\mu'$ buy the good for $p$ with the appropriate probability—i.e.,

$$\mu'(\theta) = \Big(\sum_{i \in I} \hat{q}_i(\theta)\Big)\delta_{\hat{m}} + \Big[1 - \Big(\sum_{i \in I} \hat{q}_i(\theta)\Big)\Big]\delta_{NP},$$

where $\hat{m}$ is as in ($\text{AP}^{\hat{m}}_{=p}$) for $(M', q', t', \sigma')$—which satisfies (2) and gives a weakly lower revenue due to ($\text{AP}_{\geq p}$) for $(M, q, t, \sigma)$. This gives an IC collusion scheme $(\mu', \hat{q}, \hat{t})$ that yields the same bidder *ex-ante* utility profile as $(\mu, \hat{q}, \hat{t})$, and since $(\mu, \hat{q}, \hat{t})$ is *ex-ante* Pareto-improving for $(M, q, t, \sigma)$ and $\sum_{i \in I} V_i(M, q, t, \sigma) \geq \sum_{i \in I} V_i(M', q', t', \sigma')$, the collusion scheme $(\mu', \hat{q}, \hat{t})$ can be changed into an *ex-ante* Pareto-improving collusion scheme for $(M', q', t', \sigma')$ by adding appropriate lump-sum side payments. This implies that the worst-case revenue of mechanism $(M, q, t, \sigma)$ over all IC and *ex-ante* Pareto-improving collusion schemes is at least as high as that of $(M', q', t', \sigma')$. In particular, an optimal worst-case revenue is obtained by maximizing bidders' total *ex-ante* expected utility over the class of mechanisms under consideration.

*Details of the Construction of* $(M, q^{p\text{-}DNP}, t^{p\text{-}DNP}, \sigma)$. Fix a price $p \in (0,1)$.[71] To see that $\check{p}$ is well-defined in $(0,p)$, note that the function $\zeta(x) = \frac{1-x^2}{2} + p\log x$ approaches $-\infty$ as $x \to 0^+$. As $\log x < x-1$ for all $x \in (0,1)$, we have

$$\zeta(p) = \frac{1-p^2}{2} + p - 1 - \int_p^1 \log x\, dx > -\frac{(1-p)^2}{2} + \int_p^1 (1-x)\, dx = 0.$$

By the Intermediate Value Theorem, $\zeta(p)$ has a zero in $(0,p)$. To see that it is unique, note that $\zeta$ is strictly concave and $\zeta(1) = 0$; hence, $\zeta$ has at most one zero in $(0,1)$.

The constructed mechanism $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ satisfies $t \geq 0$ and ($\mathrm{AP}_{\geq p}$) by construction and ($\mathrm{AP}^{\hat{m}}_{=p}$) for each message profile $\hat{m} \in [\check{p}, 1]^2$. To see that truthtelling is a Bayes–Nash equilibrium, note that the mechanism's interim allocation rule is

$$\begin{aligned}\overline{q}_i^{p\text{-DNP}}(\theta_i) = \int_0^1 q_i^{p\text{-DNP}}(\theta_i, \tilde{\theta}_{-i})\, d\tilde{\theta}_{-i} &= \mathbf{1}(\theta_i \geq \check{p}) \int_0^{\theta_i} \big(1 - \rho(\tilde{\theta}_{-i})\big)\, d\tilde{\theta}_{-i} + (1-\theta_i)\rho(\theta_i) \\ &= p\big(1 + \log(\theta_i/\check{p})\big)\mathbf{1}(\theta_i \geq \check{p}),\end{aligned}$$

which is nondecreasing, and the interim payment rule satisfies the envelope formula.

*Optimality of* $(M, q^{p\text{-}DNP}, t^{p\text{-}DNP}, \sigma)$ *for a Given* $p$. Consider any mechanism $(M, q, t, \sigma)$ with $t \geq 0$ that satisfies ($\mathrm{AP}_{\geq p}$) for which $\sigma$ is a Bayes–Nash equilibrium. (Here, we show optimality even when ($\mathrm{AP}^{\hat{m}}_{=p}$) is dropped.) By the Revelation Principle, we can let $M_i = \Theta_i \cup \{NP_i\}$ and $\sigma_i(\theta_i) = \theta_i$. Let $\overline{u}_i(\theta_i) = \int_0^1 \big(\theta_i q_i(\theta_i, \tilde{\theta}_{-i}) - t_i(\theta_i, \tilde{\theta}_{-i})\big) d\tilde{\theta}_{-i}$ denote the interim expected utility of type $\theta_i$ of agent $i$ under truthtelling.

We bound $\sum_{i\in I} V_i(M, q, t, \sigma)$ from above by a Lagrangian constructed as follows. The construction of multipliers relies on a threshold $c = \frac{(1-p)\check{p}^2}{p-\check{p}^2}$. Since $\check{p} < p < 1$, we have $c > 0$, and as $\frac{(1-p)\check{p}^2}{p-\check{p}^2}$ is decreasing in $p$, we also have $c < \frac{(1-\check{p})\check{p}^2}{\check{p}-\check{p}^2} = \check{p} \leq p < 1$.

- Since truthtelling is a Bayes–Nash equilibrium, the Envelope Theorem implies that for almost all $\theta_i \in [0,1]$, we have

$$\overline{u}_i(\theta_i) - \limsup_{\theta_i' \to 1^-} \overline{u}_i(\theta_i') = -\int_{\theta_i}^1 \int_0^1 q_i(\theta_i', \tilde{\theta}_{-i})\, d\tilde{\theta}_{-i}\, d\theta_i'$$

  We place a multiplier of $\lambda(\theta_i) = \frac{c-\theta_i^2}{2\theta_i^2}\mathbf{1}(\theta_i \geq c)$ on this constraint.

[71] If $p = 0$ or $p \geq 1$, then any mechanism that satisfies ($\mathrm{AP}_{\geq p}$) and ($\mathrm{AP}^{\hat{m}}_{=p}$) yields $R(p) = 0$ expected revenue under *ex-ante* incentive-efficient collusion (by Propositions 1(b) and 2(a)).

As $\int_c^1 \lambda(\theta_i)d\theta_i = 0$, writing $\Lambda(\theta_i) = \int_c^{\theta_i} \lambda(\theta_i')d\theta_i' = \frac{(\theta_i - c)(1-\theta_i)}{2\theta_i}\mathbf{1}(\theta_i \geq c)$,[72] the corresponding terms in the Lagrangian become

$$\sum_{i\in I}\int_c^1 \lambda(\theta_i)\left[\overline{u}_i(\theta_i) - \limsup_{\theta_i'\to 1^-}\overline{u}_i(\theta_i') + \int_{\theta_i}^1\int_0^1 q_i(\theta_i', \tilde{\theta}_{-i})\,d\tilde{\theta}_{-i}\,d\theta_i'\right]d\theta_i$$

$$= \sum_{i\in I}\int_c^1 \lambda(\theta_i)\overline{u}_i(\theta_i)d\theta_i + \sum_{i\in I}\int_c^1 \lambda(\theta_i)\int_{\theta_i}^1\int_0^1 q_i(\theta_i', \tilde{\theta}_{-i})\,d\tilde{\theta}_{-i}\,d\theta_i' d\theta_i$$

$$= \sum_{i\in I}\int_c^1 \lambda(\theta_i)\overline{u}_i(\theta_i)d\theta_i + \sum_{i\in I}\int_c^1 \Lambda(\theta_i)\int_0^1 q_i(\theta_i, \tilde{\theta}_{-i})\,d\tilde{\theta}_{-i}\,d\theta_i$$

$$= \sum_{i\in I}\int_c^1\int_0^1 \Big(\big(\theta_i\lambda(\theta_i) + \Lambda(\theta_i)\big)q_i(\theta_i, \tilde{\theta}_{-i}) - \lambda(\theta_i)t_i(\theta_i, \tilde{\theta}_{-i})\Big)\,d\tilde{\theta}_{-i}\,d\theta_i,$$

where the second equality is by Fubini's Theorem.

- Due to the presence of the nonparticipation messages in $(M, q, t, \sigma)$ truthtelling being a Bayes–Nash equilibrium requires that $\overline{u}_i(\theta_i) \geq 0$ for almost all $\theta_i \in [0,1]$. We place a multiplier of $\frac{1-c}{2c}\mathbf{1}(\theta_i < c)$ on this constraint; this multiplier is nonnegative as $0 < c < 1$.

To define multipliers on the payment and allocation constraints, let $\psi = \max_{i\in I}\theta_i$.

- Due to ($\mathrm{AP}_{\geq p}$), we have $\sum_{i\in I} t_i(\theta) - p\sum_{i\in I} q_i(\theta) \geq 0$ for all type profiles $\theta \in \Theta$. We place a multiplier of

$$\kappa^t(\theta) = \begin{cases} \lambda(\psi) + 1 & \text{if } \max_{i\in I}\theta_i \geq \check{p} \\ \lambda(\check{p}) + 1 & \text{if } c \leq \max_{i\in I}\theta_i < \check{p} \\ \frac{1+c}{2c} & \text{if } \max_{i\in I}\theta_i < c \end{cases}$$

on this constraint. We have $\kappa^t(\theta) \geq 0$ as $0 < c < 1$.

- Allocative feasibility requires that $\sum_{i\in I} q_i(\theta) \leq 1$. We place a multiplier of

$$\kappa^q(\theta) = \Big(\frac{1+c}{2} - p\kappa^t(\theta)\Big)\mathbf{1}(\psi \geq \check{p}) = \Big(\frac{1-p+c}{2} - \frac{cp}{2\psi^2}\Big)\mathbf{1}(\psi \geq \check{p})$$

on this constraint. We have $\kappa^q(\theta) \geq 0$ because for $\psi \geq \check{p} > 0$, we have

$$\frac{1-p+c}{2} - \frac{cp}{2\psi^2} \geq \frac{1-p}{2} + \frac{c}{2} - \frac{cp}{2\check{p}^2} = \frac{1-p}{2} - \frac{c(p-\check{p}^2)}{2\check{p}^2} = 0.$$

[72]For interpretation, $\Lambda(\theta_i)$ gives the Lagrange multiplier on the local upward incentive constraint of type $\theta_i$. Intuitively, upward incentive constraints are relevant because ruling out payments to lower types can incentivize them to pretend to be higher types.

We then have

$$\sum_{i\in I} V_i(M,q,t,\sigma) \le \mathcal{L}(q,t), \tag{G.1}$$

where the Lagrangian $\mathcal{L}(q,t)$ is defined by

$$\begin{aligned}\mathcal{L}(q,t) = &\int_0^1\int_0^1 \sum_{i\in I}\big(\theta_i q_i(\theta) - t_i(\theta)\big)\,d\theta_1 d\theta_2 \\ &+ \sum_{i\in I}\int_c^1\int_0^1 \Big(\big(\theta_i\lambda(\theta_i)+\Lambda(\theta_i)\big)q_i(\theta_i,\tilde\theta_{-i}) - \lambda(\theta_i)t_i(\theta_i,\tilde\theta_{-i})\Big)\,d\tilde\theta_{-i}\,d\theta_i \\ &\quad + \frac{1-c}{2c}\sum_{i\in I}\int_0^c\int_0^1 \Big(\theta_i q_i(\theta_i,\tilde\theta_{-i}) - t_i(\theta_i,\tilde\theta_{-i})\Big)\,d\tilde\theta_{-i}\,d\theta_i \\ &\quad + \int_0^1\int_0^1 \Big[\kappa^t(\theta)\Big(\sum_{i\in I} t_i(\theta) - p\sum_{i\in I} q_i(\theta)\Big) + \kappa^q(\theta)\Big(1-\sum_{i\in I} q_i(\theta)\Big)\Big]d\theta_1\,d\theta_2,\end{aligned}$$

In this Lagrangian, the coefficients on $q_i(\theta)$ and $t_i(\theta)$ are

$$\begin{aligned}\mathcal{L}_{q_i(\theta)} &= \theta_i + \theta_i\lambda(\theta_i) + \Lambda(\theta_i) + \theta_i\frac{1-c}{2c}\mathbf{1}(\theta_i<c) - p\kappa^t(\theta) - \kappa^q(\theta) \\ \mathcal{L}_{t_i(\theta)} &= -1 - \lambda(\theta_i) - \frac{1-c}{2c}\mathbf{1}(\theta_i<c) + \kappa^t(\theta).\end{aligned}$$

We now show that these coefficients are nonpositive, and that $\mathcal{L}_{q_i(\theta)} = 0$ for $\theta_i \ge \check p$ and $\mathcal{L}_{t_i(\theta)} = 0$ for $\theta_i \ge \max\{\theta_{-i},\check p\}$. To show the claim for $\mathcal{L}_{q_i(\theta)}$, note that

$$\begin{aligned}\mathcal{L}_{q_i(\theta)} &= \theta_i\mathbf{1}(\theta_i\ge c) + \theta_i\lambda(\theta_i) + \Lambda(\theta_i) + \theta_i\frac{1+c}{2c}\mathbf{1}(\theta_i<c) - p\kappa^t(\theta) - \kappa^q(\theta) \\ &\le \Big(\theta_i\mathbf{1}(\theta_i\ge c) + \theta_i\lambda(\theta_i) + \Lambda(\theta_i) + \frac{1+c}{2}\mathbf{1}(\theta_i<c)\Big) - \Big(p\kappa^t(\theta)+\kappa^q(\theta)\Big) \\ &= \frac{1+c}{2} - \Big(p\frac{1+c}{2c}\mathbf{1}(\psi<c) + p\frac{c+\check p^2}{2\check p^2}\mathbf{1}(c\le\psi<\check p) + \frac{1+c}{2}\mathbf{1}(\psi\ge\check p)\Big) \\ &= \frac{c+c^2-p-pc}{2c}\mathbf{1}(\psi<c) + \frac{\check p^2+\check p^2 c - pc - p\check p^2}{2\check p^2}\mathbf{1}(c\le\psi<\check p) \le 0,\end{aligned}$$

with equality for $\theta_i \ge c$, where the first equality is by subtracting $\theta_i\mathbf{1}(\theta_i<c)$ from the first term and adding it to the third term, the second equality is by the definitions of the multipliers, the third equality is by grouping terms, and the second inequality is because $c<p$ and by the definition of $c$. To show the claim for $\mathcal{L}_{t_i(\theta)}$, note that

$$\begin{aligned}\mathcal{L}_{t_i(\theta)} &= -\big(1+\lambda(\theta_i)\big)\mathbf{1}(\theta_i\ge c) - \big(1+\lambda(c)\big)\mathbf{1}(\theta_i<c) \\ &\qquad + \Big(\big(1+\lambda(c)\big)\mathbf{1}(\psi<c) + \big(1+\lambda(\max\{\check p,\psi\})\big)\mathbf{1}(\psi\ge c)\Big) \\ &= -\big(1+\lambda(\max\{c,\theta_i\})\big)\mathbf{1}(\psi\ge c) + \big(1+\lambda(\max\{\check p,\psi\})\big)\mathbf{1}(\psi\ge c) \le 0,\end{aligned}$$

with equality for $\theta_i \geq \max\{\theta_{-i}, \check{p}\}$, where the first equality is by the definitions of the multipliers as $\lambda(\theta_i) = 0$ for $\theta_i < c$ and $\lambda(c) = \frac{1-c}{2c}$, the second equality by grouping the first three terms as $\theta_i \leq \psi$, and the inequality as $\lambda(\theta_i)$ is decreasing on $[c, 1]$ and we have $c \leq \check{p}$ and $\theta_i \leq \psi$.

Hence, since $q, t \geq 0$, using (G.1), we have

$$\sum_{i \in I} V_i(M, q, t, \sigma) \leq \mathcal{L}(q, t) \leq \int_0^1 \int_0^1 \kappa^q(\theta)\, d\theta_1\, d\theta_2.$$

On the other hand, the mechanism $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ satisfies

$$\sum_{i \in I} V_i(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma) = \mathcal{L}(q^{p\text{-DNP}}, t^{p\text{-DNP}}) = \int_0^1 \int_0^1 \kappa^q(\theta)\, d\theta_1\, d\theta_2,$$

where the first equality is due to the complementary-slackness conditions $\overline{u}_i^{p\text{-DNP}}(\theta_i) = 0$ for $\theta_i \in [0, c]$, ($\text{AP}_{\geq p}$) holding with equality for all messages, and $\sum_{i \in I} q_i^{p\text{-DNP}}(\theta) = 1$ for $\psi \geq \check{p}$; and the second equality is due to the complementary-slackness conditions $q_i^{p\text{-DNP}}(\theta) = 0$ for $\theta_i < \check{p}$, and $t_i^{p\text{-DNP}}(\theta) = 0$ for $\theta_i < \max\{\theta_{-i}, \check{p}\}$.

Therefore, for each $p \in (0, 1)$, $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ is optimal among all mechanisms $(M', q', t', \sigma')$ with $t' \geq 0$ that satisfy ($\text{AP}_{\geq p}$) and ($\text{AP}^{\hat{m}}_{=p}$) for which $\sigma'$ is a Bayes–Nash equilibrium.

*Worst-Case Expected Revenue of $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ for a Given $p$.* We claim that the worst-case expected revenue of $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ over all IC and *ex-ante* Pareto-improving collusion schemes is $p(1 - \hat{p}^2)$, where $\hat{p}$ is the unique solution in $[p, 1]$ to

$$\int_{\hat{p}}^1 (\theta - p)\theta\, d\theta = V_i\big(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma\big). \tag{G.2}$$

Here, the existence of a solution $\hat{p} \in [p, 1]$ to (G.2) follows from the Intermediate Value Theorem as ($\text{AP}_{\geq p}$) implies

$$\int_p^1 (\theta - p) 2\theta\, d\theta = \mathbb{E}_{\tilde{\psi}}\big[(\tilde{\psi} - p)\mathbf{1}(\tilde{\psi} \geq p)\big] \geq 2V_i\big(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma\big),$$

while its uniqueness follows from the left-hand side of (G.2) being strictly decreasing on $[p, 1]$. (We continue to write $\psi = \max_{i \in I} \theta_i$.)

To show the claim, first consider the collusion scheme with the deterministic message function

$$\mu_i(\theta) = \begin{cases} \theta_i & \text{if } \theta_i \geq \theta_{-i} \text{ and } \theta_i \geq \hat{p} \\ 0 & \text{otherwise} \end{cases},$$

the allocation rule $\hat{q}(\theta) = q^{p\text{-DNP}}(\mu(\theta))$, and the payment rule (similar to Example 1)

$$\hat{t}_i(\theta) = p\hat{q}_i(\theta) + \Big(\tfrac{1}{2}(\theta_i - \hat{p})^2 + \theta_i(\hat{p} - p)\Big)\mathbf{1}(\theta_i \geq \hat{p}) - \Big(\tfrac{1}{2}(\theta_{-i} - \hat{p})^2 + \theta_{-i}(\hat{p} - p)\Big)\mathbf{1}(\theta_{-i} \geq \hat{p}).$$

As in Example 1, $(\mu, \hat{q}, \hat{t})$ is IC.[73] It delivers an *ex-ante* expected utility of

$$U_i(\hat{q}, \hat{t}) = \frac{1}{2}\mathbb{E}_{\tilde{\psi}}\big[(\tilde{\psi} - p)\mathbf{1}(\tilde{\psi} \geq \hat{p})\big] = V_i\big(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma\big)$$

to each bidder, so it is *ex-ante* Pareto-improving. Expected revenue under $(\mu, \hat{q}, \hat{t})$ is $p \Pr_{\tilde{\psi}}[\tilde{\psi} \geq \hat{p}] = p(1 - \hat{p}^2)$, so we have shown that the worst-case expected revenue of $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ is at most $p(1 - \hat{p}^2)$.

It remains to show that the worst-case expected revenue of $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$ is at least $p(1 - \hat{p}^2)$. Consider any *ex-ante* Pareto-improving collusion scheme $(\mu', \hat{q}', \hat{t}')$. Let $\hat{p}' \in [0, 1]$ be such that $\Pr_{\tilde{\psi}}[\tilde{\psi} \geq \hat{p}'] = \sum_{i \in I} \mathbb{E}_{\tilde{\theta}}[\hat{q}'_i(\tilde{\theta})]$. Write $\psi = \max_{i \in I} \theta_i$ and consider the conditional expected aggregate collusive allocation

$$\hat{\bar{q}}'^{,\max}(\psi) = \mathbb{E}_{\tilde{\theta}}\Big[\sum_{i \in I} \hat{q}'_i(\tilde{\theta})\Big|\psi\Big].$$

As $x = \psi^2$ is distributed uniformly on $[0, 1]$, the function $\mathbf{1}(\sqrt{x} \geq \hat{p}')$ majorizes $\hat{\bar{q}}'^{,\max}(\sqrt{x})$. By Fan and Lorentz's (1954) inequality, we then have

$$\begin{aligned}
2\int_{\hat{p}'}^{1} (\theta_i - p)\theta_i \, d\theta_i &= \int_0^1 (\sqrt{x} - p)\mathbf{1}(\sqrt{x} \geq \hat{p}') \, dx \\
&\geq \int_0^1 (\sqrt{x} - p)\hat{\bar{q}}'^{,\max}(\sqrt{x}) \, dx \\
&= \mathbb{E}_{\tilde{\psi}}\big[(\tilde{\psi} - p)\hat{\bar{q}}'^{,\max}(\tilde{\psi})\big] \\
&= \mathbb{E}_{\tilde{\theta}}\Big[(\tilde{\psi} - p)\sum_{i \in I} \hat{q}'_i(\tilde{\theta})\Big] \geq \mathbb{E}_{\tilde{\theta}}\Big[\sum_{i \in I} \tilde{\theta}_i \hat{q}'_i(\tilde{\theta}) - p\sum_{i \in I} \hat{q}'_i(\tilde{\theta})\Big] \geq \sum_{i \in I} U_i(\hat{q}', \hat{t}'),
\end{aligned}$$

where the first two equalities are by changes of variables, the third equality is by the law of iterated expectations, the second inequality is since the bidders' total utility is maximized by allocating the good to a bidder with the highest value, and the third inequality is by ($\mathrm{AP}_{\geq p}$) and the collusion feasibility conditions (1) and (2). Since $(\mu', \hat{q}', \hat{t}')$ is *ex-ante* Pareto-improving, we then have

$$\int_{\hat{p}'}^{1} (\theta - p)\theta \, d\theta \geq V_i\big(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma\big) = \int_{\hat{p}}^{1} (\theta - p)\theta \, d\theta,$$

[73] Indeed, the allocation rule is monotone and expected payments satisfy an envelope formula.

and since $\hat{p} \in [p, 1]$, it follows that $\hat{p}' \leq \hat{p}$. As ($\mathrm{AP}_{\geq p}$) holds with equality for $(M, q^{p\text{-DNP}}, t^{p\text{-DNP}}, \sigma)$, expected revenue under $(\mu', \hat{q}', \hat{t}')$ is

$$p \sum_{i \in I} \mathbb{E}_{\tilde{\theta}}\left[\hat{q}_i'(\tilde{\theta})\right] = p \cdot \Pr_{\tilde{\psi}}\left[\tilde{\psi} \geq \hat{p}'\right] \geq p \cdot \Pr_{\tilde{\psi}}\left[\tilde{\psi} \geq \hat{p}\right] = p(1 - \hat{p}^2).$$

*Optimization Over $p$.* In particular, the seller must choose $p$ to maximize $p(1 - \hat{p}^2)$, where $\hat{p}$ is defined implicitly (as a function of $p$) by (G.2). Numerical optimization yields an optimal choice of $p \approx 0.63$, which yields $\hat{p} \approx 0.69$ and hence $p(1 - \hat{p}^2) \approx 0.33$.

## References


Border, K. C. (1991). Implementation of reduced form auctions: A geometric approach. *Econometrica 59*(4), 1175–1187.

Börgers, T. (2015). *An Introduction to the Theory of Mechanism Design*. Oxford University Press.

Csóka, E., A. Pongrácz, and A. Rodivilov (2025). Guaranteed utility equilibrium. Working paper.

Dworczak, P. and E. V. Muir (2025). A mechanism design approach to property rights. Working paper.

Edlin, A. S. and C. Shannon (1998). Strict monotonicity in comparative statics. *Journal of Economic Theory 81*(1), 201–219.

Fan, K. and G. G. Lorentz (1954). An integral inequality. *American Mathematical Monthly 61*(9), 626–631.

Jagadeesan, R., I. Segal, and A. Tordjman (2026). Robustly efficient implementation under collusion: Holdout threats and the anti-core. Working paper.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory 93*(1), 1–47.

Liu, B. (2025). The optimal design of countervailing incentives. Working paper.

Manelli, A. M. and D. R. Vincent (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory 127*(1), 1–35.

McAfee, R. P. and J. McMillan (1987). Auctions and bidding. *Journal of Economic Literature 25*(2), 699–738.

Milgrom, P. and I. Segal (2002). Envelope theorems for arbitrary choice sets. *Econometrica 70*(2), 583–601.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press.